The Recency Bias in the Attribution of Responsibility for Joint Work

Andrzej Baranski[†] Manu Muñoz-Herrera[‡]

October 20, 2022

Abstract

In collaborative production process it is often the case that individuals contribute sequentially to achieve a common goal. Does the order of contributions affect perceptions of credit and responsibility for the outcome of the group? We design an experiment in which teams of three subjects collaborate sequentially in a building task with the objective of finishing within a given time limit. Uninvolved reviewers evaluate the builders' performance and decide which builder to hire (our incentivized measure of responsibility attribution). We find robust evidence of a recency bias for blame attribution: Hiring rates of final builders are substantially lower when the team fails. However, there is no difference in hiring rates between first and last builders for successful teams. We control for perceived task difficulty, objective and subjective individual contributions, and attention, none of which explain or dampen the effect. Implications for teamwork, organizational design, and managerial practices of team performance evaluation are discussed.

Keywords: credit attribution, attribution bias, joint production, laboratory experiment

We would like to thank Laila Al-Eisawi, Marielle Caballero, Amr Yakout, and Kwabena Sekyi-Djan for their excellent research assistance. Alex Gotthard-Real, Moritz Janas and Robert Stuber offered very valuable suggestions. The authors are grateful for financial support from the Center for Behavioral Institutional Design and Tamkeen under the NYUAD Research Institute award for Project CG005.

[†]Division of Social Science and Center for Behavioral Institutional Design, New York University Abu Dhabi. *Email*: a.baranski@nyu.edu

[‡]Division of Social Science and Center for Behavioral Institutional Design, New York University Abu Dhabi. *Email*: manumunoz@nyu.edu

1 Introduction

In many collaborative production processes between partners, employees or team members, contributions often occur in a sequential manner. Various departments within a firm are typically involved in the creation of a product or provision of a service. For shareholders and management, it is crucial to determine the extent to which each contributing individual is responsible for the collective outcome of the joint production process (Holmstrom, 1982; Milgrom and Roberts, 1992; Lazear and Shaw, 2007).[1] *Proper* attribution of responsibility to individuals collaborating in a task is especially important within firms, as promotions and bonuses are likely to be awarded to employees whose contributions are most recognized, while those deemed responsible for negative outcomes can be penalized, demoted, or fired (Lazear and Gibbs, 2014; Villeval, 2022).[2] As such, failing to *properly* attribute credit for joint work may lead to feelings of inequity (Adams, 1963; Carrell and Dittrich, 1978), which can be deleterious to team or firm performance (Bradler et al., 2016).[3]

In this paper, we study whether the order in which individuals contribute to a sequential production task impacts how their contributions are perceived and recognized by an evaluator (e.g., a manager). Are those who contribute early in the production process deemed to be *more causal* (i.e. primacy bias) or do last movers receive more credit for the outcome (i.e. recency bias)? Furthermore, does success or failure alter who is credited or blamed for the outcome?

While the questions we pose in this study are simple, answering them empirically based on observational data proves elusive. Most joint-production settings that one may envision are beset with confounding factors that prevent a clear comparison between the tasks within a team or firm. Usually, tasks performed by individuals may differ in their complexity or perceived relevance. Employees are likely to differ in their effort provision, the quality of their contributions, and their status (i.e. seniority, power, etc). Evaluators' judgements may be *motivated*, leading to favoritism or discrimination in their assessments. Consequently, existing data on employee evaluations may not allow us to cleanly identify the existence of recency or primacy biases in credit attribution. In this article, we design an experiment to control for the aforementioned factors and isolate the temporal order of contributions.

In our study, uninvolved evaluators (hereafter referred to as *managers*) are shown a video of a team of three other subjects who collaborate sequentially in a building task (hereafter referred to as *builders*). Managers were exposed to one of two conditions: Successful teams in which builders completed their task within a time limit, and unsuccessful teams who did not. Upon observing the builders, managers were asked who they would hire to perform the same task for them, a response that we incentivized by

[1] By joint production, we refer to a situation in which the outcome of a group of people depends on each individual member's effort or contribution. The production technology, or way in which inputs are aggregated to yield an outcome, is not of immediate concern here. Our focus will be on the timing of individual contributions.

[2] Even in the absence of financial incentives, people typically care about being adequately recognized in relation to their peers. For example, Kosfeld and Neckermann (2011) and Bradler et al. (2016) find that managers giving a purely symbolic award (a thank-you card) to good workers, elicit more effort from recipients.

[3] See Sarsons et al. (2021) for evidence of how gender stereotypes in attribution of credit for joint work negatively affect female academics' tenure promotion chances. See Dalton and Landry (2020) for how a recent performance of an NBA player can have an outsized effect on likelihood of being traded compared to the effect of more informative historical data.

compensating managers for selecting high ability builders.

We find strong evidence of a *recency bias* in the attribution of blame for failed outcomes, as the last builder is 3 times less likely to be hired than the first builder (14% vs 47%). There is no significant difference in hiring rates between first and last builders for successful teams (36% vs 40%), which highlights that the recency bias is outcome dependent in our setting. We are able to cleanly rule out multiple confounds such as objective and perceived contribution, task difficulty and limits in memory and recall from managers. Our findings are further substantiated by subjects' responses (non-incentivized) when asked who was responsible for the group outcome.

In order to address important aspects that are central to managerial settings, we conducted two variants of our main experiment. First, we considered a treatment in which managers were paid a bonus contingent on the outcome of the team they evaluated (hereafter, *Involved* treatment). This is motivated by the observation that manager's payoffs may be contingent on the outcome of their team, which in turn can trigger an emotional response that exacerbates biases. Second, managers' decisions typically affect workers, which may trigger more thoughtful or careful considerations if managers are have equity concerns (Adams, 1963). To this end, we designed a treatment in which hired builders were rewarded monetarily (hereafter, *Rewards* treatment). Consistent with our main results, we find a recency bias in blame attribution for failed outcomes and no significant order effect for successful outcomes, but there are qualitative differences in line with our expectations. Last builders in failed teams are marginally less likely to be hired in the *Involved* treatment and marginally more likely in the *Rewards* treatment. These results highlight that causal attribution judgments can be affected by experienced disappointment when the outcome has material consequences over the evaluator and that equity concerns may mildly temper the recency bias.

The remainder of this article proceeds as follows. In Section 2 we discuss the existing related work. Next, in Section 3 we present the experimental design followed by the results in Section 4. The results from the *Involved* and *Rewards* treatments are presented in Section 5. We discuss our results and conclude in Section 6. Robustness treatments and analyses are in the Online Appendix.

2 Previous Work

In the field of cognitive psychology, there is a voluminous literature studying how the order in which information is presented to people affects their judgements. In these studies, subjects express their opinion about a specific proposition after being presented with several pieces of information, which may arrive sequentially. We refer the reader to the meta-analysis by Hogarth and Einhorn (1992), which documents evidence both of primacy (over-reliance on initial evidence or information) when the information is *simple* and recency (over-reliance on late information) when information is complex. We consider our experimental building task to be very simple, which would trigger primacy according to Hogarth and Einhorn's conceptual framework, but this is not what we find.

Closer to our experiment, several recent studies have focused on order effects in determining *causal* *judgements*. As defined by Reuter et al. (2014), "[c]ausal selection is the cognitive process through which one or more elements in a complex causal structure are singled out as actual causes of a certain effect" (p. 1). Spellman (1997) argues that causality is typically attributed to the event which raises the probability of occurrence of an outcome the most, given what has happened earlier. Vignette experiments by Henne et al. (2021) reveal that events that change the probability of the outcome the most, are deemed more causal. In the context of a collective decision-making problem Bartling et al. (2015) find that pivotality affects responsibility attribution in sequential voting, with pivotal voters being blamed more for an unpopular outcome than non-pivotal voters by recipients of a collective decision. Anselm et al. (2022) find that pivotal voters are rewarded more often for supporting a fair outcome. Supporters after the pivotal voter are blamed and rewarded less, thus there is no recency or primacy.

There is a growing interest in understanding credit attribution for joint work in teams (Lazear and Shaw, 2007; Isaksson, 2018), especially in academic research collaborations (see Shen and Barabási (2014) for a review and Sarsons et al. (2021) for gender differences). Survey evidence suggests there are self-serving biases in credit claim (Herz et al., 2020), but there is no objective way of establishing percentage shares of contribution.[4] Importantly, none of these studies focus on the temporal order of contributions to a team or firm as we do here.

Finally, experiments by Cappelen et al. (2017) and Almås et al. (2020) are examples of studies employing uninvolved parties that are informed about the outcomes of an individual production task. Workers in these experiments participate in real-effort tasks and the value of their production is partially determined by luck. Third parties make distributional decisions which affect the worker's payoffs. Besides the motivating question, there are several methodological differences in our settings. First, observers' decisions do not affect the workers' payoffs in our main experiments. We intentionally do so to turn off the effect that fairness concerns regarding payoff distribution may have. Second, observers witness the production process visually as opposed to being merely informed about it.[5]

3 Experimental Design

Our experiment consisted of two stages: a sequential building real-effort task and an assessment of performance by uninvolved third parties, which leads to hiring decisions.

Stage 1: Sequential building task

Builders were invited to the SSEL laboratory at NYU Abu Dhabi to assemble a Lego® set individually. Builders followed a manual in order to put together 40% of the pieces in Part 1, 20% in Part 2 and the remaining 40% in Part 3.[6] The building process was recorded on video, with the camera focusing only

[4] Kornhaber et al. (2015) offers a review of articles discussing the challenges of co-authorship in the biomedical sciences.

[5] See Khubulashvili et al. (2021) for a discussion on the effects of observing versus imagining someone's contribution.

[6] One may argue that first and last builders faced different difficulty levels because first builders had more pieces at their disposal from where to search and then assemble. We further ensured that this was not the case by separating the pieces in bags for each part of the set. We also control for the perceived difficulty by eliciting the managers' perceptions (see Instructions in Appendix A).

on the hands of a builder putting together the Lego® pieces. For each participant three videos were recorded, one of each part of the process. In addition to a show-up fee, a bonus was paid to builders who completed the task in less than 10 minutes.

We then combined recordings to produce videos resembling a sequential team production line. Each team-production video shows a sequence of three builders completing the set, each one taking off from where the previous one had ended. At the end of every video there is a 15-second effect highlighting the team's success or failure in completing the building task.

The goal of our videos was two show a realistic but easily comprehensible task for third-party observers. To reduce distraction and increase quality of our responses we shortened the videos so that only the moments when participants were putting pieces together were visible. This means that every step of the building process is observable by the evaluators. We also sought to minimize differences between builders so that it would be difficult to identify *the true ability* of each builder. Thus, the duration of each part is not informative of the actual performance of each builder.

Stage 2: Hiring

We invited an online sample of third-party evaluators (which we refer to as *managers*) who were payed a fixed fee of $2 for watching a video clip and answering a few questions about the content of the video.[7] Subjects that did not answer attention checks correctly could not proceed with the experiment. The task was described to managers as an evaluation of a team that had built a Lego® set together and whose success depended on how long it took them to finish the task. Half of the managers in each treatment where randomly assigned to a condition in which they watched a video where the team failed, the other half watched a successful team.[8]

Upon watching their randomly assigned video, managers were asked which builder they would want to *hire*. They were also informed that each builder had assembled the entire set individually and that some had succeeded and others failed to complete it in time. If managers hired a builder who successfully built the entire set in the individual session, they received a bonus payment of $2.

We also asked managers directly which builder they considered to be the most responsible for the team's outcome. This question was not incentivized as it did not affect managers' payoffs. Our conjecture was that the answers to this question and the hiring decision would be highly positively (negatively) correlated in success (failure) conditions.[9]

To control for the perception of contributions to the building task, we asked managers to estimate the number of pieces each builder had placed. Those who were within 5 pieces of the correct amount would earn a bonus payment of $1. We also asked them which part of the assembly process did they find was more difficult to build (unincentivized). Both of these measures allow us to investigate the

[7] In total, we recruited 1600 managers via Prolific, 400 per treatment. Below we describe the Baseline condition and results.

[8] In total, we created 4 videos, two for each condition. In our analysis, we control for which video was shown to the managers.

[9] Part of our methodological inquiry was to evaluate if incentivizing the causal attribution questions could affect subjects' responses systematically compared to stated judgments. We find a high correlation and no systematic differences: causal attribution is positively correlated with hiring in success conditions and negatively in failure conditions.

determinants of hiring decisions. Accuracy of answers on pieces assembled also serves a measure of attention.

4 Results

4.1 Hiring Decisions

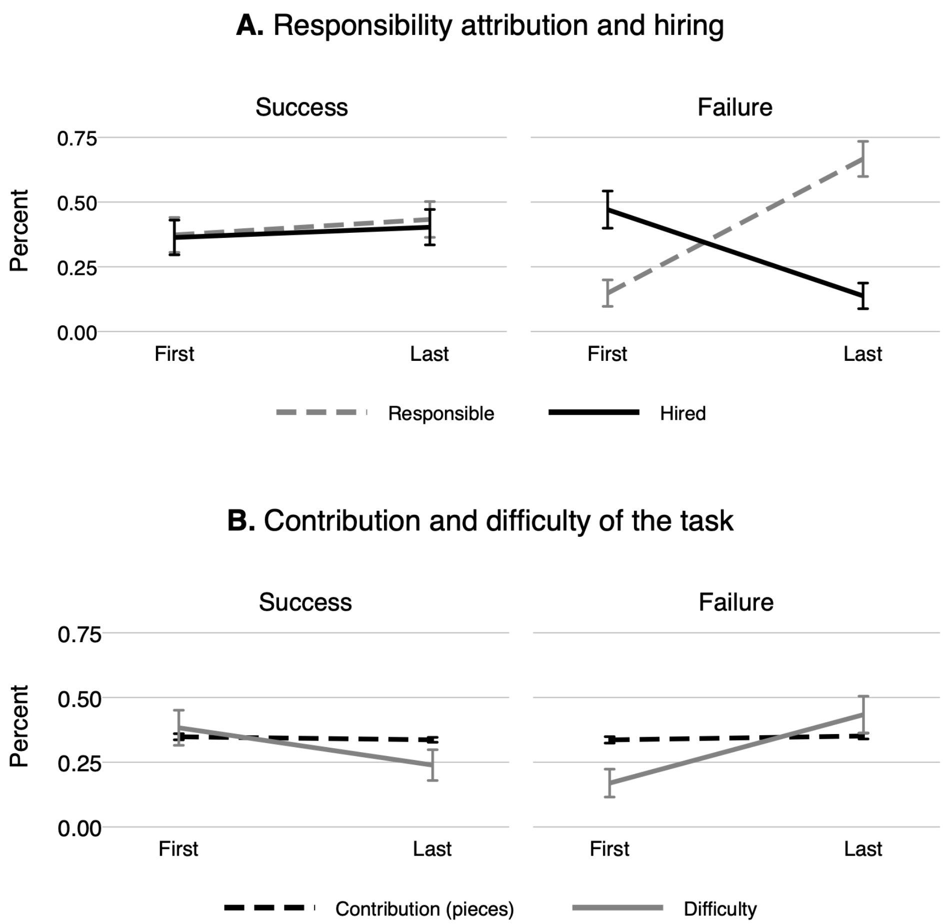

Figure 1: **Main Experimental Outcomes.** *Panel A:* The solid line (in both panels) displays the fraction of managers who hired the worker in either the First or Last position. The dashed line displays the fraction of managers who made the worker in the first or last position responsible for the success (left panel) or failure (right panel) of the outcome. *Panel B:* The dashed line displays the perceived contribution by the manager of each player to the task, expressed as the share of total pieces put together. The solid line displays the perceived difficulty by the managers of the task each player was given according to their position.

Figure 1.A shows the proportion of times each builder gets hired by the manager, both when teams

were successful (left) or failed (right). When teams succeed, first builders get hired 36% of the time while last builders 40% of the time (p=0.412, Chi-squared test). When teams fail the difference becomes significant: first builders get hired 47% of the time while last builders 14% of the time (p<0.001, Chi-squared test). These findings provide direct evidence for the existence of a recency bias in blame attribution when hiring: Last movers are disproportionately penalized for failure and not evidently prized for success.[10] The differential effect of the outcome on hiring probability of each builder is notable: failure leads to a larger hiring gap between the first and third builder (p<0.001).[11]

Repeating the same analysis as before, we find that managers are less likely to blame the first builder for a team's failure than the they are to blame the last mover (15% vs 67%, p<0.001, Chi-squared test). However, as illustrated in Figure 1.A, managers appear to be equally likely to attribute responsibility between the first and last movers when teams succeed (37% vs 43%, p=0.222, Chi-squared test), as illustrated in Figure 1.A. Moreover, when looking at the relation between hiring and attributing responsibility (credit or blame), we find a strong correlation between the hiring decision and credit (in success conditions) and no-hiring and blame (in failure conditions). The correlation coefficients are 0.500 (p<0.001) and -0.419 (p<0.001), respectively.[12]

Result 1. *There is a recency bias in the attribution of responsibility for failed teams but not for successful ones. The last builder in the sequence of a failed team is blamed more often and hired less often than the first builder, but there is no difference for first and last movers of successful teams.*

4.1.1 Controlling for perceived contributions and difficulty

In our experiment, we have aimed to control for each builder's *objective* contribution as well as to maintain the task difficulty homogeneous across stages. We now ask if managers effectively believe that contributions were equal between the first and last builder and if perceptions of difficulty vary within building stages or success conditions. We also investigate if managers' perceptions affect their hiring decisions and if they may explain the previously documented hiring bias.

With respect to our first question, it is natural to conjecture that, if last builders are blamed more for failures, it might be because managers also underestimate their material contribution to the building task. But this is not the case. There is no difference in the proportion of the pieces that managers believe first and last movers assembled (26% vs 25% for success, p=0.970; 24% vs 25% for failure, p=0.480).

We asked managers which builder had the most difficult part, with the option of stating that all were equally difficult. Figure 1.B shows that when teams are successful the first part is deemed the most difficult 38% of the time and the third part 24%, and the difference is significant (p=0.002). When

[10] The true performance of each builder (whether he builds the entire Lego set on time or not) is not significantly correlated with hiring decisions. The correlation coefficient is 0.018 (p=0.535).

[11] P-value obtained from a linear regression on the probability of hiring the last builder.

[12] This suggests that methodologically the incentivized measures, such as the hiring decision developed here, provide similar insights regarding credit attribution compared to stated judgements in non-incentivized questions. Thus, researchers seeking to study this phenomena in other settings where incentivizing responses can prove difficult may rely confidently on stated attributions of responsibility.

teams are unsuccessful, the first part is perceived as the most difficult 17% of the time and the third part 43% ($p<0.001$). Across outcomes, 28% of managers state that all parts were equally difficult.

	Success		**Failure**	
	I	**II**	**III**	**IV**
Builder 3	0.040	0.079	−0.333***	−0.350***
	(0.062)	(0.059)	(0.052)	(0.052)
Contribution		1.400***		0.850***
		(0.274)		(0.284)
Difficulty		0.117*		0.008
		(0.059)		(0.052)
Constant	0.363***	−0.173	0.471***	0.186
	(0.034)	(0.94)	(0.036)	(0.101)
# Obs.	402	400	378	376
# Groups	201	200	189	199
AIC	564	533	436	428

Table 1: **Linear Regressions for Hiring.** Linear regressions with standard errors clustered on individual managers (in parenthesis). The dependent variable is probability of being hired by the manager in columns **I** and **II** for successful outcomes and in columns **III** and **IV** for failed outcomes. To test for recency/primacy biases, we focus only on managers hiring either the first or the last builder. ***, ** and * indicate statistical significance at the 0.001, 0.01 and 0.05 levels.

How do managers' perceptions of pieces contributed and task difficulty correlate with their hiring decisions and attribution of responsibility? And furthermore, once we account for these perceptions, does the recency bias remain? Table 1 presents the results from linear regressions for each outcome variable in both success and failure conditions. Our results clearly indicate that perceived contributions are a key driver of hiring decisions and that perceived difficulty has no impact. Importantly, the recency bias in failed outcomes remains.

The results presented in this subsection help us rule out that managers' poor or selective memory can be driving our results. We also corroborate that subjects are trying to make objective assessments in hiring the highest contributor, and not just guessing. Furthermore, we checked if subjects accuracy in assessing the total number of pieces that builder's *contributed* differed by condition, but this is not the case. The euclidean distance between the actual and reported values is 22.5 pieces on average and not significantly different between success or failure conditions ($p=0.522$, obtained from a linear regression). In Appendix B, we replicate the analysis on hiring decisions, splitting managers into those whose accuracy on the actual contributions of the builders (number of pieces) was below and those above the median. Our findings remain for both groups of managers.

Result 2. *On average, managers accurately perceive that the first and last builder contribute equally to the task, meaning there are no order effects in assessing material contributions to the task. While contribution perceptions correlate positively with hiring decisions, these do not explain the recency bias in hiring.*

One possible reason why last builders are penalized is because they may be *more salient* than first builders. Ruling out the role of the last mover's salience is important because in many situations, such as sports competitions, the final mover typically attracts more attention. To test for the robustness of our findings, we decreased salience of the groups outcome by *toning-down* the announcement that a group had failed or succeeded in finishing the building task on time (using the same videos as before). The results are virtually unchanged. See Appendix D.

5 Follow-up Treatments

In this section, we present the results of additional treatments aimed at capturing two central aspects of managerial decision making: that the outcome of the team can affect managers' payoffs, and the their evaluations can affect workers earnings.

In the *Involved* treatment, managers' payoffs varied depending on the outcome of the team they evaluated. Every aspect of the experiment remained the same, except that managers were endowed with $1 at the beginning of the experiment. They were informed that, if the team they were assigned to evaluate was successful, they would earn an additional $1. Otherwise, they would suffer a cost of $1. Because failed teams entail a monetary loss for managers, one reasonable conjecture is that this can trigger a stronger attribution of blame.

Figure 2 depicts the hiring rates across all treatments. The data show that managers in treatment *Involved* display a recency bias in blame attribution when teams fail (as in our baseline treatment). Importantly, managers are less likely to hire the last builder in treatment *Involved* relative to the baseline albeit the difference is not significant in either success (3.69%, p=0.434, Chi-squared test) or failure (4%, p=0.217, Chi-squared test).

We also conducted a treatment in which managers' hiring decisions could potentially carry rewards for workers. Managers were told that there was a 10% probability that we would pay $1 to the worker they hired. We argue that this variation activates the possibility of manifesting and acting upon equity concerns to properly reward those deemed more responsible for a team's success, and punish those deemed more responsible for a team's failure. In REWARDS, managers are more likely to hire the last builder relative to the baseline in failure, although the difference is not significant in either failure (5.87%, p=0.116, Chi-squared test) or success (3.69%, p=0.458, Chi-squared test).

Result 3. *The recency bias in credit attribution, in which last builders are hired less often than first builders when teams fail, is present when payoffs of the managers are affected by the team's outcome and also when manager's decisions carry rewards for workers.*

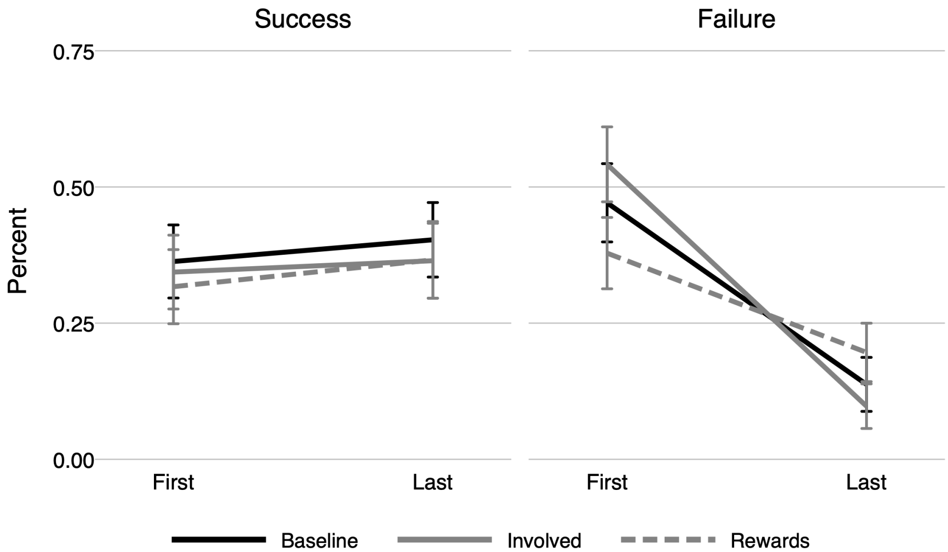

Figure 2: **Hiring decisions in Additional Treatments.** Each line displays the fraction of managers who hired the worker in either the first or last position in treatments Baseline (dark solid line), Involved (light solid line) and Rewards (dashed line). Hiring decisions are displayed separately for successful outcomes (left) and failed outcomes (right).

6 Discussion and Concluding Remarks

In this study, we present evidence of how the order of contributions to a team task affects blame attribution for the outcome. The last contributor to a sequential task is judged to be responsible more often when the team being evaluated has failed to achieve its objective. Third-parties judgements of attribution of responsibility correlate strongly with their hiring decision, a choice that carries material consequences in our study. Thus, we not only uncover a bias in judgement, but find that it can have deleterious material consequences.

The strength of the recency bias in the attribution for failure is strong and stable in subsequent treatments that mimic important conditions in managerial settings. We explored two additional conditions which enhance the external validity of our findings: When managers' payoffs depend on the outcome of the team they evaluate and when managers' hiring decisions carry monetary rewards for workers. There is an indication that last builders are more likely to be blamed for failed outcomes when the manager's payoffs are affected. It is likely that disappointment triggers an emotional response that heightens the recency bias.

In the treatment in which hiring decisions carry potential rewards for builders, we find a moderately lower blame on last builders of failed teams. One potential reason for the tempering of the bias is that evaluators may be concerned with fairness in terms of equity, that is, rewarding justly for work done. Because subjects report no differences in pieces assembled between first and last builders, they are likely to view the workers as equally deserving of the monetary reward and this mildly dampens the impulse to not hire the last builder in failed teams.

The evidence also shows that causal attribution judgements vary with the valence of the outcome (i.e. failure and success). While context is know to matter in decision-making (i.e. gains frame versus loss frame), this has not been documented in causal attribution judgements and calls for further research.

Our findings have important implications for management practices. Many business endeavors start with idea generation and planning taking place at high levels within a firm's hierarchy. If high ranking executives are deemed the initiators of projects but other employees down the hierarchy contribute at later stages in their execution, it may happen that failed endeavors are less likely to be blamed on the initiators. As such, if a rencecy bias for failed outcomes is at play, it can shield management and high ranking executive from blame which can lead to firing or sanctioning the *wrong* culprits.

In general, those in charge of evaluating employee performance in team-based firms should account for a potential bias in the attribution of responsibility, which is expected to be more pronounced in low performance teams. Failing to identify the *right* culprits for failure can lead to unjust and inefficient allocation of resources within organizations.

Importantly, we rule out several possible mechanisms that could potentially explain the bias. First, the bias does not arise due to imperfect recall or impaired memory because we find no differences in perception of pieces contributed between the first and last mover. Second, the bias does not operate through selective recall dependent on the outcome because evaluators' subjective assessments of contributions do not differ between success and failure conditions. Finally, the results from our robustness treatment (Online Appendix D) show no difference in behavior when the salience of the final builder is decreased.

Besides the novel behavioral patterns we uncover, our contribution to the existent literature is also methodological. We designed a task that controls objectively for material contributions to the team task and keeps difficulty relatively unchanged between builders. This allows to isolate the effect of the order of contributions transparently. Importantly, our incentivized hiring decision avoids any confounds that may arise due to social preferences for fairness or concerns for efficiency. This is because managers' decisions do not affect the builders. Furthermore, due to the one-shot nature of the experiment, managers' decisions do not provide builders any feedback or incentive to perform better in the future. Importantly, the strong correlation between hiring and judgments of responsibility provides a validation for the use of survey and vignette elicitation methods when providing incentives may prove difficult.

Several questions remain to be answered which may enhance our understanding of the recency bias that we have uncovered. Can workers anticipate this judgment pattern by managers? And if given a choice, which place in a production sequence would workers or teammates place themselves in? A second area to explore concerns complementarities in production, which are central to teams and firms. Note that in our task, the first builders' efforts have no direct impact on the productivity of the subsequent builders. It remains to be studied if causal judgments can be affected by the presence of synergies in joint production. We leave these and other important questions for future research.

References

Adams, J. S. (1963). Towards an understanding of inequity. *The Journal of Abnormal and Social Psychology*, 67(5):422.

Almås, I., Cappelen, A. W., and Tungodden, B. (2020). Cutthroat capitalism versus cuddly socialism: Are americans more meritocratic and efficiency-seeking than scandinavians? *Journal of Political Economy*, 128(5):1753–1788.

Anselm, R., Bhatia, D., Fischbacher, U., and Hausfeld, J. (2022). Blame and praise: Responsibility attribution patterns in decision chains. Thurgau Institute of Economics and Department of Economics at University of Konstanz Research Paper Series available at: https://www.twi-kreuzlingen.ch/research/workingpaper/. Last accessed October 9, 2022.

Bartling, B., Fischbacher, U., and Schudy, S. (2015). Pivotality and responsibility attribution in sequential voting. *Journal of Public Economics*, 128:133–139.

Bradler, C., Dur, R., Neckermann, S., and Non, A. (2016). Employee recognition and performance: A field experiment. *Management Science*, 62(11):3085–3099.

Cappelen, A. W., Moene, K. O., Skjelbred, S.-E., and Tungodden, B. (2017). The merit primacy effect. NHH Dept. of Economics Discussion Paper available at SSRN: https://ssrn.com/abstract=2963504. Last accessed 19 June 2022.

Carrell, M. R. and Dittrich, J. E. (1978). Equity theory: The recent literature, methodological considerations, and new directions. *Academy of management review*, 3(2):202–210.

Dalton, M. and Landry, P. (2020). 'overattention'to first-hand experience in hiring decisions: Evidence from professional basketball. *Journal of Economic Behavior & Organization*, 175:98–113.

Henne, P., Kulesza, A., Perez, K., and Houcek, A. (2021). Counterfactual thinking and recency effects in causal judgment. *Cognition*, 212:104708.

Herz, N., Dan, O., Censor, N., and Bar-Haim, Y. (2020). Authors overestimate their contribution to scientific work, demonstrating a strong bias. *Proceedings of the National Academy of Sciences*, 117(12):6282–6285.

Hogarth, R. M. and Einhorn, H. J. (1992). Order effects in belief updating: The belief-adjustment model. *Cognitive psychology*, 24(1):1–55.

Holmstrom, B. (1982). Moral hazard in teams. *The Bell journal of economics*, pages 324–340.

Isaksson, S. (2018). It takes two: Gender differences in group work. *Norwegian School of Economics Working Paper*.

Khubulashvili, R., Linardi, S., and Wang, X. (2021). The role of effort cost perception in outcome bias. *Working paper*.

Kornhaber, R. A., McLean, L. M., and Baber, R. J. (2015). Ongoing ethical issues concerning authorship in biomedical journals: an integrative review. *International journal of nanomedicine*, 10:4837.

Kosfeld, M. and Neckermann, S. (2011). Getting more work for nothing? symbolic awards and worker performance. *American Economic Journal: Microeconomics*, 3(3):86–99.

Lazear, E. P. and Gibbs, M. (2014). *Personnel economics in practice*. John Wiley & Sons.

Lazear, E. P. and Shaw, K. L. (2007). Personnel economics: The economist's view of human resources. *Journal of economic perspectives*, 21(4):91–114.

Milgrom, P. and Roberts, J. (1992). *Economics, organization and management*. Prentice-Hall Englewood Cliffs, NJ.

Reuter, K., Kirfel, L., Van Riel, R., and Barlassina, L. (2014). The good, the bad, and the timely: how temporal order and moral judgment influence causal selection. *Frontiers in psychology*, 5:1336.

Sarsons, H., Gërxhani, K., Reuben, E., and Schram, A. (2021). Gender differences in recognition for group work. *Journal of Political Economy*, 129(1):101–147.

Shen, H.-W. and Barabási, A.-L. (2014). Collective credit allocation in science. *Proceedings of the National Academy of Sciences*, 111(34):12325–12330.

Spellman, B. A. (1997). Crediting causality. *Journal of Experimental Psychology: General*, 126(4):323.

Villeval, M. C. (2022). The cognitive and motivational effects of performance feedback. In Erikson, T., editor, *Encyclopedia of Labor Studies*, page Edward Elgar Publishers.

Online Appendix:

The Recency Bias in the Attribution of Responsibility for Joint Work

A Instructions

Below we present the instructions for the experiment. We will indicate when a question applies to a specific treatment. For example, **[Success]** when the manager observes a team who has successfully completed the task, and **[Failure]** in the other case. Also, we will indicate if the answer options for a question were displayed in random order by using *[*randomized order]*.

WELCOME

You are participating in a study on economic decision-making. Typically, the study takes about **8 minutes** to complete.

For completing the study, you will receive $2.00. In addition, you will be able to earn **bonus payments** of up to $3.00 more. You will be paid only if you complete the entire study.

The study is anonymous. Hence, your identity will not be revealed to others and the identity of others will not be revealed to you.

Next, you will see the instructions. **Please read the instructions carefully as they describe how your earnings are determined.**

The main task is to carefully watch a short video. You will be asked questions to confirm that you have watch it with attention. **If you answer these questions incorrectly, you will be excluded from the study and you won't be eligible for payment.**

By continuing to the next screen, you consent to participate in this study. For more details about your consent, click on "See consent form".

- See consent form

[page break]

We asked participants in a different study to build parts of a Lego set. We then grouped these participants into teams, and gave a bonus to those teams for which the total time of building the Lego set was at most 3 minutes.

The following video shows you a team of participants building the Lego set. Please watch the video carefully, as we will ask you a few attention questions about the task the participants are performing.

[page break]

Please watch carefully the following video.

Note: Please pay close attention to the task participants are performing. Skipping this stage or leaving before the video ends will make you ineligible for payment.

*[*The video is displayed in this question]*

[page break]

Did the participants in the team earn the bonus by successfully building the Lego set in time?

- Yes
- No

[page break]

How many participants composed the team (built the lego set) in the video you watched? (*Hint: the number of participants is equal to the number of parts in the video*)

[options from 0 to 10]

[page break]

Please indicate how many pieces you think **each participant** put together. If your answer is within 5 units of the correct amount of pieces for each participant, you will receive a bonus of **1 USD**.

- Participant 1
- Participant 2
- Participant 3

[page break]

Please indicate **how difficult** you think the different parts of building the Lego set were *[*randomized order]*

- All three parts were equally difficult
- The first part was the most difficult
- The second part was the most difficult
- The third part was the most difficult

[page break]

[Success]

Which of the three participants in the team was **the most responsible for building** the lego set in less than three minutes *[*randomized order]*

- Participant 1
- Participant 2

- Participant 3

[page break]

Please tell us why did you say *[Chosen participant]* was the most responsible for building the Lego set in time to earn the bonus

[page break]

[Failure]

Which of the three participants in the team was **the most responsible for not building** the lego set in less than three minutes *[*randomized order]*

- Participant 1
- Participant 2
- Participant 3

[page break]

Please tell us why did you say *[Chosen participant]* was the most responsible for not building the Lego set in time to earn the bonus

[page break]

In the video you watched, each participant built one of the three parts of the Lego set. We also asked each participant to build the other two parts, meaning that each of them built the entire Lego set by him or herself.

You have the opportunity to **hire one of the three participants** in the video. If the person you hire has built the entire Lego set (the three parts) in time for him/her to earn the bonus, you will earn an additional **bonus of 2 USD**. Please indicate which of the three participant you choose to hire.

- Participant 1
- Participant 2
- Participant 3

[page break]

Please tell us why did you hire *[Hired participant]*

[page break]

Please indicate what was the gender of each of the participants

- Participant 1: Male / Female
- Participant 2: Male / Female
- Participant 3: Male / Female

[page break]

Before concluding, we would like you to answer a final set of questions.

What is your gender?

- Male
- Female
- Other

[page break]

What is your race / ethnicity? Select all that apply.

- White
- Black
- Latino
- Asian
- Native American
- Other

[page break]

How old are you (in years)?

[page break]

What is your highest educational degree?

- Less than high school
- High school graduate

- 2 year college/university degree
- 4 year college/university degree
- Masters degree or equivalent
- Doctorate or equivalent

[page break]

Are you currently employed?

- Yes
- No

[page break]

Are you currently a student?

- Yes
- No

[page break]

In which state do you currently reside?

- Alabama
- ...
- I do not reside in the United States

[page break]

Did you vote in the last election?

- Yes
- No

[page break]

Generally speaking, do you usually think of yourself as a:

- Republican

- Democrat
- Libertarian
- Independent
- Other

[page break]

Here is a 7-point scale on which political views that people might hold are arranged from extremely liberal (left) to extremely conservative (right). Where would you place yourself on this scale?

[page break]

You have reached the end of the study.

Please do not discuss the procedures or content of this study with other participants.

Your bonus payment

In the next 48 hours we will review your responses to the questions that had a bonus payment, and determine whether you earn the **bonus payment**. To receive your payment you must submit your prolific ID on the next page.

B Baseline treatment supporting analysis

We replicate the analysis from hiring for responsibility attribution. Table B1 reports linear probability models testing the decisions to attribute responsibility over the success or failure of the group's outcome. In all regressions the independent variables are dummies for the final builders, so that the first builder is the omitted category. Moreover, we use individual-level random effects in all regressions to cluster decisions at the manager's level. The dependent variable is the probability of assigning responsibility to a builder, for the case of Success in columns **I-II** and Failure in columns **III-IV**.

	Success		**Failure**	
	I	**II**	**III**	**IV**
Builder 3	0.060	0.106	0.519***	0.514***
	(0.063)	(0.058)	(0.054)	(0.056)
Contribution		1.520***		−0.091
		(0.275)		(0.283)
Difficulty		0.226***		0.053
		(0.059)		(0.060)
Constant	0.373***	−0.242*	0.148***	0.165
	(0.034)	(0.95)	(0.026)	(0.098)
# Obs.	402	400	378	376
# Groups	201	200	189	199

Table B1: **Linear Regressions for Responsibility attribution.** Linear regressions with standard errors clustered on individual managers (in parenthesis). The dependent variable is probability of assigning responsibility of the group's outcome by the manager in columns **I** and **II** for successful outcomes and in columns **III** and **IV** for failed outcomes. ***, ** and * indicate statistical significance at the 0.001, 0.01 and 0.05 levels.

In Table B2, we replicate the main analysis on hiring decisions. Managers are split into those whose accuracy on the actual contributions of the builders (number of pieces) was below the median in columns **I-II, V-VI** and above the median in columns **III-IV, VII-VIII**.

	Success				Failure			
	Below median		*Above Median*		*Below median*		*Above Median*	
	I	**II**	**III**	**IV**	**V**	**VI**	**VII**	**VIII**
Builder 3	0.050	0.096	0.030	0.054	−0.263***	−0.299***	−0.404***	−0.430***
	(0.091)	(0.081)	(0.085)	(0.085)	(0.074)	(0.074)	(0.072)	(0.071)
Contribution		1.709***		0.560		1.309***		−0.352
		(0.320)		(0.681)		(0.352)		(0.565)
Difficulty		0.113		0.140		−0.051		0.098
		(0.079)		(0.088)		(0.072)		(0.072)
Constant	0.380***	−0.264*	0.347***	0.098	0.421***	0.013	0.521***	0.626**
	(0.052)	(0.198)	(0.051)	(0.121)	(0.049)	(0.106)	(0.048)	(0.233)
# Obs.	200	198	202	202	190	188	188	188
# Groups	100	99	101	101	95	94	94	94

Table B2: **Linear Regressions for hiring - Median split on accuracy of contributions.** Linear regressions with standard errors clustered on individual managers (in parenthesis). The dependent variable is probability of being hired by the manager in columns **I-IV** for successful outcomes and in columns **V-VIII** for failed outcomes. Managers are split into those whose accuracy on the actual contributions of the builders (number of pieces) was below the median in columns **I-II, V-VI** and above the median in columns **III-IV, VII-VIII**. ***, ** and * indicate statistical significance at the 0.001, 0.01 and 0.05 levels.

C Additional treatments

C.1 Regressions

We replicate the analysis for hiring decisions as presented for the Baseline in Table 1. We report linear probability models testing hiring decisions for treatment Involved in Table C1 and for treatment Rewards in Table C2. In all regressions the independent variables are dummies for the final builders, so that the first builder is the omitted category. Moreover, we use individual-level random effects in all regressions to cluster decisions at the manager's level. The dependent variable is the probability of hiring a builder, for the case of Success in columns **I-II** and Failure in columns **III-IV**.

	Success		**Failure**	
	I	**II**	**III**	**IV**
Builder 3	0.021	0.003	−0.444***	−0.496***
	(0.061)	(0.058)	(0.047)	(0.049)
Contribution		0.876*		0.370
		(0.434)		(0.234)
Difficulty		0.269***		0.094*
		(0.064)		(0.045)
Constant	0.344***	−0.020	0.541***	0.402***
	(0.034)	(0.140)	(0.035)	(0.087)
# Obs.	384	382	410	410
# Groups	192	191	205	205

Table C1: **Linear Regressions for hiring in treatment INVOLVED.** Linear regressions with standard errors clustered on individual managers (in parenthesis). The dependent variable is probability of being hired by the manager in columns **I** and **II** for successful outcomes and in columns **III** and **IV** for failed outcomes. ***, ** and * indicate statistical significance at the 0.001, 0.01 and 0.05 levels.

Table C3 reports linear probability models testing hiring decisions by pooling treatments BASELINE, INVOLVED, and REWARDS together. In all regressions the independent variables are dummies for the final builders, so that the first builder is the omitted category. Moreover, we use individual-level random effects in all regressions to cluster decisions at the manager's level. The dependent variable is the probability of hiring a builder, for the case of Success in column **I** and Failure in column **II**.

C.2 Figures

Below we include figures of the main choices and outcomes for treatments INVOLVED (Figure C1) and REWARDS (Figure C2).

	Success		Failure	
	I	**II**	**III**	**IV**
Builder 3	0.049	0.065	−0.182***	−0.206***
	(0.061)	(0.057)	(0.050)	(0.057)
Contribution		0.875**		0.244
		(0.313)		(0.435)
Difficulty		0.277***		0.038
		(0.066)		(0.060)
Constant	0.317***	−0.050	0.379***	0.294*
	(0.035)	(0.106)	(0.033)	(0.139)
# Obs.	366	366	428	428
# Groups	183	183	214	214

Table C2: **Linear Regressions for hiring in treatment REWARDS.** Linear regressions with standard errors clustered on individual managers (in parenthesis). The dependent variable is probability of being hired by the manager in columns **I** and **II** for successful outcomes and in columns **III** and **IV** for failed outcomes. ***, ** and * indicate statistical significance at the 0.001, 0.01 and 0.05 levels.

	Success	Failure
	I	**II**
Builder 3	0.040	−0.333***
	(0.062)	(0.051)
Builder 1 X INVOLVED	−0.019	0.071
	(0.048)	(0.050)
Builder 1 X REWARDS	−0.046	−0.092
	(0.048)	(0.049)
Builder 3 X INVOLVED	−0.038	−0.040
	(0.049)	(0.033)
Builder 3 X REWARDS	−0.037	0.059
	(0.050)	(0.037)
Constant	0.363***	0.471***
	(0.034)	(0.036)
# Obs.	1152	1216
# Groups	576	608

Table C3: **Linear Regressions for hiring pooling all treatments.** Linear regressions with standard errors clustered on individual managers (in parenthesis). The dependent variable is probability of being hired by the manager in column **I** for successful outcomes and in column **II** for failed outcomes. ***, ** and * indicate statistical significance at the 0.001, 0.01 and 0.05 levels.

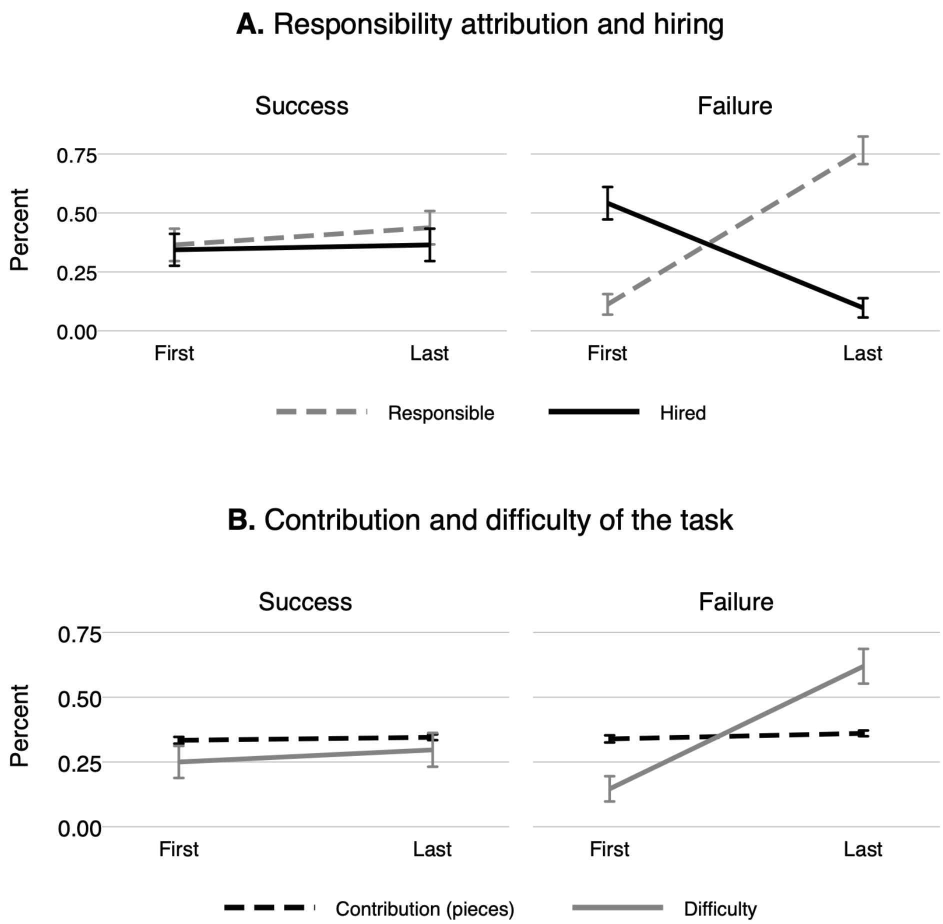

Figure C1: **Main Experimental Outcomes in treatment INVOLVED.** *Panel A:* The solid line (in both panels) displays the fraction of managers who hired the worker in either the First or Last position. The dashed line displays the fraction of managers who made the worker in the First or Last position responsible for the success (left panel) or failure (right panel) of the outcome. *Panel B:* The dashed line displays the perceived contribution by the manager of each player to the task, expressed as the share of total pieces put together. The solid line displays the perceived difficulty by the managers of the task each player was given according to their position.

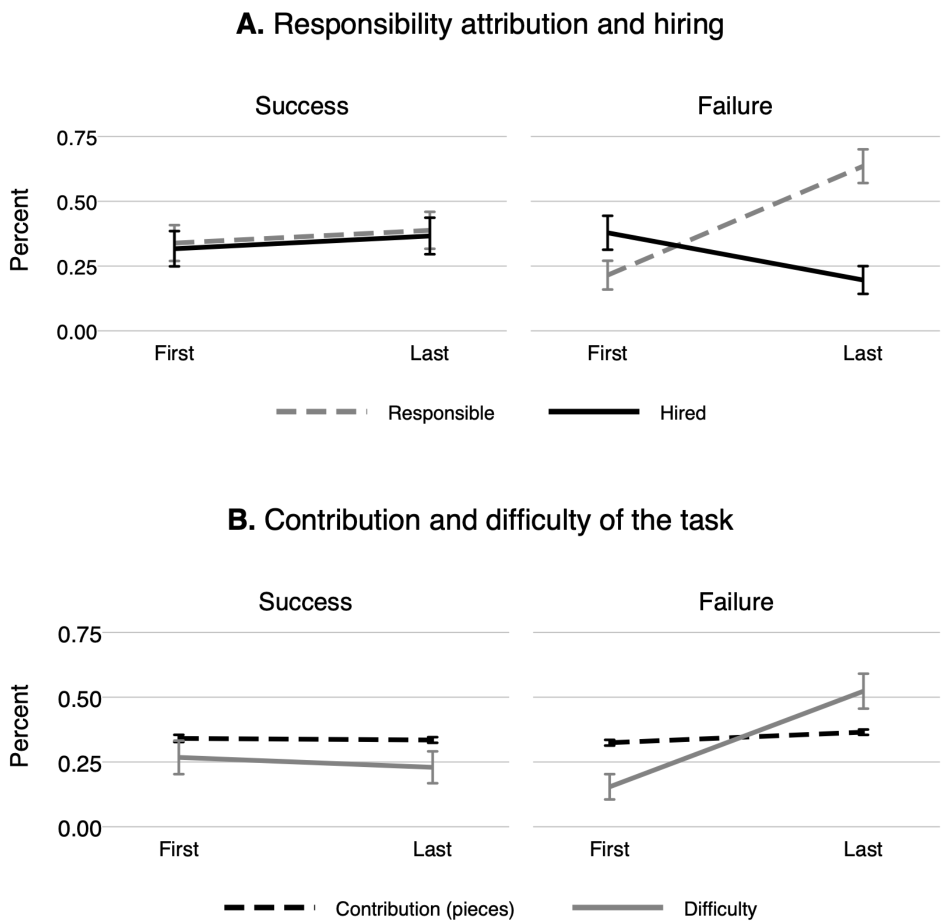

Figure C2: **Main Experimental Outcomes in treatment REWARDS.** *Panel A:* The solid line (in both panels) displays the fraction of managers who hired the worker in either the First or Last position. The dashed line displays the fraction of managers who made the worker in the First or Last position responsible for the success (left panel) or failure (right panel) of the outcome. *Panel B:* The dashed line displays the perceived contribution by the manager of each player to the task, expressed as the share of total pieces put together. The solid line displays the perceived difficulty by the managers of the task each player was given according to their position.

D Robustness treatment

Compared to the main treatments, we design an additional treatment labeled ROBUSTNESS to investigate whether the salience of last builders mattered. For this, we modified the same videos by *down playing* the success or failure of the outcomes. Specifically, in the main treatments, at the end of every video there is a 15 second effect highlighting the success or failure of the building task. For the ROBUSTNESS treatment, we eliminate the 15-second highlight and simply show a white screen with a text in black stating whether the task was built on time or not. Aside from this change at the end of the videos, everything else stayed exactly the same as in the BASELINE treatment. The results of the comparison are displayed below.

Table D1 reports linear probability models testing hiring decisions by pooling treatments BASELINE and ROBUSTNESS together. In all regressions the independent variables are dummies for the final builders, so that the first builder is the omitted category. Moreover, we use individual-level random effects in all regressions to cluster decisions at the manager's level. The dependent variable is the probability of hiring a builder, for the case of Success in column **I** and Failure in column **II**.

	Success	**Failure**
	I	**II**
Builder 3	0.040	−0.333***
	(0.062)	(0.051)
Builder 1 X ROBUSTNESS	−0.069	−0.032
	(0.047)	(0.050)
Builder 3 X ROBUSTNESS	0.023	0.077*
	(0.049)	(0.038)
Constant	0.363***	0.471***
	(0.034)	(0.036)
# Obs.	810	788
# Groups	405	394

Table D1: **Linear Regressions for hiring pooling treatments BASELINE and ROBUSTNESS.** Linear regressions with standard errors clustered on individual managers (in parenthesis). The dependent variable is probability of being hired by the manager in column **I** for successful outcomes and in column **II** for failed outcomes. ***, ** and * indicate statistical significance at the 0.001, 0.01 and 0.05 levels.