

One on One SRE

Amy Tobey

<https://github.com/tobert>

@MissAmyTobey



@MissAmyTobey

Good morning! I'm Amy Tobey and I'm an SRE at GitHub.

Agenda

- a confounding question
- story: human error
- trauma & polyvagal theory
- safety II
- story: insufficient guard rails
- human factors & ergonomics
- story: inconceivable confluence
- the 1:1 incident debrief
- divergent behavior
- spanning tree
- recapitulation
- we're hiring slide



@MissAmyTobey

We're going to start with a question, I'll tell some stories and give you some background material. Then we'll step through the 1:1 debrief, talk about some variations on it, then we'll wrap up and I'll show you my we're hiring slide.

This is the only bullet list from me today :)

How can an individual contributor influence availability at company-wide scale?



@MissAmyTobey

(breathe!)

How can an individual contributor influence availability at company-wide scale?

I've been asking myself this question for almost 20 years. I think I have some answers.

Before I dive into theory, please indulge me in a **formative** story:

human error



@MissAmyTobey

(breathe!)

Some of you have probably heard me tell this one before, but this time I have another angle.

On the second day of my first job after dropping out of college, the senior sysadmin dropped by the datacenter and casually said something like, "hey Tobey, go log into that SPARC 10 in the rack right in front of you and kill the foobar process."

Side note: it's 9:30 in the morning and I was working at a trading company.

So I log into that machine and type "killall foobar". For those of you who have yet to enjoy the blessings of Solaris, that command is literal on Solaris 2.5. Every process died at once with nary a quiver in the sound of the datacenter. 2 seconds later I'm already freaking out and the DC phone rings. I pick it up and before I can utter "datacenter," the other end screams, "WHAT THE --BEEP-- IS GOING ON DOWN THERE!?"

background material



@MissAmyTobey

(breathe!)

I want to mention some ideas I've been learning recently so I checked in on Twitter to see if I should explain. Because I only have a little of your time, I'm going to briefly introduce these ideas and I might not get them quite right. I think they are important to the future of SRE and I hope that you'll all dig into them on your own later.

trauma: extreme stress that
overwhelms a person's ability to cope



@MissAmyTobey

(breathe!)

I didn't think much of that story at the time, but looking back, I've thought about it frequently for going on 20 years now. Sometimes unbidden. It's trauma.

I think a lot of us have experienced trauma like this in incidents and that it is unavoidable. The only solutions to be had are in how we respond when it inevitably occurs.

[The Body Keeps the Score](#)

polyvagal theory
proposes a third
level of evolved
response to stress:
social engagement

social engagement



mobilization



shutdown



@MissAmyTobey

(breathe!)

social engagement: most recently evolved part of the vagus nerve that enables social response to stress by looking to those around us

mobilization: the social engagement branch of the vagus nerve is turned off, the so-called "lizard brain" - fight, flight, and active freezing

shutdown: the lowest & oldest branch, where we disassociate and become immobile

[Polyvagal Theory](#)

quote: "Porges's [Polyvagal] theory made us look beyond the effects of fight or flight and put social relationships front and center in our understanding of trauma."

"It also suggested new approaches to healing that focus on strengthening the body's system for regulating arousal."

[-Bessel van der Kolk](#)

Quote source: https://en.wikipedia.org/wiki/Polyvagal_theory

Video of Dr Porges describing polyvagal theory:

<https://www.youtube.com/watch?v=ec3AUMDjtKQ>

insufficient guard rails



@MissAmyTobey

(breathe!)

A couple years ago, I was an incident commander at a large company and participated in a lot of incidents and subsequent reviews. One review in particular stands out in my memory.

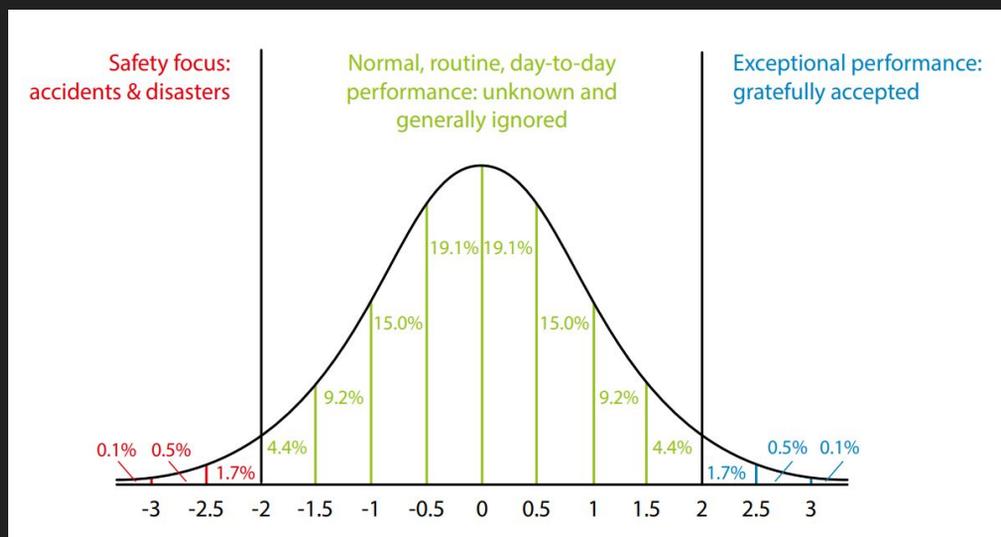
The technical details aren't that important but it goes something like this: The system that loads thousands of scripts had a glitch where it took down the main API service. It was a complex system that injected code into the hottest of hot paths while thousands of RPS of traffic was flowing. It usually worked fine, but it had recently caused a couple outages in quick succession. Someone on my team did the investigation and presentation brilliantly and that's not where I feel like things went wrong.

After the presentation, we had Q&A and much like some conferences, the Q&A turned into an airing of grievances and "this is more of a statement than a question...."

I can't say what that service team was feeling, but I remember sitting there and feeling awful. I have been in their shoes. I've been shouted at. I've seen friends fired. While they were in no real danger, they were being cross-examined by their peers, in front of their peer.

That was when I realized I that group retrospectives were too easily toxic.

Safety-II



@MissAmyTobey

(breathe!)

I'm not going to spend a lot of time on this because it's a whole series of talks on its own. There's a lot more to Safety-II and what I want to point out is the idea of branching out from our focus on what went wrong to spend more time on what went right.

The main thing to take away is that focusing on failures isn't enough. The next few epochs of systems complexity require a more thorough approach to safety and that means looking at all the work that we do to keep our systems going. It means looking at that 99% of work that goes right and figuring out why.

"Safety-II is the system's ability to succeed under varying conditions, so that the number of intended and acceptable outcomes (in other words, everyday activities) is as high as possible." -- EUROCONTROL white paper, "From Safety-I to Safety-II: A White Paper".

image & quote source: EUROCONTROL

<https://www.skybrary.aero/bookshelf/books/2437.pdf>

human factors & ergonomics

the application of psychological and physiological principles to the engineering and design of products, processes, and systems
(Wikipedia)



@MissAmyTobey

(breathe!)

I'm not going to dig into human factors, but it gets lumped with Safety-II a lot because they're (in my mind) closely related. In short, human factors & ergonomics focus on the human & system interfaces and how they impact safety, efficiency, and accuracy.

I want you to learn about these ideas because I believe they are a big part of the future of SRE for all of us.

unknown unknowns



@MissAmyTobey

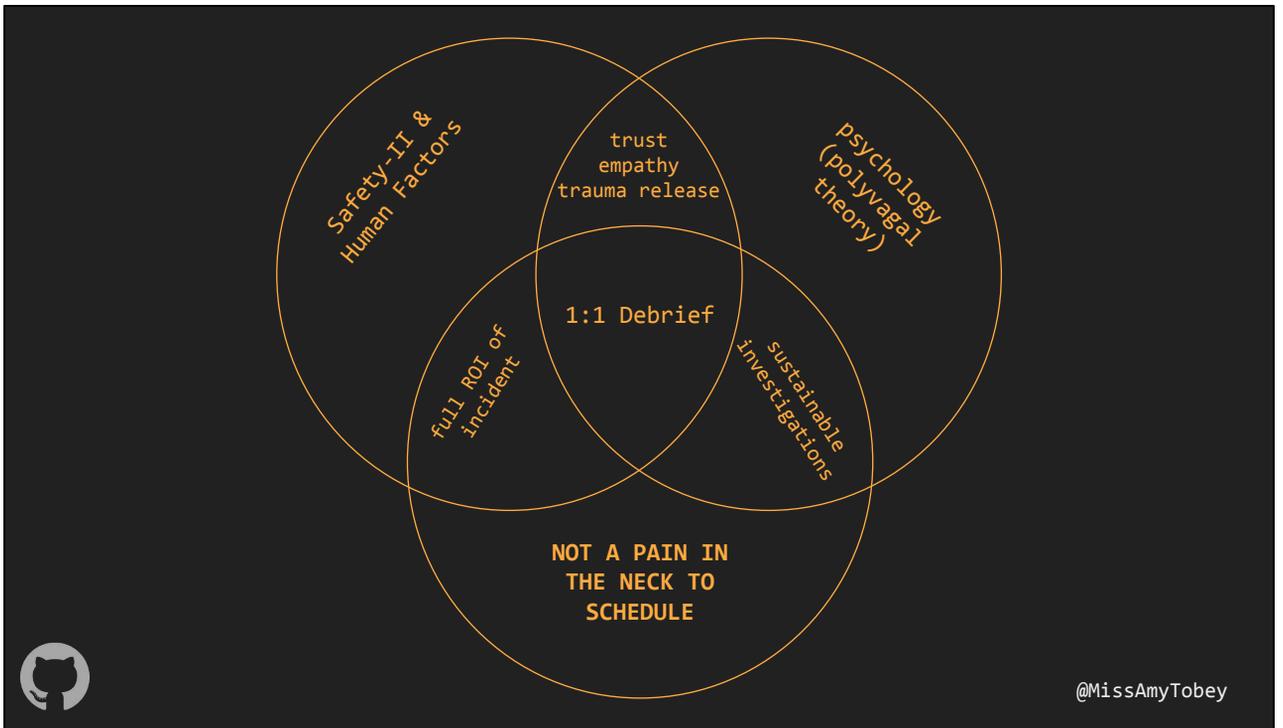
(breathe!)

last story: I joined GitHub in June of last year. In late October, GitHub experienced its longest outage in its history. I wasn't directly involved in the incident itself, but I followed along, as one does. TL;DR:

- 1.) an optic started going bad, so our people started work to repair it
- 2.) a mistake was made when replacing the optic, resulting in 43 seconds of downtime
- 3.) our database failover mechanism elected one leader in each datacenter, resulting in a split-brain
- 4.) we chose to protect our customers' data integrity and restored from backups

A few things stood out to me: everyone was calm. People took care of each other, taking over shifts, encouraging folks to go eat & rest, and a general tone of kindness. I've seen teams handle incidents professionally plenty of times, but this was something new to me. This is a team - and I mean the greater GitHub team, including a big chunk of engineering, corporate communications, legal, and even finance - I saw them all work together to bring GitHub online while caring for each other in the process. I was blown away and even though we'd just been through a major incident, I was more excited than ever to be part of it.

[Post Incident Analysis](#)



(breathe!) (enjoy the giggles)

I am not being totally serious here. I just wanted to call out the most important part of all of this: NOT A PAIN IN THE NECK TO SCHEDULE.

We were able to split the interviews across my team, so we got most of them done within a week and finished up the stragglers in the following week.

OK now let's look at the agenda, item by item:

[optional]

To explain my (tongue-in-cheek) reasoning for this travesty of a Venn diagram:

"sustainable investigations" is the result of an investigation process that preserves psychological safety and builds non-adversarial relationships.

"full ROI on incident" is realized by reducing the waste inherent in exploratory meetings and getting everything of value you can from incidents (that are de-facto pre-paid investments) (TODO: link the @allspaw tweets?)

"trust & empathy & trauma release" happen in close settings where individuals see and hear each other. Things are plainly less stressful when scheduling isn't getting you down.

The 1:1 Incident Debrief



@MissAmyTobey

(breathe!)

We usually schedule 30-60 minutes for a debrief. For the MySQL incident, most of them were an hour long and we didn't use all the time. A few were a half hour and that felt crunched but was still effective.

- introductions and agenda review



@MissAmyTobey

This is pretty straightforward meeting etiquette and agendas are magic. We've all had that meeting where someone starts talking and forgets to introduce themselves.

● informed consent



@MissAmyTobey

This is the part where we set up the context for vulnerability: in the debrief instructions it says to tell the interviewee a few things:

"I will be taking notes while we talk.

These notes will be available to myself and the incident investigators only.

After this meeting, I will clean up the notes and share them with you. Please review them as soon as possible and make sure I didn't misrepresent you."

Sometimes I will briefly talk about how I take the notes. I type continuously and transcribe nearly word-for-word.

I recommend strict share of the document by handle/email, not using groups. In my earliest experiment I tried using a git repo, because GitHub. It turned out to be the wrong choice: the interviewee wanted to rephrase something because what I wrote sounded harsh. When I re-read it, I agreed, so we pushed a commit. But there's the history sitting there in GitHub systems. We didn't want to take a chance, so we wiped that repo and moved to a Google Doc. YMMV.

- what was your role in the incident?



@MissAmyTobey

Sometimes you'll get the whole story right here & now. It really depends on the person. Sometimes I'll let them go and just write it all down. Sometimes I'll interrupt, though I prefer not to.

- what surprised you?



@MissAmyTobey

This usually works by itself. Sometimes it helps to talk about things that surprised you to get the ideas going.

For example, in the mysql incident I told people that even though it's common sense, I was surprised to learn that we tested restores every day.

We talked to a couple people who were surprised how much stuff broke when the DB went away. That, in turn, surprised me. This was a useful finding!

- how long did you work on the incident?



@MissAmyTobey

I ask people to be coarse and use estimates. The goal isn't to measure. The goal is to keep an eye out for fatigue, burnout, and heroism.

- were you able to get the support you needed?



@MissAmyTobey

These next few often get overlapping answers. That's OK. These questions are meant to get the conversation going and cover similar ground. Let people ramble a bit.

This can also raise flags for burnout and frustration.

- do you feel that the incident was preventable?



@MissAmyTobey

This is when folks get salty.

That's a good thing. Stay calm and alert. Some emotion is going to come at you. Pace your breathing. ← ties back to polyvagal theory

Sometimes you get a legit straight-talking explanation of what they think went wrong.

There is opportunity here to lead through anger & frustration.

Sometimes when emotions come into it and we're all going to say foolish things. Work from there. Ask more questions.

- what actions do you feel good about?



@MissAmyTobey

This one touches on the Safety-II concept of looking into what works and not just what went wrong.

Celebrating good things is a healthy and effective way to reinforce behaviors that you want to sustain.

In my experience, folks sometimes struggle to come up with something at first. Almost everybody comes up with something.

- what do you think could have been better?



@MissAmyTobey

This overlaps a bit with "do you feel that the incident was preventable?" and sometimes it's OK to say "yeah we already covered this earlier" and move on.

- what did you learn from this incident?



@MissAmyTobey

This uncovers all kinds of hidden insecurities and gaps in knowledge.

Some insights are invaluable.

- what do you think we can do to prevent reoccurrence?



@MissAmyTobey

This one is kinda redundant and still manages to uncover new ideas.

- did our tools and documentation serve you well?



@MissAmyTobey

This is pretty self-explanatory. For example, last October we learned about a couple playbooks that need work and found some dependencies in our CI that needed to be cleaned up.

- did you practice self-care during this process?



@MissAmyTobey

I love this question.

Asking reinforces the behavior!

Most of our folks were thrilled to answer this one. We had one person who said they stayed up all night because they wanted to be there. That's OK too. This is an opportunity to make sure they take time off or whatever they need to do to recharge.

- can you think of anyone else we should talk to?



@MissAmyTobey

For the big MySQL incident we started with about 15 people on the interview list. By the time we were done we interviewed 34 people from all over GitHub! I don't always ask this one, but sometimes it can be critical to getting the whole story.

spanning tree



@MissAmyTobey

SRE Outreach is something I was working on when I proposed this talk and have not made the same progress as with the debriefs. That said, it seems to be effective as well.

As an individual contributor, it can be difficult to get a team's time to create relationships and provide guidance. In most orgs, setting up 1:1's flies under the radar and is, in my experience, an effective way to build relationships and get to a place where people will come to you & SRE for advice.

I keep an eye out for interesting people and when I say to myself, "self, you should find out what that interesting person is up to!", I schedule a 1:1. So far, this has been a great way to build up relationships and has born fruit.

deviant behavior



@MissAmyTobey

If you choose to run with these ideas, you *should* adapt to your environment. Your culture will need different things than I did.

A common variation on the debriefs is written forms. I've had a couple people write up their own interview. It worked OK but it was missing the raw takes you get in an interview, as well as the catharsis. You might find folks who are more comfortable writing than speaking and that's OK. It might be worth trying these as a live Slack conversation, but I haven't tried it.

How can I, an individual contributor, impact reliability at organizational scale?



@MissAmyTobey

just read it :)

The secret to leading
organizational change is empathy.

@PattiSan



@MissAmyTobey

I was trying to come up with a TL;DR for this talk and googled "empathy change" for inspiration. The second link was an hbr.org (Harvard Business Review) article by Patty Sanchez with this quote as the title. I love it, so here you are.

<https://hbr.org/2018/12/the-secret-to-leading-organizational-change-is-empathy>

-Amy

<https://github.com/tobert>
[@MissAmyTobey](https://github.com/MissAmyTobey)



<https://github.com/about/careers>
<https://boards.greenhouse.io/github/jobs/1558732>



@MissAmyTobey

THANK YOU

This just in: Everybody is hiring!

We have a bunch of open positions in infrastructure, including my team. We are remote-first. No pressure. Call me :)