



One on One SRE

Amy Tobey
SRE Tech Lead @ Elastic
@MissAmyTobey
SRECon EMEA 2019

(pranayama)

agenda

- a confounding question
- story: human error
- trauma & polyvagal theory
- story: insufficient guard rails
- resilience engineering & human factors & safety II
- story: emergent behavior
- the 1:1 incident debrief
- 1:1 outreach
- recapitulation
- final slide

2

@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 SREs influence availability at company-wide scale?

@MissAmyTobey

(breathe!)

The template for this slide said *“Place a quote from someone really, really important and it will shrink to fit this space...”* (Author name here)

so I did ;)

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!)

I've told this story more than a few times, but I'd like to come at it from a new 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 9:30AM ET 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!?"

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

5

@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, maybe not so bad, but still trauma. I still remember the fear.

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.

Please read: [The Body Keeps the Score](#)

polyvagal theory
proposes a third level
of response to stress:

social engagement

social engagement

mobilization

shutdown

6

@MissAmyTobey

 elastic

(breathe!)

The impact of polyvagal theory on trauma treatment and psychology is still unfolding. In particular, PTSD treatment has benefited greatly from the building on polyvagal theory. I believe the impact on our organizations will be just as big as the research builds up and spreads. I hope you think a little about how it applies in your interactions, and in particular in incidents, because it's gives us a more precise way to consider decisions made under pressure.

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, unfortunate power dynamics, 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 peers and no matter how seasoned we get, this will always suck.

resonance



8

@MissAmyTobey



(breathe!)

Maybe it's starting to be a thing here at SRECon, but I'm still geeking out about resilience engineering. For me, it's a way to study and talk about things many of us in the SRE community have been promoting for years.

(don't linger, move on, explain in next slide)

resilience engineering

The essence of resilience is therefore the intrinsic ability of an organisation (system) to maintain or regain a dynamically stable state, which allows it to continue operations after a major mishap and/or in the presence of a continuous stress.

9

@MissAmyTobey



(breathe!)

(maybe mention the origin of 'resilience' if time looks good, doubly so if the storm comes through)

"The essence of resilience is therefore the intrinsic ability of an organisation (system) to maintain or regain a dynamically stable state, which allows it to continue operations after a major mishap and/or in the presence of a continuous stress."
[Resilience Engineering: Concepts and Precepts, 2006](#), Hollnagel, Woods, Levison

"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".

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

Human Factors, Wikipedia: "the application of psychological and physiological principles to the engineering and design of products, processes, and systems"

https://en.wikipedia.org/wiki/Human_factors_and_ergonomics



*Resilience engineering is
intersectional, like
feminism, which it
intersects with frequently.*

@MissAmyTobey

(breathe!)

A common complaint of fellow devops practitioners is that it all went wrong when things got too wrapped around technology. I agree with this assessment and I hope this framing helps. Resilience engineering is intersectional; most of us shouldn't be resilience engineers, but instead SWEs, SREs, testers, operators, lawyers, accountants, PMs, leaders, and so on, who apply resilience engineering principles to our roles. The same could be said for devops done right!

Endlessly shipping tech doesn't create resilience. Change across the socio-technical system towards a common goal - resilience - does. We can start that change by reaching out and having discussions. Sure, it'd be nice if our companies stated resilience as a go, but we have work to do to get there and I hope this is one way to get started.

emergent behavior

@MissAmyTobey



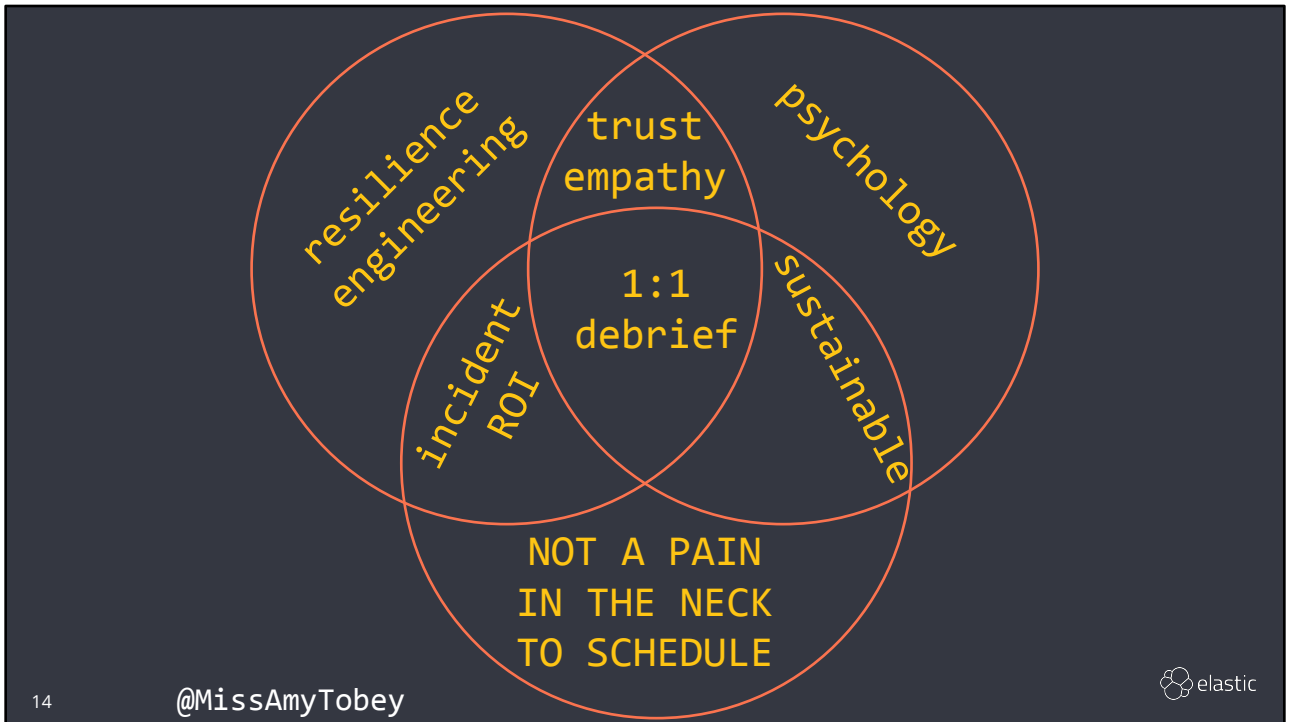
(breathe!)

last story: In late October 2018, GitHub experienced its longest outage in its history. I wasn't directly involved in the incident itself, but I followed along, as one does. This account is drawn from the public analysis. TL;DR:

- 1.) a physical network fault occurred on a serviceable part
- 2.) technician made a simple mistake, resulting in 43 seconds of network downtime
- 3.) the MySQL cluster got into a split-brain situation
- 4.) responders chose to protect 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. Following that, we implemented the 1:1 debrief approach for the first time.

[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.

I'd like to point out again that resilience engineering is intersectional; at the intersection of psychology, resilience engineering, and everyday time constraints, the 1:1 provides an opportunity to learn more about what happened while providing a safe space for folks to talk through their experience, which is usually healthy.

We were able to split the interviews across the 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" 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 scheduled 30-60 minutes for debriefs. 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.

This is what it was in late 2018. It might have changed since my time at GitHub. You'll have to ask them :)

- introductions & agenda review

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

15

@MissAmyTobey



This is the part where you 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. My style is to type as they talk and record almost everything. Others take a pause to type. It's up to you.

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. Some folks weren't comfortable with that as the medium so we switched to gdocs. YMMV.

- what was your role in the incident?

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?

17

@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 GitHub MySQL incident I told people that even though it's common sense, I was surprised to learn that restores are tested 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?

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?

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?

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 folks through their 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?

This one touches on the resilience engineering & safety-ii/"new view" 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?

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?

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

The individual insights here are invaluable; sometimes people will mention stuff in the 1:1 setting that they wouldn't in a group setting where vulnerability is more difficult.

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

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

- did our tools and documentation serve you well?

This is pretty self-explanatory. For example, in October 2018, GitHub learned about a couple playbooks that need work and found some dependencies in CI that needed work.

- did you practice self-care during this process?

I love this question.

Asking reinforces the behavior!

Most of the folks were thrilled to answer this one. There was 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?

(breathe!)

For the October 2018 MySQL incident we started with about 15 people on the interview list based on Slack history and guesses. 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.

vulnerability



(breathe!)

The 1:1 debriefs were from my time at GitHub. When I started working on this talk, I also wanted to talk about how 1:1's can be used to build a network in your organization. Since joining Elastic, I've had to do just that and 1:1's were key in bootstrapping in just a few weeks and learning the state of the system by absorbing many viewpoints.

The key feature of a 1:1 is that vulnerability comes more easily. When people feel safe being vulnerable, we have the opportunity to explore people's expectations and how they're matched or not. Since the increments of time spent on 1:1's is low, it's easy to slowly build relationships and connections across organizational boundaries. In most organizations, this requires no permission and in my experience, reaps rewards for everyone.



redundancy



(breathe!).

To bring things back around to resilience, think about how information flows through your organization. Do messages get lost? Does context spread effectively? 1:1's can fill gaps in standard communications channels, strengthening the social network, which ideally in turn, strengthens the system over time.

flying under the radar



(breathe!)

How many of you have experienced packet loss in inter-team or heirarchy communications?

Another neat attribute of 1:1 networking is you rarely need permission or group consensus to get started. Scheduling 1:1's is cheap - it's a single unit of time for you & your partner. This is especially critical when building networks outside of your own area, and can open the path to inter-disciplinary coordination that would otherwise be difficult and/or expensive to establish.



*The secret to Leading
organizational change is
empathy.*

Patty Sanchez | @PattiSan

@MissAmyTobey

(breathe)

How can SREs influence availability at organizational scale?

I was trying to come up with a summary 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>

@MissAmyTobey

. <https://www.elastic.co/about/careers/>



THANK YOU

This just in: everybody is hiring!

Elastic is a remote-first company and we're hiring SREs as well as other E's.