Singapore Management University

Institutional Knowledge at Singapore Management University

Research Collection Lee Kong Chian School Of Business

Lee Kong Chian School of Business

2-2024

Learning from machines: How negative feedback from machines improves learning between humans

Tengjian ZOU

Gokhan ERTUG

Thomas ROULET

Follow this and additional works at: https://ink.library.smu.edu.sg/lkcsb_research

Part of the Educational Methods Commons, Graphics and Human Computer Interfaces Commons, Strategic Management Policy Commons, and the Technology and Innovation Commons

This Journal Article is brought to you for free and open access by the Lee Kong Chian School of Business at Institutional Knowledge at Singapore Management University. It has been accepted for inclusion in Research Collection Lee Kong Chian School Of Business by an authorized administrator of Institutional Knowledge at Singapore Management University. For more information, please email cherylds@smu.edu.sg.

Journal of Business Research 172 (2024) 114417

Contents lists available at ScienceDirect

Journal of Business Research

journal homepage: www.elsevier.com/locate/jbusres

Learning from machines: How negative feedback from machines improves learning between humans

Tengjian Zou [a,*], Gokhan Ertug [b], Thomas Roulet [c]

[a] Zhejiang University, China
[b] University of Adelaide, Australia
[c] University of Cambridge, United Kingdom

ARTICLE INFO

Keywords:
Learning
Failure
Machine-human interaction

ABSTRACT

Prior studies on learning from failure primarily focus on how individuals learn from failure feedback given by other individuals. It is unclear whether and how the advent of machine feedback may influence individuals' learning from failures. We suggest that failure feedback provided by machines facilitates learning in two ways. First, it focuses individuals' attention on their failures, leading them to learn from these failures. Second, it serves as a catalyzer, motivating individuals to learn more from failure feedback given to them by other individuals as well. In addition, this catalyzing effect is stronger if the failure feedback from machines and by other individuals pertain to related tasks. Using a dataset of 1.5 million observations from an online programming contest community, we find support for our predictions. We contribute to the learning literature by demonstrating both the direct effect and the catalyzing effect of machine failure feedback on individuals' learning.

1. Introduction

Learning is a process in which knowledge is created through the transformation of prior experience (Kim, 1998), potentially translating this experience into superior performance (Kolb, 1984). A key dimension of prior experience is whether its outcome was a success or a failure (KC et al., 2013). Because failures provide valuable learning opportunities for individuals to adjust their practices, learning from failure is critical to improve performance (Sitkin, 1992). Failures can be communicated to an individual through feedback given by different sources (Ilgen & Davis, 2000; Jordan & Audia, 2012), such as a supervisor (Fedor et al., 2001) or a peer (Ashford & Tsui, 1991).

The literature on learning from failure focuses primarily on how humans learn from failure feedback given by other humans (see a review by Dahlin et al., 2018). However, in many contexts, it is becoming increasingly common for individuals to receive failure feedback given by machines as well as by humans (Rahwan et al., 2020). For example, coders who work in companies receive feedback about coding errors from their supervisors or peers, as well as from machines (e.g., Microsoft Visual C++). Considering a different setting, Luo et al. (2021) note that, in addition to human coaches, organizations are now using artificial intelligence coaches to provide feedback to their sales agents and to foster learning. Despite its increasing prevalence, we have limited understanding about how machine feedback influences individuals' learning, including how machine feedback and human feedback might interact to affect individuals' learning. These are important issues, with implications for whether and how organizations should implement machines as a feedback source.

Building on the literature on learning from failure, we propose that machine failure feedback has both a direct effect and a catalyzing effect on individuals' learning from failure. The direct effect reflects the idea that machine failure feedback draws individuals' attention to their failures, stimulating them to learn from these failures. The catalyzing effect reflects the idea that machine failure feedback also increases individuals' awareness of those failures that are identified by human failure feedback, which then motivates these individuals to deploy more resources to learn from human failure feedback as well. In addition, we expect this catalyzing effect to be stronger if the failure feedback from machines and by other individuals pertain to related tasks. We test our predictions using data from the activities of 93,145 contestants in an online programming contest community, yielding about 1.5 million observations. This setting is well suited to test our hypotheses because contestants can receive failure feedback from humans, who are members in the community (whom we refer to as "peers" when discussing our

* Corresponding author.
E-mail address: zoutengjian@zju.edu.cn (T. Zou).

https://doi.org/10.1016/j.jbusres.2023.114417
Received 2 February 2022; Received in revised form 30 January 2023; Accepted 20 November 2023
Available online 6 December 2023
0148-2963/© 2023 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

setting), as well as from machines. In addition, in this setting learning outcomes can be measured precisely and objectively. Our analyses of this large dataset provide support for our hypotheses. Our study enriches the learning literature by demonstrating that machine failure feedback plays a key and thus far understudied role on individual learning, and also contributes to the literature on human–machine interaction by showing that machines can significantly impact the dynamics of learning between individuals.

2. Theory and hypotheses

Failure is performance that falls short of a desired outcome (Cannon & Edmondson, 2001). For example, an employee's creative idea might be rejected by an organization's evaluation committee (Wilhelm et al., 2019), the cardiac procedure performed by a surgeon might be unsuccessful (KC et al., 2013), a project leader might not find the expected outcome in their project (Burgers, Van Den Bosch, & Volberda, 2008; Shepherd et al., 2011), researchers might revisit interventions in organizations that ended up being unsuccessful (Hodgkinson & Wright, 2002), or an entrepreneur's new venture might not survive (Ucbasaran et al., 2013; Boso et al., 2019; Amankwah-Amoah, Adomako, & Berko, 2022). Learning from failure is a process by which individuals reflect on what might have gone wrong in their practices and develop solutions to avoid similar mistakes in the future, ultimately improving performance (Dahlin et al., 2018; Zellmer-Bruhn & Gibson, 2006).

2.1. Learning from failure as a consequence of peer feedback

One key mechanism that connects failure and subsequent learning is motivation to learn (Dahlin et al., 2018). Motivation to learn from failure refers to the resources that are expended by individuals in activities that relate to learning from failure. These resources can include time, attention, and effort. Learning from failure requires individuals to allocate resources to identify and analyze the causes of their failures, and search for and implement solutions to improve performance and prevent similar failures from recurring (Ellis et al., 2014). Even though failure feedback *does* offer an opportunity for correcting mistakes and learning new ways of doing things, not taking any actions following failure feedback would keep an individual from learning (Eskreis-Winkler & Fishbach, 2019). Hence, the performance-enhancing effect of failure feedback depends on an individual's willingness to expend time, attention, and effort to learn from the feedback received.

Prior studies have examined how people learn from their failures (Dahlin et al., 2018). In many cases, failure is revealed to individuals by feedback from external parties, i.e., by sources other than the individuals themselves (Ilgen & Davis, 2000). Peers are one external source of failure feedback, insofar as failure feedback from them can make failures visible to a focal actor and trigger learning (Baum & Dahlin, 2007). This visibility raises the focal individual's awareness that some of his or her practices might be problematic (Kim & Miner, 2007). As a result, this individual will be motivated to learn from these failures, for example by taking actions to scrutinize the failure information, figuring out the cause of failures, and remedying problematic practices. This process results in learning from failures, i.e., it generates subsequent performance improvement by reducing the chances that those mistakes happen again.

Research in different settings has documented that individuals learn from their failures when those failures are brought to their attention by peers. For example, in a field experiment, Ellis and Davidi (2005) showed that the performance of soldiers improved when they were debriefed on their failures after each training day. Similarly, Ellis, Mendel, and Nir (2006) reported that students' performance improved after they experienced failed events (i.e., they did not win a prize in the game), which were pointed out to them in a laboratory study.

In line with the findings in this body of literature, our baseline prediction is that peer failure feedback promotes learning, which leads to performance improvement.

Baseline hypothesis: *An individual learns from prior peer failure feedback, such that there is a positive relationship between the amount of peer failure feedback that the individual previously received and the individual's current performance.*

2.2. The role of machines in learning from failure

Research has shown that machine feedback can influence human–human interactions. For example, in a set of experiments, Shirado and Christakis (2017) showed that machines can improve coordination between humans. In another experimental study, Jung, Martelaro, and Hinds (2015) found that machines can enhance team functioning by regulating core team processes, such as conflict. In a related study, Traeger and colleagues (2020) demonstrated that machines can positively shape human conversational dynamics in human-robot teams.

Machines are increasingly being used as a source of feedback (e.g., Merle, St-Onge, & Sénécal, 2022; Van der Kleij, Feskens, & Eggen, 2015). As a result, humans can receive negative feedback provided by machines. Such negative feedback can shape human behaviors. For example, two studies found that negative feedback from machines promotes people's energy conservation behavior (Midden & Ham, 2009; Ham & Midden, 2014). The authors suggest that machines can exert a "social influence" on humans, which in turn changes human behavior. In other words, machines can serve as sources of feedback, which then leads to behavioral change in humans. Just as humans can provide both positive and negative feedback to other individuals to affect their learning (e.g., Arbel, Murphy, & Donchin, 2014), machines can provide positive as well as negative feedback to humans in their feedback-providing role. Humans tend to be selective and biased in how they learn from other humans (Tourish & Hargie, 2012), but machines have the potential to be a source that yields more objective learning. The research by Midden and Ham (2009) and Ham and Midden (2014) found that negative feedback provided by machines is more effective in shaping humans' behaviors, as compared to positive feedback given by machines. We build on insights from this stream of literature to propose that machine failure feedback has both a direct effect and a catalyzing effect on individuals' learning from failures, as we detail below in turn.

First, machine failure feedback can directly impact individuals' learning from failures. The mechanism here is similar to how peer failure feedback enhances individuals' learning from failures, as we outlined in developing the baseline hypothesis. Machine failure feedback raises individuals' attention to their failures, increasing their awareness that there might be errors in their existing practices. As a result, individuals will be motivated to expend time, attention, and effort to identify the cause of these failures and search for and implement solutions to adjust their practices. These arguments lead us to hypothesize that:

Hypothesis 1. *An individual learns from prior machine failure feedback, such that there is a positive relationship between the amount of machine failure feedback that the individual previously received and that individual's current performance.*

In addition to its direct effect on individuals' learning, we argue that machine failure feedback also has a catalyzing effect, by which it can enhance how much individuals learn from failure feedback provided by peers. KC and colleagues (2013) argued that other people's failures capture the attention of a focal individual and elicit a motivational response for this individual to learn from their own failures, so that they can evade similar failures and adverse consequences. Complementing this argument, we suggest that machine failure feedback ends up being motivational for individuals to learn from failure feedback that is given by peers as well. When individuals receive more machine failure feedback, we expect that they would be more likely to be aware of the potential to learn from peer failure feedback as well. Accordingly, when they have received more machine failure feedback, individuals will be more motivated to allocate resources to learn from peer failure feedback. In contrast, a lower level of machine failure feedback might lead individuals to believe that there are not many genuinely problematic issues with their routines, and that there is no urgency to take remedial action. As a result, lower level of machine failure feedback – compared to higher level of machine failure feedback – leaves individuals less likely to be stimulated to learn from peer failure feedback. For example, coders who work in organizations can receive feedback about coding errors from their supervisors or peers, as well as from machines. In this scenario, our suggestion is that more machine failure feedback would also raise individuals' attention to failure feedback from humans and highlight the potential to learn from that type of failure feedback as well.

In summary, we argue that machine failure feedback raises individuals' awareness of the potential to learn in general. This motivates individuals to allocate resources to learn more from peer failure feedback as well. Accordingly, our expectation is that machine failure feedback positively moderates the relationship between peer failure feedback and individual performance, such that this relationship is more positive if the individual received more machine failure feedback, and less positive if the individual received less machine failure feedback.

Hypothesis 2. *Prior machine failure feedback amplifies the positive relationship between prior peer failure feedback and current performance.*

To the extent that individuals can receive failure feedback on different types of tasks, the relatedness between machine failure feedback and peer failure feedback may vary, in terms of the tasks on which that feedback is provided (Schilling et al., 2003). We propose that how much machine failure feedback improves learning from peer failure feedback, i.e., the catalyzing effect in H2, depends on the relatedness of the tasks on which feedback is given. Related stimuli can more easily activate people's awareness to the learning process (Müller et al., 2016; Wang & Nickerson, 2019), In contrast, people are less responsive to unrelated stimuli (Hodgkinson & Wright, 2002). In line with this reasoning, we suggest that when machine failure feedback is received on tasks that are related to tasks on which peer failure feedback is received, machine failure feedback will more strongly amplify the relationship between peer failure feedback and learning. Machine failure feedback that is more related to peer failure feedback, in terms of the tasks on which that feedback is given, is more motivational. As a result, it drives individuals to allocate more time, attention, and effort to scrutinize the peer feedback and learn from it. In comparison, a lower level of relatedness between machine failure feedback and peer failure feedback does not raise the awareness of individuals to the potential to learn from peer failure feedback to the same degree. Therefore, we predict that:

Hypothesis 3. *The moderating effect of prior machine failure feedback on the relationship between prior peer failure feedback and current performance is stronger if prior machine failure feedback is more related to prior peer failure feedback.*

3. Data and methods

3.1. Setting

We collected data from an online community that hosts programming contests. These contests offer opportunities for members of this community, which range from students to professionals, to practice and improve their coding skills. Our data covers the period from February 2010, when this online community was launched, to October 2018. There are 988 contests in this time period. On average, this online community organizes a contest every four days. About 1,775 contestants are registered in each contest. Contests last for 4.8 h on average. Our sample has 1,474,753 contest-participant observations, which come from the activities of 93,145 contestants.

Machine failure feedback. Contestants typically have five to six programming problems to solve in a contest. A problem solution is a piece of code written in a programming language (e.g., C, C++, Java, or Python). Once the problem solution is submitted, it is automatically and immediately evaluated by the system (i.e., an automated machine). This process is referred to as a "machine test" in this community. After this test is completed, the machine provides two types of feedback. If the solution solves the problem, the machine returns feedback indicating that the solution is a success. If the solution does not solve the problem, the machine provides feedback suggesting that the solution contains errors. However, the machine does not specify what these errors are (the feedback is simply that the code has errors, in that it does not solve the problem). Hence the contestants need to inspect their code and figure out why their solution failed, and correct the errors. In addition to such feedback during contests, this community also allows contestants to practice (in non-contest settings) using problems that were used in earlier contests. During such practice, contestants can again receive machine feedback. In other words, individuals in this community can receive machine feedback in contests, as well as during practice sessions, which take place outside of contests.

Peer failure feedback. In addition to machine failure feedback, contestants can also receive failure feedback from other contestants (i.e., peers) during contests. This is possible because this community allows contestants to view and identify errors in others' codes during the contest. Contestants can view another person's code for each problem, but only if they submit and "lock" their own solution first, so that they cannot change it anymore for that problem. If a contestant finds errors in the code of another person, then the contestant can suggest a test (i.e., a set of input data) on which the contestant expects that the solution by that other person will fail. This procedure is called "hacking" another contestant's solution. If the hacking is successful, the person receives a piece of feedback suggesting that his or her solution failed. Similar to the failure feedback given by machine, the failure feedback provided by peers does not specify what the errors are (as in the machine case, the feedback is simply that the code has errors, in that it does not solve the problem). Hence the feedback receiver needs to inspect their code and figure out why their solution failed, and correct the errors.

Calculation of performance. Contestants' performance is calculated when a contest is over. Performance is a sum of two parts. The first part is the points a contestant earns by solving problems during a contest, which are calculated as follows: Each minute taken by the contestant to solve a problem decreases the problem's value by $X/250$ points per minute (where X is a problem's initial, or full, points). The number of points a contestant gets for solving a problem equals the current value of the problem (in points) minus penalties. The penalties are determined based on the number of failed solutions by the contestant that have been submitted for this problem, multiplied by 50 points. In order to maximize the points that they earn for a problem, a contestant should submit a solution as soon as possible, but also make sure that the solution will pass the tests, i.e., be a success. For example, if a problem has initial value of 500 points, and a contestant solves the problem in the third trial, after two minutes, then the points they will earn for this problem are $500 - (500/250) * 2 - 2 * 50 = 396$.

The second part is the points a contestant earns in a contest as a result of successful and unsuccessful hacks. A contestant earns 100 points for a successful hack (of another contestant's submitted solution) and loses 50 points for an unsuccessful hack.

Based on the sum of these two parts, a total sum of points is calculated for each contestant. The contestants are ranked in a leaderboard, which indicates the final ranking in that contest. In cases when multiple contestants obtain the same points, their ranks are the same. An illustration of the process of a contest is shown in Fig. 1.

3.2. Dependent variable

Performance. We measure *Performance* by a contestant's rank in a given contest, which reflects a contestant's relative performance compared to others in this contest. Rank is a frequently used measure to

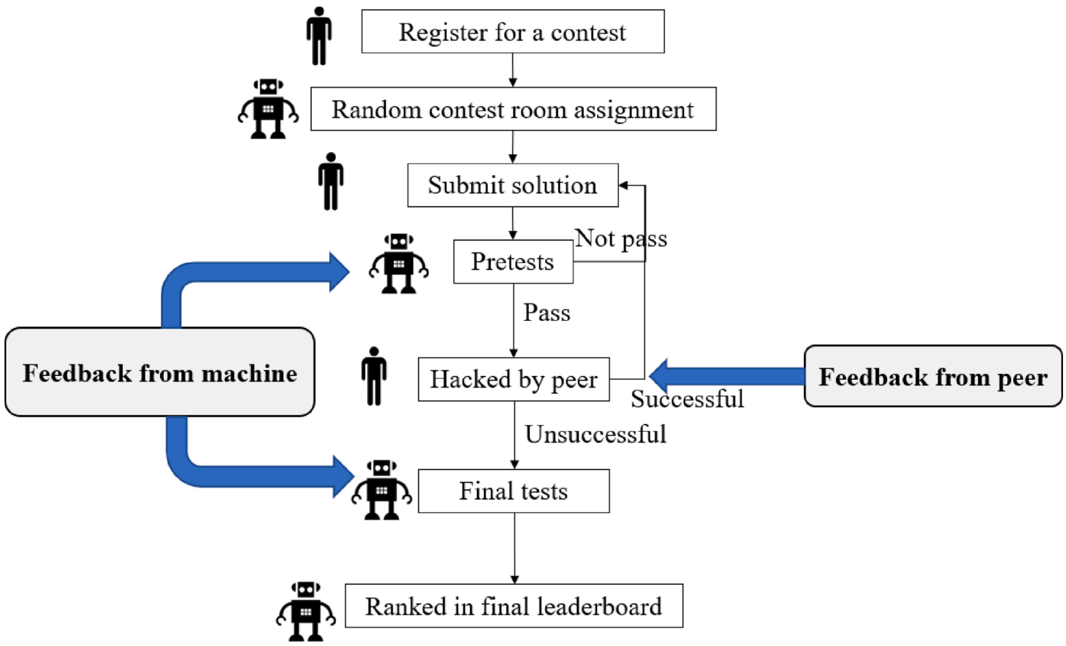

Fig. 1. An illustration of the process of a contest. ***Note***: A human icon on the left of the box indicates that the step is completed by a contestant. A robot icon on the left of the box indicates that the step is performed by the system.

assess individuals' relative performance (e.g., Bhattacharjee et al., 2007; Groysberg et al., 2008; Kuhnen & Tymula, 2012). As we described above, contestants are ranked based on the total points they obtain in a contest, which is the sum of points earned from solving problems and from hacking. For example, if a contestant solved three problems in a contest and obtained 236, 358, and 532 for each problem, and also made two successful hacks and one unsuccessful hack, the final points are 1,276 (236 + 358 + 532 + 100*2 – 50*1).

The highest rank in a contest is 1 and the lowest rank is equal to the number of contestants in that contest. As the number of contestants varies across contests, we normalize the rank to 0–1 range using a min–max normalization. The smaller the normalized rank, the better the performance. Then, for easier interpretation, we reverse code this normalized rank by subtracting it from 1 and multiplying it by 100. As a result, performance ranges from 0 to 100, where the larger the number, the better the performance. The way we normalize rank and calculate performance is shown in formula (1) below.

$$Performance = \left(1 - \frac{rank - \min(rank)}{\max(rank) - \min(rank)}\right)*100, \quad (1)$$

where *rank* is focal contestant's rank in a contest. min*(rank)* and max *(rank)* are the minimum and maximum rank, respectively, in a contest. In our setting, min*(rank)* is 1 and max*(rank)* is the total number of contestants in a contest. In our additional analyses, we also consider a different measure of performance, which we discuss after presenting the main results.

3.3. Independent variables

Prior peer failure feedback. Hacking is a process through which a contestant identifies failures in other contestants' code and makes the presence of failures in their code known to them. Based on this, we consider a hack to constitute failure feedback provided by a human peer to the feedback receiver. Accordingly, to assess our baseline hypothesis, we count the number of times that a contestant was hacked by other contestants in prior contests (excluding the current contest) to measure *Prior peer failure feedback*.

Prior machine failure feedback. Whenever a contestant's submitted solution to a problem does not pass the system test, we take this to be an instance of a contestant receiving failure feedback from a machine. In line with this, to test Hypothesis 1, we measure *Prior machine failure feedback* by counting the number of times that a focal contestant's solutions failed machine tests in previous contests (excluding the current contest).

Because the distribution of these two independent variables is skewed, we log transform them. We test the moderation effect postulated in Hypothesis 2 by using an interaction variable between *Prior peer failure feedback* and *Prior machine failure feedback*.

3.4. Moderating variables

The tasks on which participants receive machine failure feedback in this setting can be more or less related to tasks on which participants also receive peer failure feedback. To test Hypothesis 3, we disaggregate *Prior machine failure feedback* into two categories, based on the tasks on which that feedback is provided. One category consists of machine failure feedback received for tasks that are more related to tasks on which *Prior peer failure feedback* is received. The other category consists of machine failure feedback received for tasks that are less related to those tasks on which *Prior peer failure feedback* is provided. We use two ways to perform this disaggregation.

The first way we disaggregate tasks into those that are more or less related to each other leverages the fact that this online community has two types of problems. One type is contest problems, which are originally developed by this onine community and used in contests that are hosted by this online community. The other type is gym problems, which are collected by this online community from external sources (e.g., Facebook Hacker Cup) and provide practice opportunities for members in this community. Contestants can receive failure feedback from peers (i.e., hacks) only in contest problems, but not in gym problems. Therefore, we consider failure feedback given by a machine in contest problems to be more related to failure feedback given by peers, as compared to failure feedback given by a machine in gym problems. Accordingly, to test Hypothesis 3, we disaggregate *Prior machine failure feedback* into *Prior machine failure feedback in contest problems* (the number of failed solutions a contestant had in previous contest problems) and *Prior machine failure feedback in gym problems* (the number of failed solutions a contestant had in previous gym problems).

The second way we disaggregate *Prior machine failure feedback* leverages the fact that problems A and B in a contest are the easier problems, compared to the rest (i.e., problems C, D, E, …). Because a clear majority (72.7 %) of failure feedback given by peers (i.e., hacks) occurs in problems A and B, we suggest that failure feedback given by a machine in problems A and B would be more related (in terms of the tasks on which feedback is provided) to failure feedback given by peers. Based on this argument, we also test Hypothesis 3 by disaggregating *Prior machine failure feedback* into *Prior machine failure feedback in problems A and B* (the number of failed solutions of a contestant had in problems A and B) and *Prior machine failure feedback in the rest of the problems* (the number of failed solutions of a contestant had in the rest of the problems).

All of the moderating variables above are log transformed in light of their skewed distribution.

3.5. Control variables

Past performance. We measure a contestant's past performance by using the average performance of this contestant in prior contests. We control for past performance to account for contestants' underlying skills or ability. If a contestant performed well in past contests, she is likely to continue to perform well.

Status. We measure status using a contestant's Elo rating at the time that they join the contest. A person's status indicates their hierarchical position in a social system (e.g., Piazza & Castellucci, 2014). Elo ratings (which are also used in chess, tennis, and some video games) indicate a contestant's relative standing compared to other contestants in a hierarchy. Status confers multiple advantages that might positively affect performance. The highest Elo rating in our sample is 3,739. We divided this measure by 100 to facilitate the reporting of the coefficients and standard errors.

Knowledge specialization. Every problem in a contest has a list of tags (e.g., dynamic programming, data structure) that specify the knowledge required to solve it. We use these tags to measure a contestant's *Knowledge specialization* with respect to problems asked in the current contest. Specifically, we first aggregate the tags for all the problems solved by the contestant in past contests, and then sum the number of times each tag in the current contest appears in this aggregated tag list. *Knowledge specialization* is measured as the number of times that tags (for all the problems in the current contest) appear in that aggregated tag list, normalized by the total length of the aggregated list. For example, if the aggregated list of tags for a contestant is [a, a, a, b, b, c, d, e, f, g, h, h] and the tags in focal contest are [a, b, c, d, e], then the contestant's *Knowledge specialization* is $(3 + 2 + 1 + 1 + 1) / 12 = 0.667$.

Prior unsuccessful hacks from peers. As mentioned earlier, to assess the baseline hypothesis, we use the number of times that a contestant was hacked – successfully – by peers in prior contests. However, because a hack attempt can also be unsuccessful, we control for *Prior unsuccessful hacks from peers*, which is the number of times that a contestant was unsuccessfully hacked by peers in prior contests. An unsuccessful hack attempt occurs when a contestant believes they identified an error in the code of another contestant, but the code is in fact able to process the suggested input as expected, without an error. We control for this variable because if a contestant was unsuccessfully hacked by peers in prior contests, she might become confident about her solutions, which can positively affect her performance in current contest, or she might become complacent, which can negatively affect performance in current contest. Given its skewness, we used a logarithmic transformation of this measure.

Experience of hacking peers in prior contests. Similar to how a contestant can be hacked by peers, she can also hack peers, and these hacks may be successful or unsuccessful. We use two variables to control for the focal contestant's experience of hacking peers in prior contests. The first is *Prior successful hacks to peers*, which is the total number of times that a contestant successfully hacked peers in prior contests. The second, *Prior unsuccessful hacks to peers*, is the total number of times that a contestant unsuccessfully hacked peers in prior contests. We control for *Prior successful hacks to peers* because successful hacks to peers can increase the contestant's confidence, which can positively affect her performance in a focal contest, or make her complacent, which can negatively affect performance. We control for *Prior unsuccessful hacks to peers* because these can discourage the contestant and dent her confidence, which could negatively affect performance in current contest, or the contestant might learn from this experience, therefore increasing performance in the current contest. We use a log transformation, in light of the skewed distribution of these two measures.

Influence of hacks in the focal contest. During the current contest (i.e., the one for which we assess performance), a contestant can hack others or be hacked by others. We use four variables to control for the influence of such hacking events on a contestant's performance in a focal contest. *Number of successful hacks sent by focal contestant in focal contest* is the number of times the contestant hacked others successfully in the current contest. We control for this because a contestant earns 100 points for a successful hack, which positively contributes to performance. *Number of unsuccessful hacks sent by focal contestant in focal contest* is the number of times the contestant was unsuccessful in hacking others in the current contest. We control for this because contestants lose 50 points for an unsuccessful hack, negatively affecting the total points they gather in a contest and their ranking. *Number of successful hacks received by a contestant in focal contest* is the number of times the contestant was hacked successfully by others in the current contest. If a contestant's solution is successfully hacked by others, this solution is marked as failed, which reduces performance. Finally, *Number of unsuccessful hacks received by a contestant in focal contest,* is the number of times the contestant was hacked unsuccessfully by others in the current contest. If a contestant's solution is hacked unsuccessfully by others, she may become more confident, which can positively influence performance. As with other similar measures, we logarithmically transformed these four measures.

3.6. Analyses

We used ordinary least squares (OLS) regressions that consistently include fixed effects for contests, contestants, and rooms. We used fixed-effects estimations based on the results of a Hausman test ($p < 0.001$). Because of the multiple fixed-effects that we include in all of our models, contest characteristics (e.g., number of participants in the contest, duration of contest), non-time-varying contestant attributes (e.g., ethnicity, sex), and room characteristics (e.g., number of contestants in the room) are not separately estimated, since they do not vary within the fixed-effects control variables in our observation period. We use robust standard errors that are clustered at the contest, contestant, and room level, to take into account the possible non-independence of observations across the same contest, contestant, and room.

3.7. Model specification

In our baseline hypothesis, we proposed that prior peer failure feedback is positively related to individuals' current performance, and then proposed in Hypothesis 1 that prior machine failure feedback is also positively related to individuals' current performance. The model shown in equation (1) is used to test the baseline hypothesis and Hypothesis 1:

$$Performance_{ic} = \beta_0 + \beta_1 PeerFeedback_i + \beta_2 MachineFeedback_i + Controls_i + \delta_c + \delta_i + \delta_r + \varepsilon_i \tag{1}$$

where $Performance_{ic}$ refers to contestant i's performance in contest c, $PeerFeedback_i$ refers to the amount of peer failure feedback that contestant i received before contest c, $MachineFeedback_i$ refers to the amount of machine failure feedback that contestant i received before contest c, δ_c, δ_i, and δ_r refer to contest, contestant, and room fixed effects, and ε_i refers to random error term. If β_1 and β_2 are positive and significant, then the baseline hypothesis and Hypothesis 1 are supported.

We proposed in Hypothesis 2 that prior machine failure feedback amplifies the positive relationship between prior peer failure feedback and current performance. The model shown in equation (2) is used to test Hypothesis 2:

$$\begin{aligned} Performance_{ic} = {} & \beta_0 + \beta_1 PeerFeedback_i + \beta_2 MachineFeedback_i \\ & + \beta_3 MachineFeedback_i{}^{*}PeerFeedback_i + Controls_i \\ & + \delta_c + \delta_i + \delta_r + \varepsilon_i \end{aligned} \tag{2}$$

where $Performance_{ic}$ refers to contestant i's performance in contest c, $PeerFeedback_i$ refers to the amount of peer failure feedback that contestant i received before contest c, $MachineFeedback_i$ refers to the amount of machine failure feedback that contestant i received before contest c, $MachineFeedback_i{}^{*}PeerFeedback_i$ refers to the interaction term between prior machine failure feedback and prior peer failure feedback, δ_c, δ_i, and δ_r refer to contest, contestant, and room fixed effects, and ε_i refers to random error term. Hypothesis 2 is supported if β_3 is positive and significant.

We proposed in Hypothesis 3 that the moderating effect of prior machine failure feedback on the relationship between prior peer failure feedback and current performance is stronger if prior machine failure feedback is more related to prior peer failure feedback. The model shown in equation (3) is used to test Hypothesis 3:

$$\begin{aligned} Performance_{ic} = {} & \beta_0 + \beta_1 PeerFeedback_i + \beta_2 MachineFeedbackMoreRelated_i \\ & + \beta_3 MachineFeedbackLessRelated_i \\ & + \beta_4 MachineFeedbackMoreRelated_i{}^{*}PeerFeedback_i \\ & + \beta_5 MachineFeedbackLessRelated_i{}^{*}PeerFeedback_i \\ & + Controls_i \\ & + \delta_c + \delta_i + \delta_r + \varepsilon_i \end{aligned} \tag{3}$$

where $Performance_{ic}$ refers to contestant i's performance in contest c, $PeerFeedback_i$ refers to the amount of peer failure feedback that contestant i received before contest c, $MachineFeedbackMoreRelated_i$ refers to machine failure feedback that is more related to peer failure feedback, $MachineFeedbackLessRelated_i$ refers to machine failure feedback that is less related to peer failure feedback, $MachineFeedbackMoreRelated_i{}^{*}PeerFeedback_i$ refers to the interaction term between prior machine failure feedback that is more related to peer failure feedback and prior peer failure feedback, $MachineFeedbackLessRelated_i{}^{*}PeerFeedback_i$ refers to the interaction terms between prior machine failure feedback that is less related to peer failure feedback and prior peer failure feedback, δ_c, δ_i, and δ_r refer to contest, contestant, and room fixed effects, and ε_i refers to random error term. Hypothesis 3 is supported if β_4 has a stronger positive effect than β_5.

4. Results

We present descriptive statistics and correlations in Table 1.[1]

Table 1
Descriptive statistics and correlations (N = 1,474,753).

		Mean	S.D.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	Performance	50.25	28.55																
2	Prior peer failure feedback	1.09	0.92	0.26															
3	Prior machine failure feedback	4.64	1.48	0.26	0.74														
4	Prior machine failure feedback in contest problems	4.58	1.44	0.26	0.75	0.99													
5	Prior machine failure feedback in gym problems	1.22	1.84	0.24	0.50	0.61	0.54												
6	Prior machine failure feedback in problems A and B	4.07	1.41	0.13	0.71	0.95	0.95	0.53											
7	Prior machine failure feedback in the rest of the problems	3.55	1.82	0.42	0.71	0.89	0.87	0.63	0.73										
8	Past performance	45.89	20.96	0.58	0.32	0.31	0.29	0.31	0.11	0.55									
9	Status	14.84	2.87	0.49	0.33	0.31	0.30	0.38	0.13	0.52	0.79								
10	Knowledge specialization	0.58	0.21	−0.11	−0.08	−0.12	−0.12	−0.08	−0.07	−0.17	−0.18	−0.18							
11	Prior unsuccessful hacks from peers	0.99	1.00	0.28	0.76	0.67	0.67	0.49	0.62	0.68	0.38	0.43	−0.10						
12	Prior successful hacks to peers	0.62	1.13	0.28	0.57	0.51	0.50	0.47	0.44	0.54	0.40	0.50	−0.09	0.58					
13	Prior unsuccessful hacks to peers	0.71	1.12	0.25	0.60	0.54	0.54	0.47	0.49	0.56	0.35	0.43	−0.08	0.61	0.85				
14	Number of successful hacks received by focal contestant in focal contest	0.09	0.24	−0.08	0.01	0.02	0.02	0.01	0.02	0.01	−0.01	−0.01	0.04	0.01	0.00	0.00			
15	Number of unsuccessful hacks received by focal contestant in focal contest	0.07	0.23	0.05	0.03	0.03	0.03	0.03	0.02	0.05	0.04	0.04	0.01	0.04	0.03	0.04	0.17		
16	Number of successful hacks sent by focal contestant in focal contest	0.06	0.28	0.19	0.12	0.11	0.11	0.11	0.09	0.13	0.13	0.14	−0.01	0.13	0.26	0.23	0.05	0.05	
17	Number of unsuccessful hacks sent by focal contestant in focal contest	0.05	0.24	0.09	0.09	0.09	0.09	0.09	0.08	0.11	0.09	0.09	−0.01	0.10	0.22	0.23	0.07	0.08	0.49

[1] The correlation table shows that *Status* and *Past performance* are correlated ($r = 0.79$), including both of these variables in our estimations may lead to concerns about collinearity. We conducted a robustness check by removing *Status* from our estimations and the hypotheses remain supported ($p < 0.001$ for H1, H2, H3).

Table 2
OLS fixed-effects models predicting *Performance.*

	(1)	(2)	(3)	(4)	(5)
Prior unsuccessful hacks from peers	5.224***	2.768***	2.293***	1.556***	1.166***
	(0.084)	(0.073)	(0.068)	(0.068)	(0.068)
Prior successful hacks to peers	0.688***	0.607***	0.794***	0.748***	0.473***
	(0.085)	(0.081)	(0.078)	(0.077)	(0.077)
Prior unsuccessful hacks to peers	1.525***	0.737***	0.699***	0.450***	0.139
	(0.090)	(0.086)	(0.083)	(0.083)	(0.083)
Past performance	0.110***	0.084***	0.049***	0.044***	0.044***
	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)
Status	0.468***	0.540***	0.702***	0.708***	0.537***
	(0.049)	(0.046)	(0.046)	(0.046)	(0.046)
Knowledge specialization	−3.643***	−2.732***	0.487	0.438	0.417
	(0.663)	(0.641)	(0.621)	(0.618)	(0.618)
Number of successful hacks received by focal contestant in focal contest	−10.874***	−9.699***	−10.911***	−10.405***	−10.504***
	(0.420)	(0.419)	(0.413)	(0.416)	(0.416)
Number of unsuccessful hacks received by focal contestant in focal contest	5.048***	4.393***	4.248***	4.053***	3.932***
	(0.212)	(0.212)	(0.210)	(0.210)	(0.210)
Number of successful hacks sent by focal contestant in focal contest	8.461***	8.404***	8.318***	8.309***	8.260***
	(0.235)	(0.234)	(0.231)	(0.231)	(0.231)
Number of unsuccessful hacks sent by focal contestant in focal contest	−2.274***	−2.440***	−2.536***	−2.579***	−2.598***
	(0.205)	(0.203)	(0.202)	(0.202)	(0.202)
Prior failure feedback given by peers		5.592***		2.386***	−3.075***
		(0.109)		(0.098)	(0.256)
Prior failure feedback given by machine			5.451***	4.874***	4.884***
			(0.073)	(0.070)	(0.071)
Prior failure feedback given by machine * Prior failure feedback given by peers					1.064***
					(0.048)
Contest fixed effects	*Included*	*Included*	*Included*	*Included*	*Included*
Contestant fixed effects	*Included*	*Included*	*Included*	*Included*	*Included*
Room fixed effects	*Included*	*Included*	*Included*	*Included*	*Included*
Constant	34.083***	30.531***	9.139***	10.266***	13.109***
	(0.752)	(0.729)	(0.786)	(0.783)	(0.804)
N	1,474,753	1,474,753	1,474,753	1,474,753	1,474,753
R^2	0.608	0.612	0.617	0.617	0.618

Robust standard errors are in parentheses. Two-tailed tests. *** $p < 0.001$. All models include fixed effects for contestant, contest, and room.

4.1. Tests of baseline hypothesis and hypotheses 1 and 2

To test the baseline hypothesis and Hypotheses 1 and 2, we use the regression results in Table 2. Model 1 includes all the control variables. We enter *Prior peer failure feedback* and *Prior machine failure feedback* separately in Models 2 and 3. Both variables have positive and significant coefficients ($\beta = 5.592$, $p < 0.001$ for the former; $\beta = 5.451$, $p <$ 0.001 for the latter). To test the baseline hypothesis and Hypothesis 1, we enter these two variables together in Model 4. The two variables remain positive and significant ($\beta = 2.386$, $p < 0.001$ for the former; $\beta =$ 4.874, $p < 0.001$ for the latter), which is consistent with our baseline hypothesis and Hypothesis 1. An individual learns from *Prior peer failure feedback* and *Prior machine failure feedback*, such that *Prior peer failure feedback* and *Prior machine failure feedback* have a positive effect on the individual's *Performance* in a current competition. In Model 5, we test Hypothesis 2 by adding the interaction variable between *Prior peer failure feedback* and *Prior machine failure feedback*. The coefficient of this interaction is positive and significant ($\beta = 1.064$, $p < 0.001$), suggesting that *Prior machine failure feedback* positively moderates the relationship between *Prior peer failure feedback* and *Performance*, supporting Hypothesis 2.

The plot of this interaction effect is shown in Fig. 2. The dashed line depicts the relationship between *Prior peer failure feedback* and *Performance* when there is a high level of *Prior machine failure feedback*, and the

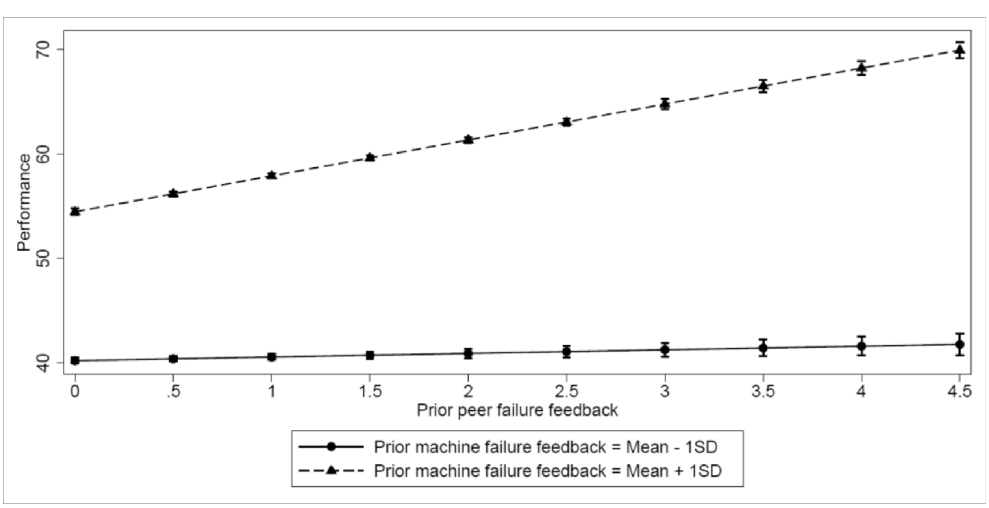

Fig. 2. Plot of the interaction between *Prior machine failure feedback* and *Prior peer failure feedback* in Model 5 of Table 2.

Table 3
OLS fixed-effects models predicting *Performance* (moderator is decomposed into two parts).

	(1)	(2)	(3)	(4)
Past performance	0.045***	0.043***	0.045***	0.042***
	(0.005)	(0.005)	(0.005)	(0.005)
Status	0.629***	0.509***	0.624***	0.510***
	(0.045)	(0.046)	(0.045)	(0.046)
Knowledge specialization	0.432	0.443	0.434	0.437
	(0.620)	(0.619)	(0.620)	(0.617)
Prior unsuccessful hacks from peers	1.461***	1.106***	1.450***	1.083***
	(0.067)	(0.068)	(0.067)	(0.067)
Prior successful hacks to peers	0.638***	0.430***	0.617***	0.496***
	(0.077)	(0.077)	(0.076)	(0.077)
Prior unsuccessful hacks to peers	0.346***	0.085	0.337***	0.076
	(0.082)	(0.083)	(0.082)	(0.083)
Number of successful hacks received by focal contestant in focal contest	−10.425***	−10.519***	−10.423***	−10.547***
	(0.416)	(0.415)	(0.416)	(0.415)
Number of unsuccessful hacks received by focal contestant in focal contest	4.021***	3.916***	4.017***	3.915***
	(0.210)	(0.210)	(0.209)	(0.210)
Number of successful hacks sent by focal contestant in focal contest	8.281***	8.249***	8.278***	8.262***
	(0.231)	(0.231)	(0.231)	(0.231)
Number of unsuccessful hacks sent by focal contestant in focal contest	−2.587***	−2.605***	−2.585***	−2.620***
	(0.202)	(0.202)	(0.201)	(0.202)
Prior machine failure feedback in contest problems	4.714***	4.818***	4.750***	4.646***
	(0.072)	(0.072)	(0.070)	(0.070)
Prior machine failure feedback in gym problems	1.045***	0.855***	0.947***	1.344***
	(0.041)	(0.041)	(0.059)	(0.060)
Prior peer failure feedback	2.061***	−3.090***	2.002***	−3.917***
	(0.097)	(0.264)	(0.106)	(0.268)
Prior machine failure feedback in contest problems * Prior peer failure feedback		1.006***		1.230***
		(0.050)		(0.054)
Prior machine failure feedback in gym problems * Prior peer failure feedback			0.055	−0.301***
			(0.031)	(0.034)
Contest fixed effects	*Included*	*Included*	*Included*	*Included*
Contestant fixed effects	*Included*	*Included*	*Included*	*Included*
Room fixed effects	*Included*	*Included*	*Included*	*Included*
Constant	11.724***	13.631***	11.719***	14.084***
	(0.786)	(0.804)	(0.786)	(0.806)
N	1,474,753	1,474,753	1,474,753	1,474,753
R^2	0.618	0.618	0.618	0.618

Robust standard errors are in parentheses. Two-tailed tests. *** $p < 0.001$. All models include fixed effects for contestant, contest, and room.

solid line demonstrates the same relationship when there is a low level of *Prior machine failure feedback*. The dashed line has a positive slope, whereas the solid line is almost flat, which is consistent with Hypothesis 2 that a higher level of prior machine failure feedback has a catalyzing effect, that is, it enables individuals to learn more from prior peer failure feedback.

4.2. Tests of hypothesis 3

The results in Tables 3 and 4 are used to test Hypothesis 3. In Table 3, we disaggregate the moderator *Prior machine failure feedback* into *Prior machine failure feedback in contest problems* and *Prior machine failure feedback in gym problems*. We enter the interaction variable between *Prior peer failure feedback* and *Prior machine failure feedback in contest problems* and the interaction variable between *Prior peer failure feedback* and *Prior*

Table 4
OLS fixed-effects models predicting *Performance* (moderator is decomposed into two parts).

	(1)	(2)	(3)	(4)
Past performance	0.025***	0.022***	0.029***	0.026***
	(0.005)	(0.005)	(0.005)	(0.005)
Status	0.693***	0.607***	0.483***	0.531***
	(0.045)	(0.046)	(0.045)	(0.045)
Knowledge specialization	0.710	0.667	0.692	0.674
	(0.606)	(0.603)	(0.606)	(0.604)
Prior unsuccessful hacks from peers	1.411***	1.051***	1.071***	1.010***
	(0.067)	(0.067)	(0.067)	(0.067)
Prior successful hacks to peers	0.830***	0.619***	0.540***	0.549***
	(0.076)	(0.076)	(0.075)	(0.075)
Prior unsuccessful hacks to peers	0.341***	0.042	0.079	0.018
	(0.081)	(0.082)	(0.082)	(0.082)
Number of successful hacks received by focal contestant in focal contest	−10.508***	−10.625***	−10.565***	−10.611***
	(0.414)	(0.413)	(0.414)	(0.413)
Number of unsuccessful hacks received by focal contestant in focal contest	4.008***	3.899***	3.908***	3.888***
	(0.210)	(0.209)	(0.209)	(0.209)
Number of successful hacks sent by focal contestant in focal contest	8.308***	8.269***	8.261***	8.260***

(*continued on next page*)

Table 4 (*continued*)

	(1)	(2)	(3)	(4)
	(0.230)	(0.231)	(0.231)	(0.231)
Number of unsuccessful hacks sent by focal contestant in focal contest	−2.627***	−2.656***	−2.635***	−2.649***
	(0.201)	(0.201)	(0.201)	(0.201)
Prior machine failure feedback in problems A and B	2.152***	2.301***	2.593***	2.472***
	(0.076)	(0.075)	(0.073)	(0.074)
Prior machine failure feedback in the rest of the problems	3.065***	2.994***	2.685***	2.826***
	(0.065)	(0.064)	(0.068)	(0.073)
Prior peer failure feedback	1.746***	−3.412***	−1.500***	−3.163***
	(0.095)	(0.229)	(0.185)	(0.230)
Prior machine failure feedback in problems A and B * Prior peer failure feedback		1.112***		0.701***
		(0.049)		(0.081)
Prior machine failure feedback in the rest of the problems * Prior peer failure feedback			0.807***	0.413***
			(0.041)	(0.066)
Contest fixed effects	*Included*	*Included*	*Included*	*Included*
Contestant fixed effects	*Included*	*Included*	*Included*	*Included*
Room fixed effects	*Included*	*Included*	*Included*	*Included*
Constant	15.073***	16.525***	17.721***	17.343***
	(0.772)	(0.782)	(0.795)	(0.794)
N	1,474,753	1,474,753	1,474,753	1,474,753
R^2	0.618	0.619	0.619	0.619

Robust standard errors are in parentheses. Two-tailed tests. *** $p < 0.001$. All models include fixed effects for contestant, contest, and room.

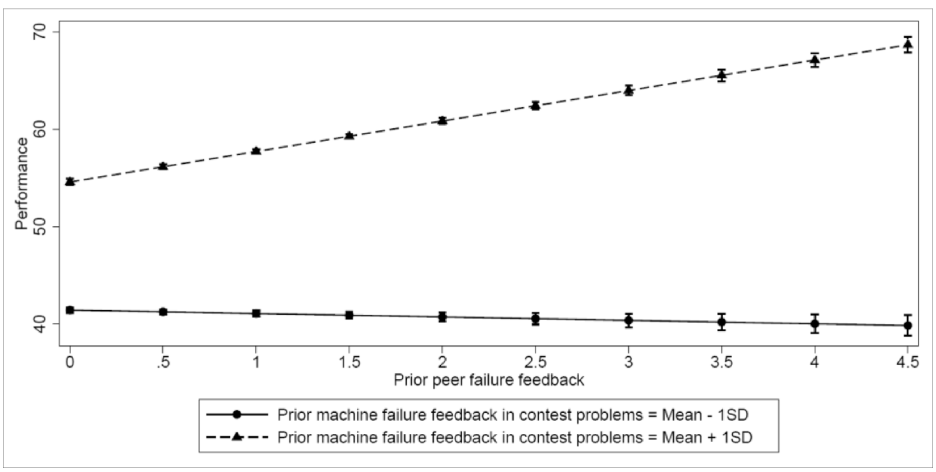

Fig. 3. Plot of the interaction between *Prior machine failure feedback in contest problems* and *Prior peer failure feedback* in Model 4 of Table 3.

machine failure feedback in gym problems in Models 2 and 3. Whereas the coefficient of the former interaction variable is positive and significant ($\beta = 1.006, p < 0.001$), the coefficient of the latter interaction variable is only marginally significant ($\beta = 0.055$, $p = 0.072$). We include the two interaction variables together in Model 4. The coefficient of the former remains positive and significant ($\beta = 1.230$, $p < 0.001$), but the coefficient of the latter is now negative and significant (β = -0.301, $p <$ 0.001).[2] The results in Model 4 suggest that while *Prior machine failure feedback in contest problems* enhances the relationship between *Prior peer failure feedback* and *Performance*, *Prior machine failure feedback in gym problems* does not have such an effect. A Wald test also indicates that these two interactions are significantly different from each other (F = 418.4, $p < 0.001$). These results support Hypothesis 3. Machine failure feedback that is more related to peer failure feedback has a stronger moderating effect on the relationship between peer feedback and learning outcome, such that it enhances the catalyzing effect of prior machine failure feedback.

The implications of the two interaction variables in Model 4 of Table 3 are plotted in Figures 3 and 4. Fig. 3 shows that when there is a high level of machine failure feedback (i.e., dashed line), the relationship between peer failure feedback and performance is positive. However, when there is low level of machine feedback (i.e., solid line), performance almost does not change with peer feedback. Fig. 3 suggests that *Prior machine failure feedback in contest problems* positively moderates the relationship between *Prior peer failure feedback* and *Performance*. However, we observe in Fig. 4 that *Prior machine failure feedback in gym problems* negatively moderates the relationship between *Prior peer failure feedback* and *Performance*. The visualization in Figures 3 and 4 is consistent with Hypothesis 3.

In Table 4, we disaggregate *Prior machine failure feedback* into *Prior machine failure feedback in problems A and B* and *Prior machine failure feedback in the rest of the problems*. The coefficients of both of the resulting interaction variables are positive and significant ($\beta = 1.112$, p < 0.001; $\beta = 0.807$, $p < 0.001$) in Models 2 and 3, and remain so in Model 4 ($\beta = 0.701$, $p < 0.001$; $\beta = 0.413$, $p < 0.001$). However, the moderating effect of *Prior machine failure feedback in problems A and B* is stronger than the moderating effect of *Prior machine failure feedback in the rest of the problems*. For *Prior machine failure feedback in problems A and B*, when this moderator increases from one standard deviation below the mean to one standard deviation above the mean, the slope

[2] We speculate that this negative interaction effect can result from the following: if individuals receive a greater amount of less related machine failure feedback (i.e., failure feedback provided by a machine in gym problems), they might be distracted to learn from peer failure feedback. In comparison, a smaller amount of less related machine failure feedback might not distract individuals' as much to learn from peer failure feedback.

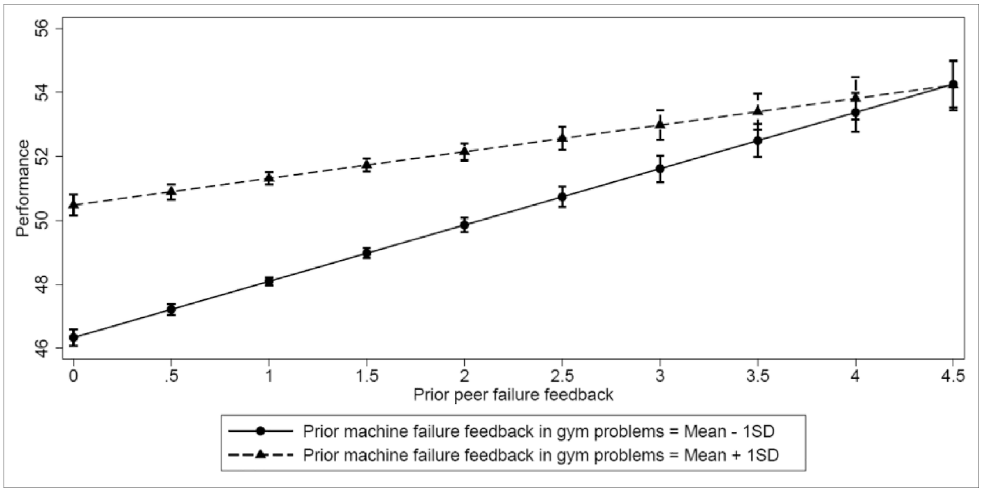

Fig. 4. Plot of the interaction between *Prior machine failure feedback in gym problems* and *Prior peer failure feedback* in Model 4 of Table 3.

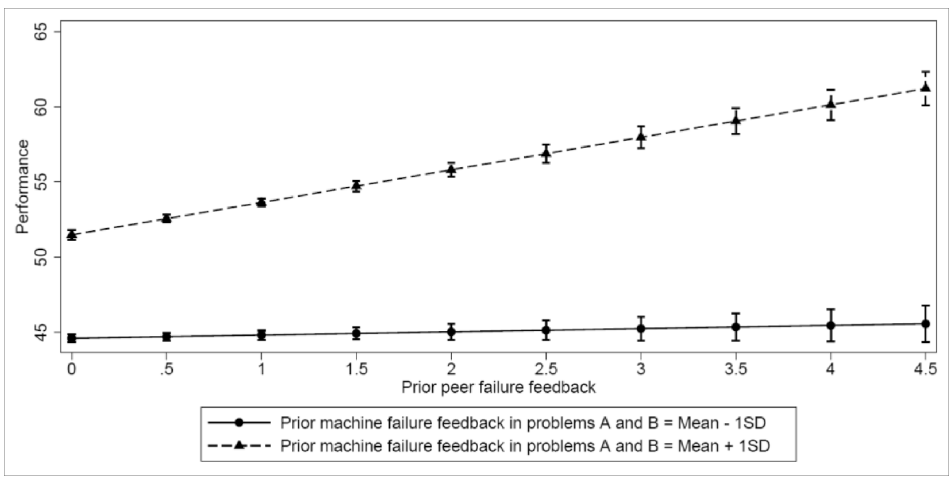

Fig. 5. Plot of the interaction between *Prior machine failure feedback in problems A and B* and *Prior peer failure feedback* in Model 4 of Table 4.

increases by 1.95. In comparison, for *Prior machine failure feedback in the rest of the problems*, when this moderator increases from one standard deviation below the mean to one standard deviation above the mean, the slope increases by 1.49 (a 30.9 % weaker moderating effect). A Wald test also shows that these two interactions are significantly different from each other ($F = 4.2$, $p < 0.05$). These results are consistent with Hypothesis 3.

The implications of the two interactions in Model 4 of Table 4 are plotted in Figs. 5 and 6. Figs. 5 and 6 demonstrate that when there is a high level of machine feedback (dashed line), the relationship between peer feedback and performance is positive. However, when there is low level of machine feedback (solid line), performance hardly changes with peer feedback. As we discussed in the previous paragraph, the moderating effect in Fig. 5 is stronger than the moderating effect in Fig. 6, consistent with Hypothesis 3.

4.3. Additional analyses using an alternative dependent variable

To test our hypotheses, we measured learning outcomes (i.e., performance) using contestants' ranking in contests. Accordingly, the unit of analysis is a contest-contestant observation, based on the contestant's final ranking in these contests. However, a contestant can make multiple submissions during a contest, for the different problems that are part of that contest. Each of these submissions can be a success (i.e., it passes the test) or a failure (i.e., it does not pass the test). Therefore, an alternative way to assess learning outcomes is to investigate the success for each submission made by a contestant for each of the problems that make up a contest. Accordingly, we create an alternative dependent variable, *Success of a submission*, to measure performance in a more granular way. The unit of analysis in this instance is a contest-contestant-submission observation. *Success of a submission* is coded 1 if a particular submission was successful, and 0 if it failed.

To test Hypotheses 1 and 2, we create two variables including *Prior peer failure feedback* and *Prior machine failure feedback*. *Prior peer failure feedback* is measured by the count of the number of times that a contestant was successfully hacked by other contestants before focal submission, whereas *Prior machine failure feedback* is operationalized by counting the number of times that a focal contestant's solutions failed in machine tests before focal submission. These two variables are log-transformed due to their skewed distributions.

To test Hypothesis 3, we disaggregate *Prior machine failure feedback* into two categories, based on the tasks on which that feedback is provided. One category consists of machine failure feedback given on tasks that are more related to tasks on which *Prior peer failure feedback* is provided. The other category consists of machine failure feedback on tasks that are less related to those on which *Prior peer failure feedback* is provided. We use two different ways to perform this disaggregation as we did for the results reported in Tables 3 and 4. First, we disaggregate *Prior machine failure feedback* into *Prior machine failure feedback in contest problems* and *Prior machine failure feedback in gym problems.* While *Prior*

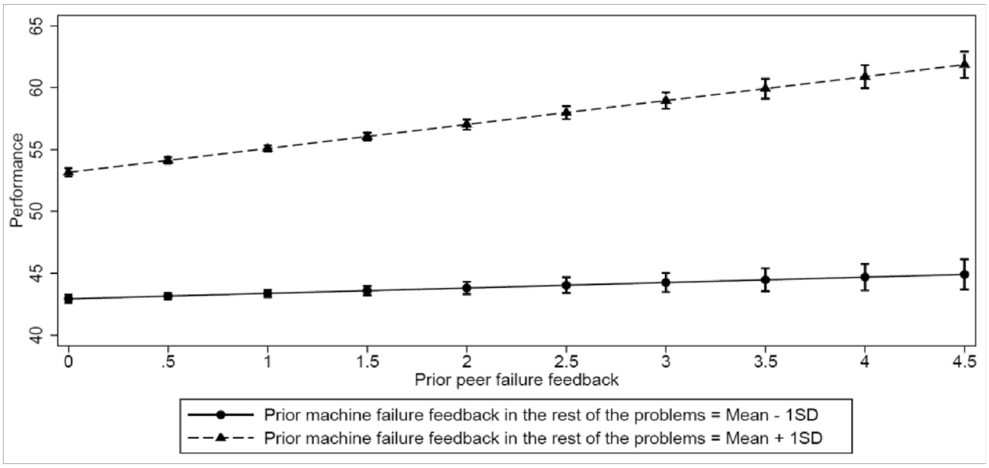

Fig. 6. Plot of the interaction between *Prior machine failure feedback in the rest of the problems* and *Prior peer failure feedback* in Model 4 of Table 4.

machine failure feedback in contest problems is measured by the total number of failed solutions in contest problems before focal submission, *Prior machine failure feedback in gym problems* measures the total number of failed solutions in gym problems before focal submission. Machine failure feedback in contest problems is more related to human failure feedback than machine failure feedback in gym problems because contestants can receive failure feedback from peers (i.e., hacks) only in contest problems, but not in gym problems. Second, we split *Prior machine failure feedback* into *Prior machine failure feedback in problems A and B* and *Prior machine failure feedback in the rest of the problems.* While *Prior machine failure feedback in problems A and B* is measured by the total number of failed solutions of a contestant had in problems A and B before focal submission, *Prior machine failure feedback in the rest of the problems* captures the total number of failed solutions that a contestant had in the rest problems before focal submission. The rationale for this disaggregation is that the majority (72.7 %) of failure feedback given by peers (i.e., hacks) occurs in problems A and B, thus failure feedback given by a machine in problems A and B would be more related to failure feedback given by peers. These four variables are log-transformed due to their skewed distributions.

We controlled for covariates that might impact the dependent variable. First, since our independent variable *Prior peer failure feedback* captures the number of successful hacks from peers before focal submission, we also control for *Prior unsuccessful hacks from peers*, which is the number of unsuccessful hacks from peers before focal submission, *Prior successful hacks to peers*, which is the number of successful hacks from a focal individual to peers before focal submission, and *Prior unsuccessful hacks to peers*, which is the number of unsuccessful hacks from a focal individual to peers before focal submission. These three variables are log-transformed due to their skewed distributions. We control for *Status*, as measured by an individual's Elo ratings. Finally, we control for *Specialization*. To operationalize this variable, we aggregate the tags for all the problems solved by the contestant in past contests, and then sum the number of times that each tag in the current contest appears in that aggregated tag list. *Specialization* is measured as the number of times that tags (for all the problems in the current contest) appear in that aggregated tag list, normalized by the total length of the aggregated list.

The models that we use to estimate this alternative dependent variable are similar to the ones we reported in Tables 2-4. Specifically, the model shown in equation (4) is used to test the baseline hypothesis and Hypothesis 1, in which we propose that prior peer failure feedback and prior machine failure feedback are positively associated with performance.

$$Success_{ijc} = \beta_0 + \beta_1 PeerFeedback_i + \beta_2 MachineFeedback_i + Controls_i + \delta_c + \delta_i + \delta_r + \varepsilon_i \tag{4}$$

where $Success_{ijc}$ refers to whether contestant i's jth submission in contest c is successful or failed, $PeerFeedback_i$ refers to the amount of peer failure feedback that contestant i received before contest c, $MachineFeedback_i$ refers to the amount of machine failure feedback that contestant i received before contest c, δ_c, δ_i, and δ_r refer to contest, contestant, and room fixed effects, and ε_i refers to random error term. If β_1 and β_2 are positive and significant, then the baseline hypothesis and Hypothesis 1 are supported.

The model shown in equation (5) is used to test Hypothesis 2, in which we posit that prior machine failure feedback strengthens the positive relationship between prior peer failure feedback and performance.

$$\begin{aligned} Success_{ijc} = {} & \beta_0 + \beta_1 PeerFeedback_i + \beta_2 MachineFeedback_i \\ & + \beta_3 MachineFeedback_i {}^* PeerFeedback_i + Controls_i \\ & + \delta_c + \delta_i + \delta_r + \varepsilon_i \end{aligned} \tag{5}$$

where $Success_{ijc}$ refers to whether contestant i's jth submission in contest c is successful or failed, $PeerFeedback_i$ refers to the amount of peer failure feedback that contestant i received before contest c, $MachineFeedback_i$ refers to the amount of machine failure feedback that contestant i received before contest c, $MachineFeedback_i{}^*PeerFeedback_i$ refers to the interaction term between prior machine failure feedback and prior peer failure feedback, δ_c, δ_i, and δ_r refer to contest, contestant, and room fixed effects, and ε_i refers to random error term. Hypothesis 2 is supported if β_3 is positive and significant.

The model shown in equation (6) is used to test Hypothesis 3, in which we propose that prior machine failure feedback has a stronger moderating effect on the relationship between prior peer failure feedback and performance if it is more related to prior peer failure feedback.

$$\begin{aligned} Success_{ijc} = {} & \beta_0 + \beta_1 PeerFeedback_i + \beta_2 MachineFeedbackMoreRelated_i \\ & + \beta_3 MachineFeedbackLessRelated_i \\ & + \beta_4 MachineFeedbackMoreRelated_i {}^* PeerFeedback_i \\ & + \beta_5 MachineFeedbackLessRelated_i {}^* PeerFeedback_i \\ & + Controls_i + \delta_c + \delta_i + \delta_r + \varepsilon_i \end{aligned} \tag{6}$$

where $Success_{ijc}$ refers to whether contestant i's jth submission in contest c is successful or failed, $PeerFeedback_i$ refers to the amount of peer failure feedback that contestant i received before contest c, $MachineFeedbackMoreRelated_i$ refers to machine failure feedback that is more related to peer failure feedback, $MachineFeedbackLessRelated_i$ refers to machine failure feedback that is less related to peer failure feedback, $MachineFeedbackMoreRelated_i{}^*PeerFeedback_i$ refers to the interaction term between prior machine failure feedback that is more

Table 5
OLS fixed-effects models predicting *Success of a submission.*

	(1)	(2)	(3)	(4)	(5)
Prior unsuccessful hacks from peers	0.0397***	0.0173***	0.0152***	0.0076***	0.0035***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Prior successful hacks to peers	0.0078***	0.0066***	0.0084***	0.0078***	0.0049***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Prior unsuccessful hacks to peers	0.0117***	0.0045***	0.0045***	0.0020*	−0.0010
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Status	−0.0011*	−0.0011*	−0.0003	−0.0003	−0.0019***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Specialization	−0.0314***	−0.0101+	0.1288***	0.1205***	0.1072***
	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)
Prior peer failure feedback		0.0497***		0.0231***	−0.0294***
		(0.001)		(0.001)	(0.003)
Prior machine failure feedback			0.0447***	0.0396***	0.0401***
			(0.001)	(0.001)	(0.001)
Prior peer failure feedback * Prior machine failure feedback					0.0106***
					(0.001)
Contest fixed effects	*Included*	*Included*	*Included*	*Included*	*Included*
Contestant fixed effects	*Included*	*Included*	*Included*	*Included*	*Included*
Room fixed effects	*Included*	*Included*	*Included*	*Included*	*Included*
Constant	0.4132***	0.3836***	0.2074***	0.2170***	0.2443***
	(0.008)	(0.008)	(0.010)	(0.010)	(0.009)
N	6,237,859	6,237,859	6,237,859	6,237,859	6,237,859
R^2	0.322	0.323	0.324	0.324	0.324

Robust standard errors are in parentheses. Two-tailed tests. *** $p < 0.001$.
All models include fixed effects for contestant, contest, room, and problem.

Table 6
OLS fixed-effects models predicting *Success of a submission* (moderator is decomposed into two parts).

	(1)	(2)	(3)	(4)
Status	−0.0007	−0.0020***	−0.0009	−0.0020***
	(0.001)	(0.001)	(0.001)	(0.001)
Specialization	0.1124***	0.1038***	0.1132***	0.1019***
	(0.007)	(0.007)	(0.007)	(0.007)
Prior unsuccessful hacks from peers	0.0085***	0.0043***	0.0082***	0.0040***
	(0.001)	(0.001)	(0.001)	(0.001)
Prior successful hacks to peers	0.0069***	0.0045***	0.0062***	0.0049***
	(0.001)	(0.001)	(0.001)	(0.001)
Prior unsuccessful hacks to peers	0.0017	−0.0012	0.0014	−0.0013
	(0.001)	(0.001)	(0.001)	(0.001)
Prior machine failure feedback in contest problems	0.0306***	0.0323***	0.0315***	0.0316***
	(0.001)	(0.001)	(0.001)	(0.001)
Prior machine failure feedback in gym problems	0.0073***	0.0051***	0.0042***	0.0081***
	(0.000)	(0.000)	(0.001)	(0.001)
Prior peer failure feedback	0.0253***	−0.0328***	0.0234***	−0.0381***
	(0.001)	(0.003)	(0.001)	(0.003)
Prior machine failure feedback in contest problems * Prior peer failure feedback		0.0112***		0.0126***
		(0.001)		(0.001)
Prior machine failure feedback in gym problems * Prior peer failure feedback			0.0017***	−0.0019***
			(0.000)	(0.000)
Contest fixed effects	*Included*	*Included*	*Included*	*Included*
Contestant fixed effects	*Included*	*Included*	*Included*	*Included*
Room fixed effects	*Included*	*Included*	*Included*	*Included*
Constant	0.2467***	0.2658***	0.2471***	0.2677***
	(0.009)	(0.010)	(0.009)	(0.010)
N	6,237,859	6,237,859	6,237,859	6,237,859
R^2	0.324	0.324	0.324	0.324

Robust standard errors are in parentheses. Two-tailed tests. *** $p < 0.001$. All models include fixed effects for contestant, contest, room, and problem.

related to peer failure feedback and prior peer failure feedback, $MachineFeedbackLessRelated_i{*}PeerFeedback_i$ refers to the interaction term between prior machine failure feedback that is less related to peer failure feedback and prior peer failure feedback, δ_c, δ_i, and δ_r refer to contest, contestant, and room fixed effects, and ε_i refers to random error term. Hypothesis 3 is supported if β_4 has a stronger positive effect than β_5.

Table 5 presents the results we use to test the baseline hypothesis, as well as Hypotheses 1 and 2. The sequence that we enter the variables in Table 5 is the same as in Table 2. Fig. A1 (Appendix A) plots the interaction in Model 5 in Table 5. Tables 6 and 7 present the results we use to test Hypothesis 3. The sequence that we enter the variables in Tables 6 and 7 is the same as in Tables 3 and 4. Figs. A2 and A3 (Appendix A) plot the interactions in Model 4 of Table 6. Figs. A4 and A5 (Appendix A) plot the interactions in Model 4 of Table 7.

Overall, the results in Tables 5-7 are consistent with those in Tables 2-4. In particular, Model 4 in Table 5 suggests that both prior peer failure feedback and prior machine failure feedback are positively

Table 7
OLS fixed-effects models predicting *Success of a submission* (moderator is decomposed into two parts).

	(1)	(2)	(3)	(4)
Status	0.0004	−0.0006	−0.0012	−0.0000
	(0.001)	(0.001)	(0.001)	(0.000)
Specialization	0.1111***	0.0972***	0.1045***	0.0971***
	(0.007)	(0.007)	(0.007)	(0.007)
Prior unsuccessful hacks from peers	0.0083***	0.0031***	0.0051***	0.0035***
	(0.001)	(0.001)	(0.001)	(0.001)
Prior successful hacks to peers	0.0081***	0.0051***	0.0056***	0.0057***
	(0.001)	(0.001)	(0.001)	(0.001)
Prior unsuccessful hacks to peers	0.0021	−0.0019	−0.0002	−0.0017
	(0.001)	(0.001)	(0.001)	(0.001)
Prior machine failure feedback in problems A and B	0.0300***	0.0323***	0.0334***	0.0311***
	(0.001)	(0.001)	(0.001)	(0.001)
Prior machine failure feedback in the rest of the problems	0.0081***	0.0075***	0.0055***	0.0087***
	(0.001)	(0.001)	(0.001)	(0.001)
Prior peer failure feedback	0.0228***	−0.0449***	−0.0072**	−0.0466***
	(0.001)	(0.003)	(0.002)	(0.003)
Prior machine failure feedback in problems A and B * Prior peer failure feedback		0.0145***		0.0182***
		(0.001)		(0.001)
Prior machine failure feedback in the rest of the problems * Prior peer failure feedback			0.0071***	−0.0037***
			(0.000)	(0.001)
Contest fixed effects	*Included*	*Included*	*Included*	*Included*
Contestant fixed effects	*Included*	*Included*	*Included*	*Included*
Room fixed effects	*Included*	*Included*	*Included*	*Included*
Constant	0.2301***	0.2468***	0.2531***	0.2392***
	(0.009)	(0.010)	(0.009)	(0.009)
N	6,237,859	6,237,859	6,237,859	6,237,859
R^2	0.324	0.324	0.324	0.324

Robust standard errors are in parentheses. Two-tailed tests. ** $p < 0.01$, *** $p < 0.001$. All models include fixed effects for contestant, contest, room, and problem.

Table 8
Robustness check following the method in Bennett and Snyder (2017).

Time window	Peer failure feedback	Baseline hypothesis	Machine failure feedback	H1
One-year	1.517***	Supported	3.612***	Supported
	(0.075)		(0.054)	
Six-month	1.353***	Supported	2.700***	Supported
	(0.066)		(0.043)	
Three-month	1.023***	Supported	1.955***	Supported
	(0.062)		(0.035)	

These models are estimated with the same set of control variables and fixed effects (including contest, contestant, and room fixed effects) as we report in Tables 2 to 4. Robust standard errors are in parentheses. Two-tailed tests. *** $p < 0.001$.

associated with performance, which supports the baseline hypothesis and Hypothesis 1. In addition, Model 5 in Table 5 shows that the interaction between prior peer failure feedback and prior machine failure feedback is positive and significant, providing support to Hypothesis 2. The results in Tables 6 and 7 are used to test Hypothesis 3. This consistency is in line with our expectation, since participants' final contest performance (the outcome we used for our main analyses) is mainly a function of the accuracy of their individual submissions throughout a given contest (as captured in this more granular alternative dependent variable). As participants' final performance decreases as the number of their failed submissions increases, we also consider the success of each of their submissions as a viable alternative measure of their performance that enables us to zoom in on a more granular performance-related event.

In Table 6, we disaggregate *Prior machine failure feedback* into *Prior machine failure feedback in contest problems* and *Prior machine failure feedback in gym problems*. The former is more related to prior peer failure feedback than the latter. The results in Model 4 of Table 6 suggest that while *Prior machine failure feedback in contest problems* enhances the relationship between *Prior peer failure feedback* and *Performance*, *Prior machine failure feedback in gym problems* does not have such an effect. A Wald test indicates that these two interactions are significantly different from each other ($F = 456.2, p < 0.001$). These results support Hypothesis 3. In Table 7, we disaggregate *Prior machine failure feedback* into *Prior machine failure feedback in problems A and B* and *Prior machine failure feedback in the rest of the problems*. The former is more related to prior peer failure feedback than the latter. The results in Model 4 of Table 7 suggest that while *Prior machine failure feedback in problems A and B* enhances the relationship between *Prior peer failure feedback* and *Performance*, *Prior machine failure feedback in the rest of the problems* does not have such an effect. A Wald test indicates that these two interactions are significantly different from each other ($F = 481.3$, $p < 0.001$). These results again support Hypothesis 3. Taken together, the results using this alternative dependent variable support our hypotheses.

4.4. Robustness checks

First, in the results that we present in Tables 2 to 4, we decided to not

Table 9
Robustness check following the method in Bennett and Snyder (2017).

Time window	Machine failure feedback * Peer failure feedback	H2	Machine failure feedback in contest problems * Peer failure feedback	Machine failure feedback in gym problems * Peer failure feedback	Machine failure feedback in problems A and B * Peer failure feedback	Machine failure feedback in the rest problems * Peer failure feedback	H3
One-year	0.870*** (0.044)	Supported	1.418*** (0.052)	−0.304*** (0.029)	0.747*** (0.073)	0.464*** (0.057)	Supported
Six-month	0.733*** (0.042)	Supported	1.366*** (0.051)	−0.217*** (0.027)	0.692*** (0.064)	0.448*** (0.052)	Supported
Three-month	0.509*** (0.045)	Supported	1.124*** (0.051)	−0.158*** (0.028)	0.635*** (0.062)	0.279*** (0.055)	Supported

These models are estimated with the same set of control variables and fixed effects (including contest, contestant, and room fixed effects) as we report in Tables 2 to 4. Robust standard errors are in parentheses. Two-tailed tests. *** $p < 0.001$.

control for success experience, or success feedback given by a machine, since success feedback from a machine and failure feedback from a machine are highly correlated ($r = 0.94$). This happens because both success and failure feedback go up with an individual's overall experience in entering contests. As a result of the high correlation, including success feedback in the models yields multicollinearity concerns. However, since success experience and failure experience can have distinct effects on learning outcomes we ran a robustness test by adding success feedback to all the models reported in Tables 2 to 4. The results we get from these models are consistent with what we present in Tables 2 to 4 and continue to support all of the hypotheses ($p < 0.001$).

Second, Bennett and Snyder (2017) suggest that a significant relationship between failure experience and learning outcomes could be induced by certain model specifications. Specifically, the authors suggest that regressing learning outcomes on *cumulative* failure experience is problematic. Following their recommendation, to calculate failure feedback we use shorter time windows, rather than a contestant's entire history in all previous contests. Whereas in the main models that are reported in Tables 2-4 we calculate failure feedback by counting all failure feedback prior to the current contest, in this robustness check, we recalculate failure feedback (for both peer failure feedback and machine failure feedback) using one-year, six-month, and three-month moving windows. We present the results in Tables 8 and 9. These results show that our baseline hypothesis and Hypotheses 1–3 continue to receive support if we use these shorter time windows to calculate failure feedback, alleviating the concern that our findings might be mechanically induced by the model specification.

5. Discussion and conclusion

Studies on machine-human interaction have shown that the presence of machines can benefit humans in various ways, such as by further engaging and motivating people in their activities (Colombo et al., 2007). Nevertheless, evidence on whether and how machines impact humans' learning remains relatively scarce. This is an important question especially in light of machines' increasing presence in workplaces. We explore this question and find that (i) failure feedback provided by machines directly contributes to individuals' learning; (ii) failure feedback provided by machines also amplifies learning from failure feedback provided by peers, i.e. the more that individuals receive failure feedback from machines, the more that they learn from failure feedback given by peers as well, and (iii) this catalyzing effect is stronger when the tasks on which individuals receive machine failure feedback are more closely related to the tasks on which they receive failure feedback from peers.

5.1. Contributions of our study

Our study makes two main contributions. First, it adds to the learning literature (Zou, Ertug, & George, 2018). While this literature's traditional focus is on the process of learning from other individuals, we demonstrate that machine failure feedback plays a key role on individual learning. Our analyses suggest that machine failure feedback facilitates learning both directly – by leading individuals to learn from their failures – and indirectly – by amplifying the learning effect of failure feedback provided by other humans. Second, a separate and broader literature has shown that machines may shape human–human interactions, such as coordination (Hinds et al., 2004; Kiesler & Hinds, 2004; Shirado & Christakis, 2017), conflict (Jung et al., 2015), and communication (Traeger et al., 2020). Our study adds to this literature by showing that machines can impact the dynamics of learning between individuals, via its catalyzing effect. Our research opens avenues for future research to continue to explore how machines can influence other aspects of human learning, both directly on their own and indirectly, in terms of how they influence the learning process between humans.

5.2. Practical implications

Our findings suggest that the use of machines in organizations to provide failure feedback to humans can help people learn more from feedback provided by their own peers. Machines can distinguish spurious failures from true ones, a distinction that is underexplored in failure learning research (Dahlin et al., 2018). The machines in our setting are designed by experts to catch almost all possible mistakes in contestants' code. As a result, contestants trust the judgement of machines. This underscores the importance of programming or designing machines such that they are a "competent" and "trustworthy" source of failure feedback to humans. Our findings are that machine failure feedback has both a direct effect on individuals' learning as well as a catalyzing effect on their learning from failure feedback given by humans. From a practical point of view, these findings suggest that organizations should provide both machine feedback and human feedback to their employees, and that machine feedback could help address biases that might hamper human-to-human learning. For example, as part of their annual performance feedback, organizations can implement an information system to provide performance feedback to employees based on rich data, such as sales, project progress, and number of patents filed, in addition to the currently common practice of providing performance feedback from employees' supervisors. As many organizations are undergoing a digital transformation, this is feasible because such data will become increasingly available.

5.3. Limitations and future research

The characteristics of our setting point to questions that can be explored by future research. For example, machines in our setting have no variation in terms of their attributes. Therefore, we cannot investigate questions about what kind of machines might enable individuals to learn even more from peer failure feedback. Recent research has found that vulnerable machines (machines making vulnerable statements) can enhance human-to-human communication more than neutral machines (machines making neutral statements) (Traeger et al., 2020). Accordingly, our findings can be expanded and refined in contexts where the attributes of machines vary. We show that machines can fuel humans' learning by providing feedback. Future research can explore other channels through which machines can facilitate humans' learning. For example, machines might also train humans to identify their own errors, which is an important precondition for individuals to take action to learn from their errors.

In addition, machines in our setting provide only one type of failure feedback, indicating that there is an error in contestants' coding. Machines can be designed to give other types of feedback as well. It is unclear whether providing only one type of feedback or multiple types of feedback would better enhance individuals' learning. It might be that multiple types of feedback enhance the learning implications of each other, such that they amplify learning overall, or they might distract individuals from each mistake, such that they undermine learning overall. Another characteristic of our setting is that the information that is in machine failure feedback and peer failure feedback does not vary. The identification of mistakes is done in the same way, with a similar amount of text, in particular by providing an example of input which the contestant's code does not process as intended, but without pointing out the exact error in the code. It might be that more detailed failure feedback, or failure feedback containing more information about the possible cause of failure and suggestions about how to remedy it, are better for learning, as these might make it easier to identify and understand the cause of failure.

While we use contestants' rank in a contest to measure their performance and to test our hypotheses, we are aware that there are other ways to capture contestants' performance, including the success of submissions made by contestants. Even though we also tested our hypotheses using that specific alternative measure of performance, both measures might be equally appropriate to capture performance in our context, we are not able to determine which measure might be *the* best measure in this context to capture performance. Future research can use different means to identify which performance measures are the most suitable ones in their setting, including the one we study, and investigate the generalizability and robustness of our results based on different performance metrics.

Finally, while our focus in this study is on learning from failure, machines can also support learning from success feedback. We investigated how machine failure feedback promotes individuals' learning from failure feedback provided by humans. Future research can look into how machine success feedback and human success feedback interact with each other to influence humans' learning.

In conclusion, we propose and find that machines have both a direct effect and a catalyzing effect on individuals' learning. The catalyzing effect refers to the finding that the more that humans receive failure feedback from machines, the more that they also learn from failures identified by human feedback. In addition, the more that the machine failure feedback is received for tasks that are closely related to tasks for which peer failure feedback is received, the more that it amplifies humans' learning from that peer failure feedback. Our findings suggest that machines can be used in organizations to foster learning from failures.

CRediT authorship contribution statement

Tengjian Zou: Writing – review & editing, Writing – original draft, Methodology, Data curation, Conceptualization. **Gokhan Ertug:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Thomas Roulet:** Writing – review & editing, Writing – original draft, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

Tengjian Zou acknowledges the support by the National Natural Science Foundation of China (72202206, 72091312, 72232009) and the support by the Fundamental Research Funds for the Central Universities.

Appendix A. Plots for the interactions in Tables 5–7

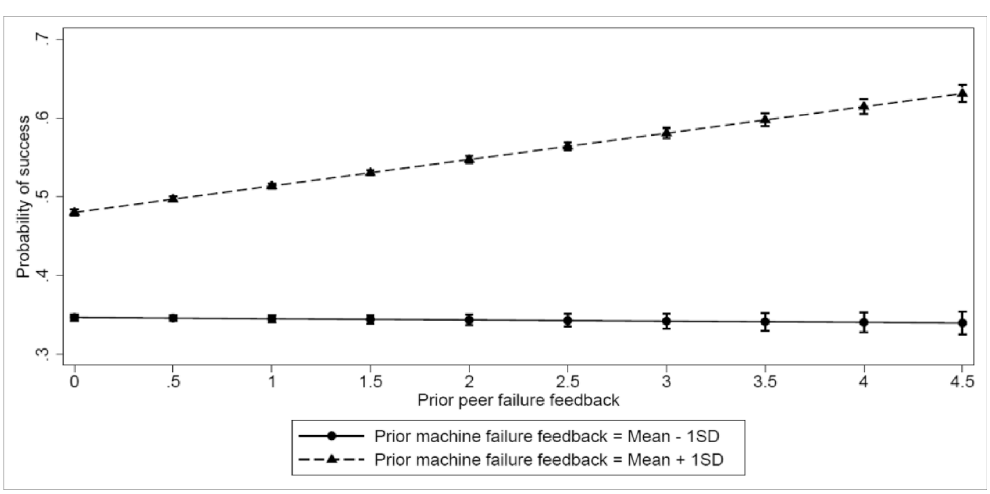

Fig. A1. Plot of the interaction between *Prior machine failure feedback* and *Prior peer failure feedback* in Model 5 of Table 5.

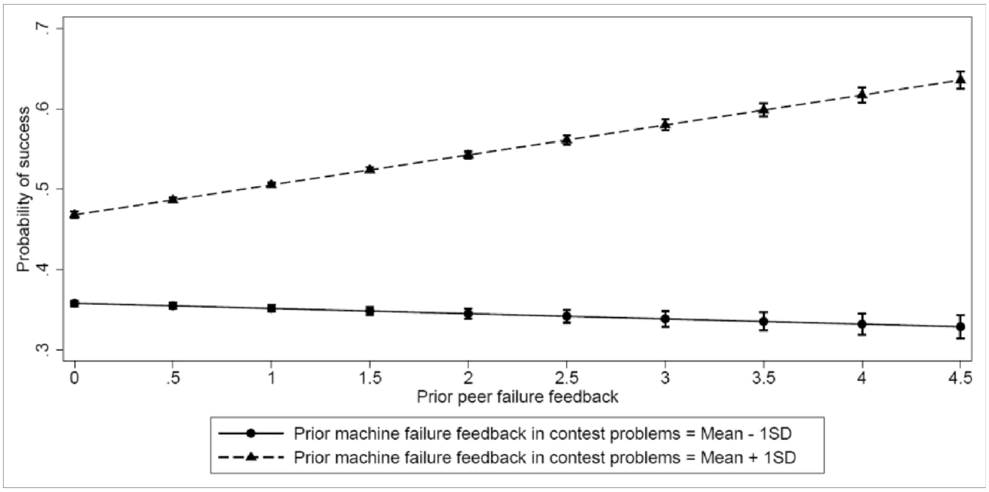

Fig. A2. Plot of the interaction between *Prior machine failure feedback in contest problems* and *Prior peer failure feedback* in Model 4 of Table 6.

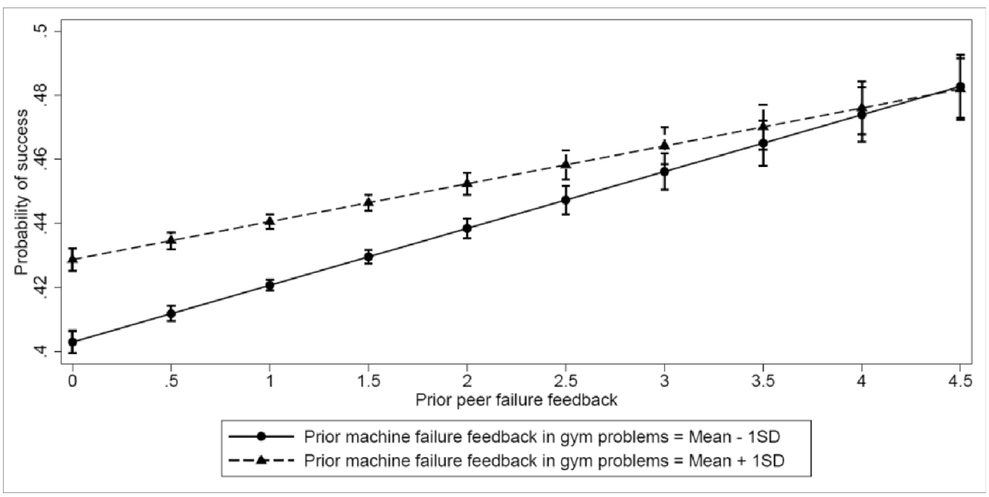

Fig. A3. Plot of the interaction between *Prior machine failure feedback in gym problems* and *Prior peer failure feedback* in Model 4 of Table 6.

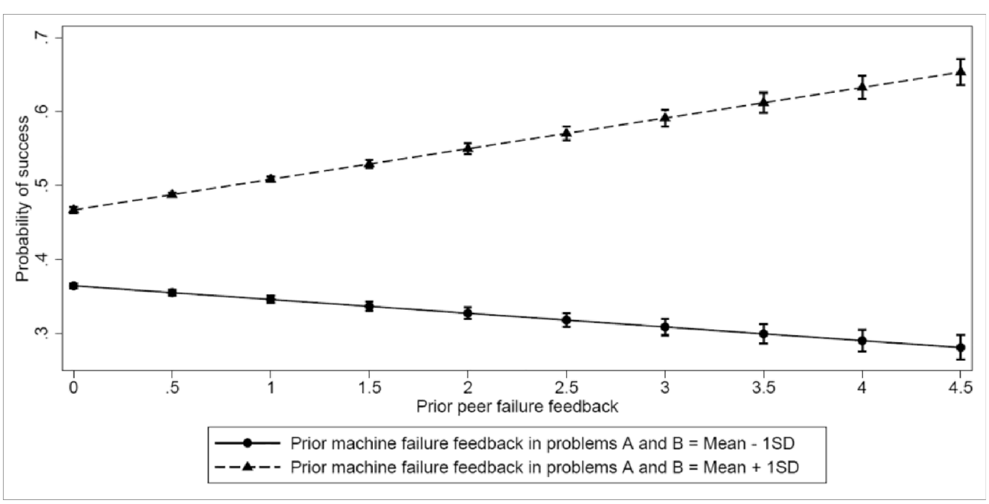

Fig. A4. Plot of the interaction between *Prior machine failure feedback in problems A and B* and *Prior peer failure feedback* in Model 4 of Table 7.

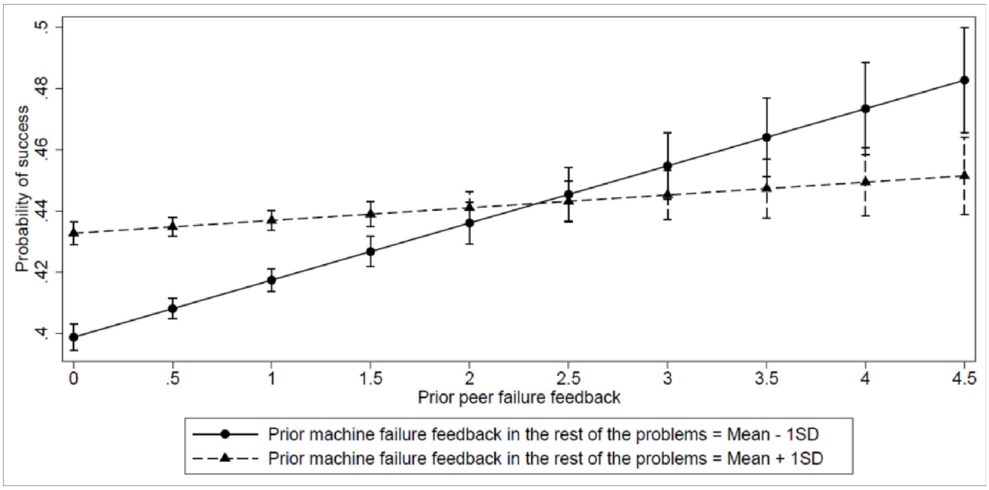

Fig. A5. Plot of the interaction between *Prior machine failure feedback in the rest of the problems* and *Prior peer failure feedback* in Model 4 of Table 7.

.

References

Amankwah-Amoah, J., Adomako, S., & Berko, D. O. (2022). Once bitten, twice shy? The relationship between business failure experience and entrepreneurial collaboration. *Journal of Business Research, 139*, 983–992.

Arbel, Y., Murphy, A., & Donchin, E. (2014). On the utility of positive and negative feedback in a paired-associate learning task. *Journal of Cognitive Neuroscience, 26*(7), 1445–1453.

Ashford, S. J., & Tsui, A. S. (1991). Self-regulation for managerial effectiveness: The role of active feedback seeking. *Academy of Management Journal, 34*(2), 251–280.

Baum, J. A., & Dahlin, K. B. (2007). Aspiration performance and railroads' patterns of learning from train wrecks and crashes. *Organization Science, 18*(3), 368–385.

Bennett, V. M., & Snyder, J. (2017). The empirics of learning from failure. *Strategy Science, 2*(1), 1–12.

Bhattacharjee, S., Gopal, R. D., Lertwachara, K., Marsden, J. R., & Telang, R. (2007). The effect of digital sharing technologies on music markets: A survival analysis of albums on ranking charts. *Management Science, 53*(9), 1359–1374.

Boso, N., Adeleye, I., Donbesuur, F., & Gyensare, M. (2019). Do entrepreneurs always benefit from business failure experience? *Journal of Business Research, 98*, 370–379.

Burgers, J. H., Van Den Bosch, F. A., & Volberda, H. W. (2008). Why new business development projects fail: Coping with the differences of technological versus market knowledge. *Long Range Planning, 41*(1), 55–73.

Cannon, M. D., & Edmondson, A. C. (2001). Confronting failure: Antecedents and consequences of shared beliefs about failure in organizational work groups. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior, 22*(2), 161–177.

Colombo, R., Pisano, F., Mazzone, A., Delconte, C., Micera, S., Carrozza, M. C., … Minuco, G. (2007). Design strategies to improve patient motivation during robot-aided rehabilitation. *Journal of Neuroengineering and Rehabilitation, 4*(1), 33–39.

Dahlin, K. B., Chuang, Y. T., & Roulet, T. J. (2018). Opportunity, motivation, and ability to learn from failures and errors: Review, synthesis, and ways to move forward. *Academy of Management Annals, 12*(1), 252–277.

Ellis, S., & Davidi, I. (2005). After-event reviews: Drawing lessons from successful and failed experience. *Journal of Applied Psychology, 90*(5), 857–871.

Ellis, S., Mendel, R., & Nir, M. (2006). Learning from successful and failed experience: The moderating role of kind of after-event review. *Journal of Applied Psychology, 91*(3), 669–680.

Ellis, S., Carette, B., Anseel, F., & Lievens, F. (2014). Systematic reflection: Implications for learning from failures and successes. *Current Directions in Psychological Science, 23*(1), 67–72.

Eskreis-Winkler, L., & Fishbach, A. (2019). Not learning from failure—the greatest failure of all. *Psychological Science, 30*(12), 1733–1744.

Fedor, D. B., Davis, W. D., Maslyn, J. M., & Mathieson, K. (2001). Performance improvement efforts in response to negative feedback: The roles of source power and recipient self-esteem. *Journal of Management, 27*(1), 79–97.

Groysberg, B., Lee, L. E., & Nanda, A. (2008). Can they take it with them? The portability of star knowledge workers' performance. *Management Science, 54*(7), 1213–1230.

Ham, J., & Midden, C. J. (2014). A persuasive robot to stimulate energy conservation: The influence of positive and negative social feedback and task similarity on energy-consumption behavior. *International Journal of Social Robotics, 6*(2), 163–171.

Hinds, P. J., Roberts, T. L., & Jones, H. (2004). Whose job is it anyway? A study of human-robot interaction in a collaborative task. *Human-Computer Interaction, 19*(1–2), 151–181.

Hodgkinson, G. P., & Wright, G. (2002). Confronting strategic inertia in a top management team: Learning from failure. *Organization Studies, 23*(6), 949–977.

Ilgen, D., & Davis, C. (2000). Bearing bad news: Reactions to negative performance feedback. *Applied Psychology, 49*(3), 550–565.

Jordan, A. H., & Audia, P. G. (2012). Self-enhancement and learning from performance feedback. *Academy of Management Review, 37*(2), 211–231.

Jung, M. F., Martelaro, N., & Hinds, P. J. (2015), March. Using robots to moderate team conflict: the case of repairing violations. In *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction* (pp. 229-236).

Kc, D., Staats, B. R., & Gino, F. (2013). Learning from my success and from others' failure: Evidence from minimally invasive cardiac surgery. *Management Science, 59*(11), 2435–2449.

Kim, L. (1998). Crisis construction and organizational learning: Capability building in catching-up at Hyundai Motor. *Organization Science, 9*(4), 506–521.

Kim, J. Y., & Miner, A. S. (2007). Vicarious learning from the failures and near-failures of others: Evidence from the US commercial banking industry. *Academy of Management Journal, 50*(3), 687–714.

Kiesler, S., & Hinds, P. (2004). Introduction to this special issue on human-robot interaction. *Human-Computer Interaction, 19*(1–2), 1–8.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development.* Englewood Cliffs, NJ: Prentice Hall.

Kuhnen, C. M., & Tymula, A. (2012). Feedback, self-esteem, and performance in organizations. *Management Science, 58*(1), 94–113.

Luo, X., Qin, M. S., Fang, Z., & Qu, Z. (2021). Artificial Intelligence Coaches for Sales Agents: Caveats and Solutions. *Journal of Marketing, 85*(2), 14–32.

Merle, A., St-Onge, A., & Sénécal, S. (2022). Does it pay to be honest? The effect of retailer-provided negative feedback on consumers' product choice and shopping experience. *Journal of Business Research, 147*, 532–543.

Midden, C., & Ham, J. (2009). Using negative and positive social feedback from a robotic agent to save energy. In *Proceedings of the 4th international conference on persuasive technology* (pp. 1-6).

Müller, S., Rothermund, K., & Wentura, D. (2016). Relevance drives attention: Attentional bias for gain-and loss-related stimuli is driven by delayed disengagement. *Quarterly Journal of Experimental Psychology, 69*(4), 752–763.

Piazza, A., & Castellucci, F. (2014). Status in organization and management theory. *Journal of Management, 40*(1), 287–315.

Rahwan, I., Crandall, J. W., & Bonnefon, J. F. (2020). Intelligent machines as social catalysts. *Proceedings of the National Academy of Sciences, 117*(14), 7555–7557.

Schilling, M. A., Vidal, P., Ployhart, R. E., & Marangoni, A. (2003). Learning by doing something else: Variation, relatedness, and the learning curve. *Management Science, 49*(1), 39–56.

Shepherd, D. A., Patzelt, H., & Wolfe, M. (2011). Moving forward from project failure: Negative emotions, affective commitment, and learning from the experience. *Academy of Management Journal, 54*(6), 1229–1259.

Shirado, H., & Christakis, N. A. (2017). Locally noisy autonomous agents improve global human coordination in network experiments. *Nature, 545*(7654), 370–374.

Sitkin, S. B. (1992). Learning through failure: The strategy of small losses. *Research in Organizational Behavior, 14*, 231–266.

Tourish, D., & Hargie, O. (2012). Metaphors of failure and the failures of metaphor: A critical study of root metaphors used by bankers in explaining the banking crisis. *Organization Studies, 33*(8), 1045–1069.

Traeger, M. L., Sebo, S. S., Jung, M., Scassellati, B., & Christakis, N. A. (2020). Vulnerable robots positively shape human conversational dynamics in a human–robot team. *Proceedings of the National Academy of Sciences, 117*(12), 6370–6375.

Ucbasaran, D., Shepherd, D. A., Lockett, A., & Lyon, S. J. (2013). Life after business failure: The process and consequences of business failure for entrepreneurs. *Journal of Management, 39*(1), 163–202.
Van der Kleij, F. M., Feskens, R. C., & Eggen, T. J. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis. *Review of Educational Research, 85*(4), 475–511.
Wang, K., & Nickerson, J. V. (2019). A Wikipedia-based method to support creative idea generation: The Role of Stimulus Relatedness. *Journal of Management Information Systems, 36*(4), 1284–1312.
Wilhelm, H., Richter, A. W., & Semrau, T. (2019). Employee learning from failure: A team-as-resource perspective. *Organization Science, 30*(4), 694–714.
Zellmer-Bruhn, M., & Gibson, C. (2006). Multinational organization context: Implications for team learning and performance. *Academy of Management Journal, 49*(3), 501–518.
Zou, T., Ertug, G., & George, G. (2018). The capacity to innovate: A meta-analysis of absorptive capacity. *Innovation, 20*(2), 87–121.

Tengjian Zou is Assistant Professor at the School of Management, Zhejiang University. He received his PhD from the Lee Kong Chian School of Business, Singapore Management University. His research interests include social networks, status and learning.

Gokhan Ertug is Professor at Adelaide Business School, the University of Adelaide. He received his PhD from INSEAD. His research interests include social networks, status, and organization theory more generally.

Thomas J. Roulet is the Professor of Organisational Sociology & Leadership at the Judge Business School, and a Fellow at King's College, both at the University of Cambridge. His work provides institutional and stakeholder perspectives on negative evaluations (stigma, scandals), failure learning and social change.