Online Supplemental Material for:

**The Exaggerated Benefits of Failure**

**This file includes:**

Pilot Study in Footnote 1 of Main Manuscript

Supplemental Studies S1-S2

Supplementary Analyses for Study 2b

Pilot for Study 6

Additional Studies in the General Discussion

Figures with Data Points

References

**Pilot Study in Footnote 1 of Main Manuscript**

Failure can be defined in different ways. Across studies, we define failure as failing to achieve a desired goal. Using this definition, failure encompasses situations in which goal achievement falls short due to a mistake on the part of the actor (e.g., a student who fails an exam due to lack of studying; a civilian who is incarcerated because they broke the law). But it also, more broadly, includes situations in which a desired goal is not achieved, through no fault of the actor’s own. We conducted this pilot study to confirm that people agreed with this broad definition of failure.

**Method**

***Participants***

Following the procedures described in Study 1, we opened a HIT on MTurk for 120 participants. MTurk returned 122 participants (48.8% female, 51.2% male; *M*age = 42.15, *SD* = 12.53).

***Procedure***

First, all participants were introduced to our definition of failure: “This is our definition of failure: failure means falling short of a desired state or goal.” After elaborating on this definition, participants had to summarize the definition of failure in an open-text box to demonstrate comprehension.

Next, we randomly assigned participants to view one of four different failures. Participants reported whether they believed their assigned failure (professional failure, criminal failure, substance abuse failure, other health failure) constituted an actual failure. For example, in the professional domain, participants answered Y/N to the following question: “Would you consider a professional (e.g., a nurse, lawyer, teacher) who does not pass their licensing exam to have experienced a failure?”

**Results**

In the professional domain, 100% of participants agreed that the situation they read about constituted a failure. In the criminal domain, 86.7% of participants agreed that the failure situation they read about constituted a failure. In the substance abuse domain, 87.1% of participants agreed that the failure situation they read about constituted a failure. In the health domain, 93.5% of participants agreed that the failure situation they read about constituted a failure.

In sum, participants agreed that the failure situations presented to participants in Study 1 did indeed constitute failures between 87% and 100% of the time.

**Supplemental Studies S1-S2**

**Method**

***Participants***

We recruited two samples. In each one, we targeted 150 participants with the goal of recruiting approximately 50 participants per failure [Supplemental Study S1: 152 respondents (41.1% female, 58.9% male; *M*age = 42.00, *SD* = 12.22); Supplemental Study S2: 145 respondents (33.3% female, 66.7% male; *M*age = 38.26, *SD* = 11.51)]. We pre-registered both studies: Study S1 (<https://osf.io/nfx6t/?view_only=795c46bd58d24b4eb255ed15f29a2c2f>) and Study S2 (<https://osf.io/krab4/?view_only=8dac82ab83ad4369a1e47d3a156b591b>).

***Procedure***

Prior to randomization, we asked all participants an open-response question (e.g., “Please tell us who your best friend is and why”). We included this question for all online studies because online participants who are not willing to invest effort tend to drop out when they see an open response question. Participants who answered this question were then randomly assigned to a condition and completed the main task.

In Supplemental Study S1, participants predicted success following failure for addicts in recovery, for one of three types of addiction (smoking, alcohol, opioids; randomly assigned). First, participants considered addicts in recovery (i.e., for opioids: “Consider people in the US who have suffered from opioid addiction and are now in recovery”). Next, participants estimated the likelihood of relapse (i.e., for opioids: “What percent of these people will stay in recovery long-term and not relapse?”). Participants indicated a percentage ranging from 0% to 100%.

In Supplemental Study S2, participants predicted success following failure for ex-convicts who committed one of three types of crimes (violent offenses, property offenses, drug offenses; randomly assigned). In each case, participants considered an inmate convicted of one of these crimes, who is subsequently released from prison. Participants then estimated the likelihood of this ex-convict successfully re-integrating into society—meaning that the convict does not re-offend (i.e., for violent offenses: “Consider a prisoner who was convicted of a violent offense. Upon release from prison, what do you think are the chances that this individual successfully re-integrates into society [does not get re-arrested in the next 9 years]?”). Participants indicated a percentage ranging from 0% to 100%.

**Results**

***Supplemental Study S1***

In truth,only 10% of recovered smokers (Sailor, 2011), 20% of recovered alcoholics (Hazelden Betty Ford Foundation, 2021), and 18% of recovered opioid addicts (Chalana et al., 2016) do not relapse after achieving recovery. Participants overestimated the actual likelihood of sustained recovery in each of these domains: smokers (*M* = 43.64% vs. 10%, *SD* = 22.17%), *t*(49) = 10.73, *p* < .001, *d* = 1.52, 95% CI = [0.81, 2.21]; alcoholics (*M* = 43.22% vs. 20%, *SD* = 20.76%), *t*(50) = 7.98, *p* < .001, *d* = 1.12, 95% CI = [0.48, 1.74]; opioid addicts (*M* = 38.41% vs. 18%, *SD* = 23.68%), *t*(50) = 6.16, *p* < .001, *d* = 0.86, 95% CI = [0.26, 1.45].

***Supplemental Study S2***

In truth, only 21% of ex-convicts who commit violent offenses, 13% of ex-convicts who commit property offenses, and 17% of ex-convicts who commit drug offenses successfully re-integrate into society—meaning, they do not re-offend—in the nine years following their release (Alper et al., 2018). Participants overestimated the true likelihood of subsequent success in each of these domains: violent offenses (*M* = 49.46% vs. 21%, *SD* = 23.66%), *t*(49) = 8.51, *p* < .001, *d* = 1.20, 95% CI = [0.54, 1.84]; property offenses (*M* = 44.30% vs. 13%, *SD* = 22.82%), *t*(46) = 9.40, *p* < .001, *d* = 1.37, 95% CI = [0.66, 2.06]; and drug offenses (*M* = 43.63% vs. 17%, *SD* = 21.39%), *t*(47) = 8.63, *p* < .001, *d* = 1.24, 95% CI = [0.57, 1.90].

**Supplementary Analyses for Study 2b**

As noted in the main text of the paper, participants in Control condition—the condition in which an acute failure, a drug overdose, was mentioned—overestimated the true statistical likelihood (16.6%; Kilaru et al., 2020) that the target would enter treatment, *t*(49) = 9.79, *p* < .001, *d* = 1.38, 95% CI [0.70, 2.05].

Participants in the Failure Erased condition also overestimated this true value (16.6%; Kilaru et al., 2020), *t*(49) = 5.13, *p* < .001, *d* = 0.72, 95% CI [0.13, 1.31].

We next conducted a more conservative follow-up analysis. Given that participants in the Failure Erased condition had imagined a target for whom there was no mention of a recent overdose, we instead tested whether the likelihood of entering drug treatment in this condition was greater than the higher lifetime incidence of entering drug treatment for someone with a drug abuse problem, which is 25% (Grant et al., 2016). They did, *t*(49) = 2.42, *p* = .019, *d* = 0.34, 95% CI [0.06, 0.63].

Finally, we tested whether participants in the Control condition overestimated success (compared to the 16.6% likelihood of entering treatment after an overdose) more than participants in the Failure Erased condition overestimated success when we benchmarked estimates in the Failure Erased condition against the actual lifetime likelihood of entering treatment for people with drug addiction—which is 25% (Grant et al., 2016). They did. Using this new benchmark of success in the Failure Erased condition, participants in the Control condition believed the target was more likely to enter drug treatment (*M*Control = 34.20%, *SD* = 24.70%) than participants in the Failure Erased condition (*M*Failure Erased = 7.52%, *SD* = 21.96%), *t*(98) = 5.71, *p* < .001, *d* = 1.14, 95% CI [0.72, 1.56], suggesting that the mention of an acute failure leads to a greater overestimate of success than not mentioning failure—including when estimates in the Failure Erased condition are benchmarked against the actual *higher* lifetime likelihood of “success”—i.e., people with drug addiction entering treatment—which is 25%.

**Pilot for Study 6**

The purpose of Study 6 was to illuminate the percent of patients who have experienced heart failure who actively think about it. In order to determine the percent of people who actively think about heart failure, we conducted a pilot study of participants who had a heart failure.

**Method**

***Participants***

We opened a study on Prolific to 50 participants. Using the panel feature on Prolific, participants were only eligible to participate if they had experienced heart failure. Prolific returned 51 participants. In an initial open-text question that asked participants to describe health problems they had experienced, only 37 participants wrote in that they had experienced heart failure (e.g., “Left-sided heart failure.”). Participants whose open-text responses did not confirm that they had indeed experienced a heart problem as their panel responses suggested were excluded from the study, leaving a final sample of *N* = 37 (28.6% female, 71.4% male; *M*age = 29.20, *SD* = 12.46).

***Procedure***

We asked participants an open-text question: “What topic (or topics) would you say you actively think about?” Participants then wrote in the topics they actively think about.

**Results**

Of the 37 participants who qualified, only one participant wrote that they actively think about “health.” Other participants wrote in a variety of other topics (e.g., “war in Ukraine,” “travel, family, work”) but did not mention anything health related. Thus, in this sample of 37 heart patients, 2.7% (that is, one participant) listed health as a topic they actively think about.

**Additional Studies in the General Discussion**

**General Discussion Study on Managerial Implications**

In the General Discussion of our paper, we mention that overestimating how much is learned from failure may also lead to another implication: over-challenging others. Indeed, we find initial support for this implication in a field test we ran with 251 managers, described below. In this study, managers who overestimated the learning benefits of failure over-challenged their direct reports, leading them to (unintentionally) undermine the very learning they aimed to promote.

***Participants***

First, we recruited managers who were attending a virtual conference to serve as Predictors. We recruited conference attendees to participate in a Qualtrics survey in which they would answer a short series of questions to learn more about themselves and their workplace attitudes. The first part of the survey was the materials for this study; afterwards, participants answered other questions that were not relevant to the current investigation. A total of 251 managers signed into the link and participated (68.5% female, 31.5% male; *M*age = 46.70, *SD* = 14.10).

Second, we recruited a separate set of participants on MTurk to be Experiencers. We opened the survey to 200 participants. MTurk returned 200 respondents (41.5% female, 58.5% male; *M*age = 43.25, *SD* = 12.92).

***Procedure***

Managers who served as Predictors were asked to consider working at a company that was trying to accelerate learning of a new, technical language among a group of online panel participants. Managers were told: “Your goal is to figure out the optimal level of challenge to accelerate learning.” Managers then saw a question from the script task (from Study 4) that online participants would answer as part of the learning process. Managers learned that the task involved guessing on the part of the participant, but that guesses were followed up with correct or incorrect feedback, depending on whether the participant answered the question right or wrong. As the dependent variable for Predictors, we asked managers: “Now imagine you are designing a quiz with 3 multiple choice questions like the one you just saw. Your goal is to maximize how much of the new technical language your group of online participants learns. Tell us: in which scenario do you think the participant would learn more from their guesses and the subsequent feedback?” Managers chose whether they thought online panel participants would learn more from a quiz in which they guessed two questions right and one question wrong, or a quiz on which they guessed all three questions right. These options allowed us to assess whether managers believe that a partial failure (getting 2 questions right and 1 wrong) leads to more learning than no failure (i.e., getting all 3 questions right).

Experiencers were randomized to a condition a two-condition Partial Failure vs. No Failure) between-subjects design. Experiencers took one of the two versions of the script task that the Managers endorsed. For a full description of how the script task works, see Study 4. In the current study, each Experiencer answered three script questions in Round 1 (learning phase). After answering each question, participants received feedback on their response. Afterwards, everyone completed Round 2 (test phase), which consisted of three similar iterations of the initial questions from Round 1.

What differed between the two conditions was whether the feedback indicated that the participant had answered mostly correct (Partial Failure condition: 2 questions correct, 1 incorrect) all three initial questions correct (No Failure condition).

***Results***

In the Predictor condition, managers were more likely to predict that Experiencers would learn more from the Partial Failure version of the quiz (96.41%; 242 of 251 participants) than the No Failure version (4.59%; 9 of 251 participants), χ2(1) = 216.29, *p* < .001.

However, the Experiencers demonstrated managers wrong. Experiencers, who answered three questions, had learning scores that ranged from 0 to 3, depending on their performance. Experiencers in the Partial Failure condition (*M* = 2.40, *SD* = 0.94) learned significantly less than Experiencers in the No Failure condition (*M* = 2.73, *SD* = 0.75), *t*(198) = 2.74, *p* = .007, *d* = 0.39, 95% CI = [0.10, 0.67]. Whereas Managers expected people to learn *more* from some failure than pure success, encountering just some failure alongside success led to a *decrease* in learning.

**General Discussion Study on Specificity**

In the General Discussion of our paper, we mention a study testing specificity of the effect. That is, we examined whether predictors overestimate all successes upon learning about failure, or if they more specifically overestimate the likelihood of the success that is tied to the failure. Specifically, in this study, we tested whether predictors who overestimate the rate at which convicts re-integrate into society are also overly optimistic about the educational achievements of convicts, expecting that they are not.

***Method***

**Participants**. We opened the survey to 100 MTurk participants. MTurk returned 100 respondents (54.0% female, 46% male; *M*age = 41.71, *SD* = 13.27).

**Procedure.** Participants estimated the percent of ex-convicts, in general, who successfully re-integrate into society following their release (true value: 17%; Alper et al., 2018). In addition to answering this question, participants estimated the percent of ex-convicts who have completed college or another post-secondary degree (true value: 15%; Ositelu, 2014). This second value (15%) was intentionally similar to the first (17%), so we could test whether overestimates occur for all values in this range, or if participants specifically overestimate success following failure.

***Results***

Participants overestimated the actual percent of ex-convicts who ‘succeed’ following failure—that is, the percent that do not re-offend upon release from prison (*M* = 42.46%, *SD* = 18.83%; 95% CI = [38.72%, 46.20%]; true value: 17%), *t*(99) = 13.52, *p* < .001, *d* = 1.35, 95% CI = [1.08, 1.62]. In contrast, participants did *not* overestimate the percent of ex-convicts with college or post-secondary degrees (*M* = 16.14%, *SD* = 12.38%; 95% CI = [13.51%, 18.77%]; true value: 15%), *t*(99) = 0.86, *p* = .393, *d* = 0.09, 95% CI = [-0.11, 0.28].

**Figures with Full Data Points**

Here we re-present the five figures from the manuscript, including all data points. Points cluster at specific values in Figure 3 because there were only several discrete values for the dependent variable.

**Figure 1**

A graph of a number of people who succeed

Description automatically generated

**Figure 2**

A graph of a study and a few failure erasing

Description automatically generated with medium confidence

**Figure 3**

A graph of a test

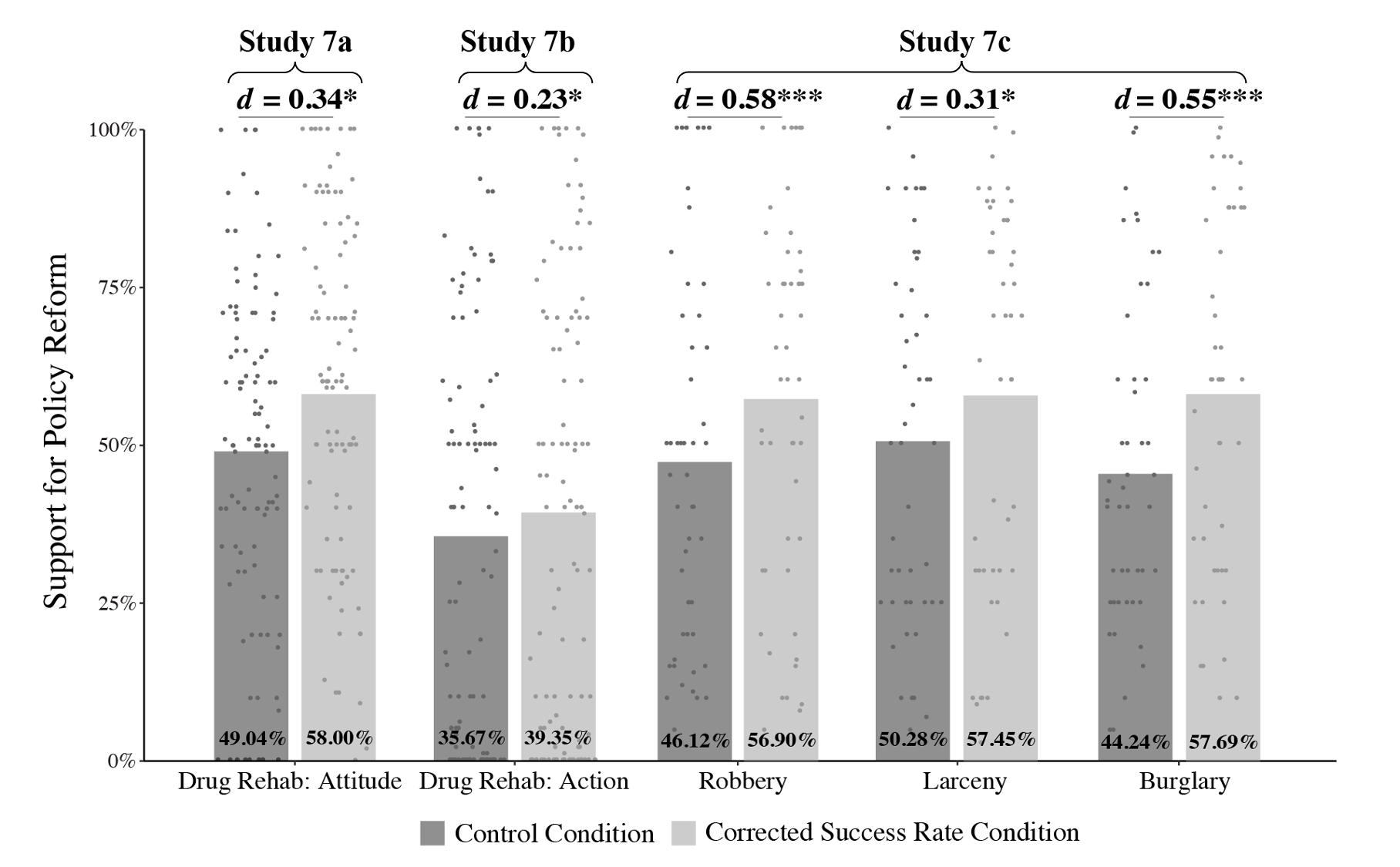
Description automatically generated

**Figure 4**

A graph of value of true value

Description automatically generated

**Figure 5**



**References**

Alper, M., Durose, M. R., & Markman, J. (2018). *2018 Update on prisoner recidivism: a 9-year follow-up period (2005-2014).* U.S. Department of Justice Office of Justice Programs Bureau of Justice Statistics.

Chalana, H., Kundal, T., Gupta, V., & Malhari, A. S. (2016). Predictors of relapse after inpatient opioid detoxification during 1-year follow-up. *Journal of addiction*, 2016.

Grant, B. F., Saha, T. D., Ruan, W. J., Goldstein, R. B., Chou, S. P., Jung, J., ... & Hasin, D. S. (2016). Epidemiology of DSM-5 drug use disorder: Results from the National Epidemiologic Survey on Alcohol and Related Conditions–III. *JAMA psychiatry, 73*(1), 39-47.

Hazelden Betty Ford Foundation. (2021, April 21). *Understanding relapse and the risks.* Retrieved from https://www.hazeldenbettyford.org/articles/relapse-risks-stats-and-warning-signs

Kilaru, A. S., Xiong, A., Lowenstein, M., Meisel, Z. F., Perrone, J., Khatri, U., ... & Delgado, M. K. (2020). Incidence of treatment for opioid use disorder following nonfatal overdose in commercially insured patients. *JAMA network open, 3*(5), e205852-e205852.

Ositelu, M. (2014). *Equipping individuals for Life Beyond Bars.* New America.

https://www.newamerica.org/education-policy/reports/equipping-individuals-life-beyond-bars/results/#:~:text=Only%2015%20percent%20of%20incarcerated,See%20figure%206.

Sailor, M. (2011, January 10). *How often do smokers relapse when they quit?* HowStuffWorks. Retrieved from https://health.howstuffworks.com/wellness/smoking-cessation/how-often-do-smokers-relapse.htm