# *Should* employees report when things go wrong?

Elaine Yao and Brandon Freiberg
SFI GWCSS 2022

June 28, 2022

## 1 Introduction

Things going wrong can sometimes be good. For example, "accidental" scientific discoveries such as the invention of penicillin originated from unintended mistakes. The ability to turn these mistakes into successes often relies on the expertise of specialists who recognize that accidental deviations from the plan can actually be beneficial.

In contrast, classical principal-agent problems are typically concerned with how uninformed decision-makers (principals) can extract truthful information from informed agents who have incentives to lie based on private information. We consider a inversion of this problem: can it be beneficial to decision-makers to allow informed agents to lie about things going wrong? We develop a model and simulation which provide a mechanistic explanation for why principals may not prefer to force agents to conform to their preferences.

We frame our problem as one of organizational experimentation. A company initially pursues a goal determined by the CEO. However, its progress towards this goal is the private information of the expert employee who can choose whether or not to report when things go wrong. If the employee reports the error to the CEO, the company reverts to the original goal. However, if he doesn't report, then he can maneuver the company towards another, possibly better goal. We show that in a baseline scenario where lying is not punished, employees frequently have the ability to direct organizations towards a better local maximum, if one exists.

Our work is partly motivated by work on experimentation in organizations using multi-armed bandit models which have received some attention in management science, economics, and political science. These models typically involve a decision to either play a "safe" alternative which offers a known and predictable payoff, or to play a "risky" alternative which offers an unknown but possibly superior payoff. What we do is, in a sense, an attempt to examine this "exploitation-exploration" trade-off in a context with an experimentation-adverse principal and an experimentation-loving agent.

# 2 Model overview

The company consists of a CEO and an employee. At $t = 0$, the company is instantiated at a random point on a continuous three-dimensional surface, which encodes both the position of the company along the $x - y$ plane and the *fitness* at different points in the plane (the height of the surface $z$). We consider different possible landscapes with various levels of "ruggedness" (the frequency and extremeness of peaks and troughs). The state consists then simply of $(x, y, z, t)$.

At the beginning of time, the company is initialized at a random point in the landscape. The CEO cannot see the entire landscape, but observes the fitness level in every point in a rectangular neighborhood with some radius. He picks the direction corresponding to the highest point in the neighborhood (the goal). In the absence of any deviations from the path, the company simply progresses towards goal one integer step at a time. After $t = 0$, the CEO does not directly observe the position of the company, but instead relies on reports from the employee.

At every time $t$, the company deviates from its initial path with probability $p < 1$ in a random direction. The employee has the choice to report the deviation to the CEO or not. If he reports it, then the company reverts to progressing towards the original goal. However, the employee observes the current position as well as some neighborhood around this position (for simplicity we assume this is the same as the radius of the CEO's original observation). If the employee detects a peak in the current neighborhood that exceeds the original goal, he will choose to not report the deviation, and allows the company to progress towards the new peak. The employee cannot consciously "commit" a mistake in order to experiment, he can only conceal and leverage a mistake that happens randomly.

# 3 Simulation

## 3.1 Parameter values and functional forms

To simulate our model, we selected parameter values for model primitives and added functional form additional assumptions. To simulate a "smooth" landscape, we simply use $f(x,y) = -(x^2 + y^2)$. To simulate "highly rugged" landscape, we use the "eggholder function", which has a high level of local maxima and minima, and is defined as

$$f(x,y) = -(y+47)\sin\sqrt{\left|\frac{x}{2} + (y+47)\right|} - x\sin\sqrt{|x-(y+47)|}$$

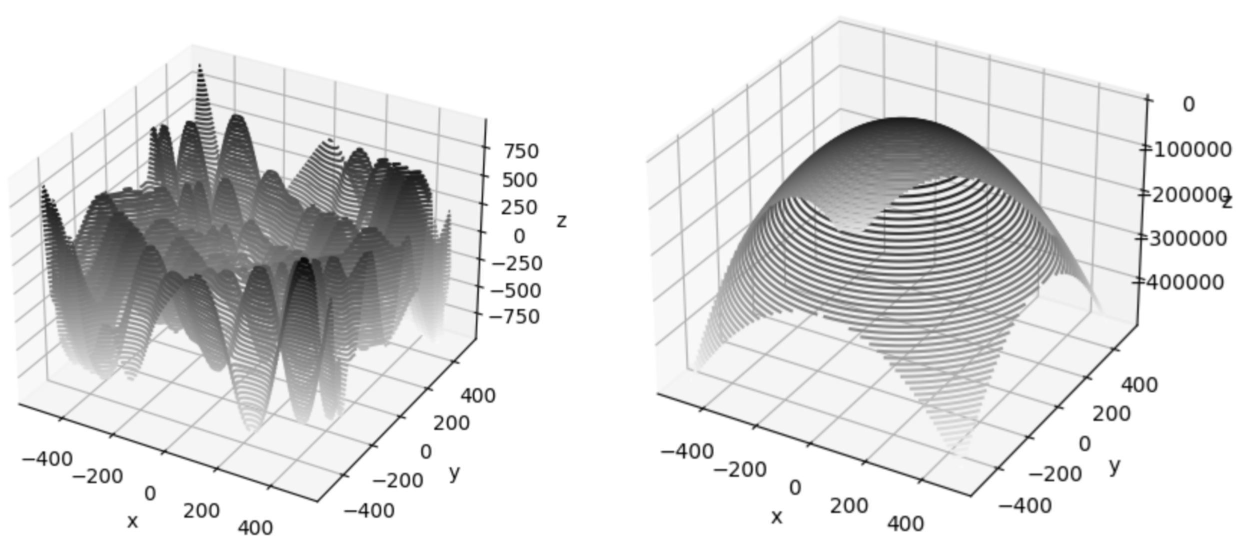


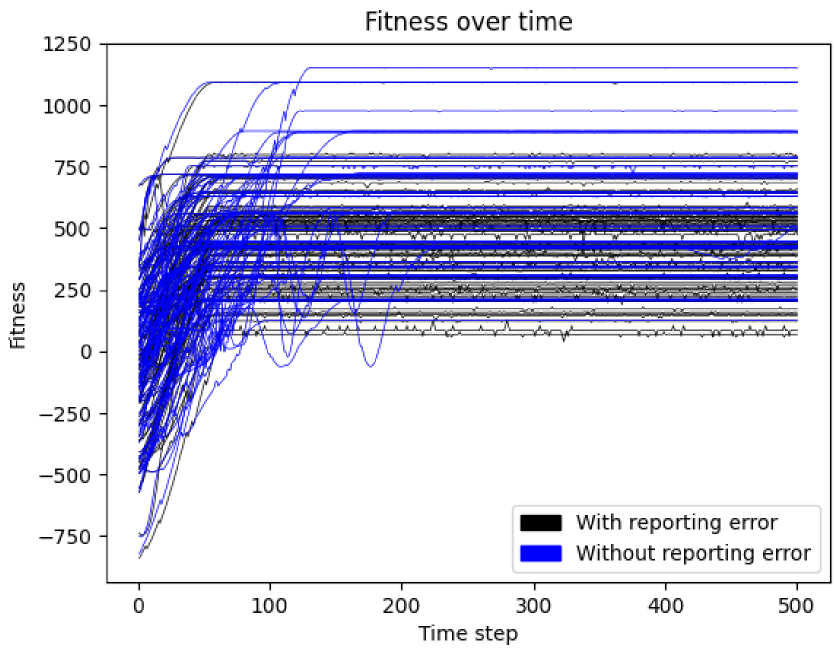

Figure 1: "Eggholder function" v. Cone function

The radius of the CEO's observable neighborhood at $t = 0$ and of the employee's observable neighborhood at the time of any possible deviation are both 50 in the "eggholder" scenario and

The probability of deviation at any given time step is 10%. We impose an exogenous time limit on organization experimentation of 500 time steps and allow the organization to move one integer step in the $x$ and/or $y$ directions.

### 3.2 Results

The primary aim of simulation was to compare what happens when employees were encouraged to report deviations. We ran 1000 simulations of this model. We plotted fitness over time in simulations where the employee did report errors (in black) and did not report errors (in blue)

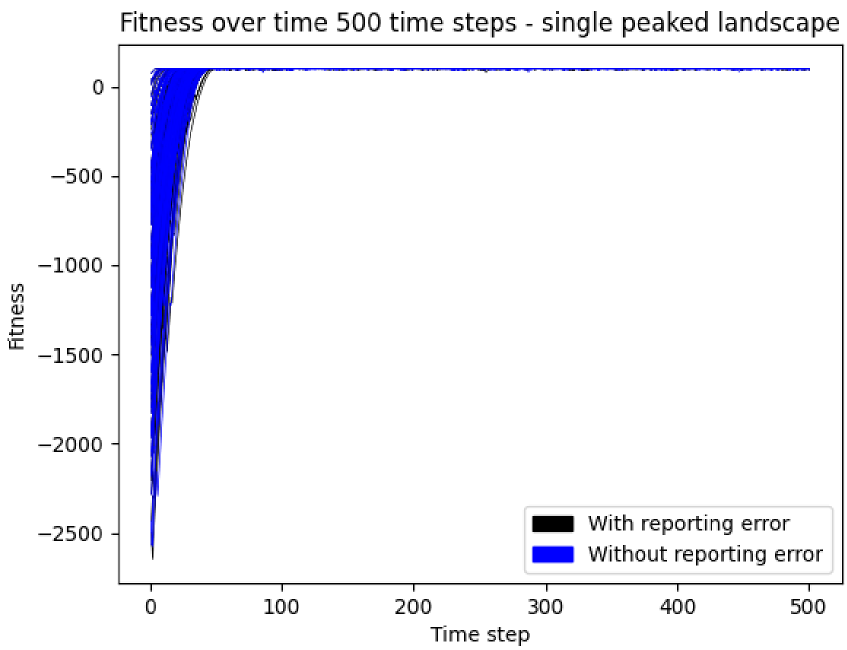



Note that because of the way we have hard-coded time steps, it is possible for small deviations to occur even once the company has converged. Note also that we can see that sometimes the blue line (where things going wrong are not reported) that sometimes performance dips before converging to a new peak. While it's hard to read a single line clearly form this graph, we can see the possible effect of "cumulative errors".

Some summary statistics (standard deviation in parentheses)

| | Employee reports errors | Employee conceals errors |
|---|---|---|
| Probability of attaining CEO Peak | 0.91 | 0.25 |
| Fitness ("aka profit") | 450.01 (18.24) | 559.28 (205.58) |
| Timesteps until reaching peak | 50.67 (8.83) | 75.88 (34.31) |

When might there be no benefit to exploration? Consider, for example, a landscape with a single global maximum, such as $z = -(x^2 + y^2)$. In this case, even if the employee doesn't report errors, the new path he selects will lead towards the same global maximum. To trivially show this, we make the same fitness plot with the new landscape. In this case, there is clearly not any benefit to allowing employees to conceal errors, which simply results in a slightly less efficient path towards the goal.




# 4 Extensions and improvements

- The current model does not account for strategic concerns on part of either actor – all the dynamics are completely mechanistic. In previous iterations of the model, we thought about ways of fixing this, such as the following: If the employee reports the problem right away, they get fired with some small probability. If they do not report the problem and the height of the surface decreases for "long enough", then they get fired for sure. However, if they do not report the problem and the height of the surface reaches or surpasses the initial goal within some time limit, then they don't get punished.

  We ended up not having enough time to both solve for the expected utility of reporting/not reporting, in part because it's not clear from our setup what the probabilities associated with firing, and because this adds a bunch of model primitives (e.g. probabilities of firing, "time limit") that don't seem essential to our model. Ideally you would like to tie beliefs of about the riskiness of the path to what you can see in the neighborhood, which would require better-specified beliefs over the distribution of fitness in the landscape.

- A related concern is that players don't really have payoffs in our model. Ideally you would tie payoffs to the performance of the company in some way, which might provide better incentive for the principal to allow the agent more freedom to allow "things going wrong" .

- Originally we wanted to use an NK landscape to encode the "landscape" of the setting. This would be preferable because it allows for tuning the "ruggedness" of the landscape, which influences how often deviations lead to new successes.

- Does anything change with multiple employees and information? One possibility we considered is that there are several employees, each of which who witness a noisy correlated signal of fitness (e.g. each employee $i$ sees $s_i(z) \sim N(z, \epsilon)$) and chooses whether or not to report an error to the CEO. This would also add an interesting strategic layer, because the CEO would have to ascertain the probability that an employee who reported had simply gotten a "bad signal" of the true location.