Microsoft Azure

Learning from incidents

Understanding how things went right

Nick Stenning
@nickstenning

written with Jessica DeVita
@ubergeekgirl

http://www.americanairmuseum.com/aircraft/10376

Agenda

1. Why learn from incidents?

Agenda

1. Why learn from incidents?

2. Four common traps.

Agenda

1. Why learn from incidents?

2. Four common traps.

3. Four helpful practices.

Why do we try to learn from incidents?

How Complex Systems Fail

(Being a Short Treatise on the Nature of Failure; How Failure is Evaluated; How Failure is Attributed to Proximate Cause; and the Resulting New Understanding of Patient Safety)

Richard I. Cook, MD

https://aka.ms/csfail

"Complex systems contain changing mixtures of failures latent within them."

https://aka.ms/csfail

"Complex systems contain changing mixtures of failures latent within them."

"Complex systems run in degraded mode."

https://aka.ms/csfail

“Complex systems contain changing mixtures of failures latent within them.”

“Complex systems run in degraded mode.”

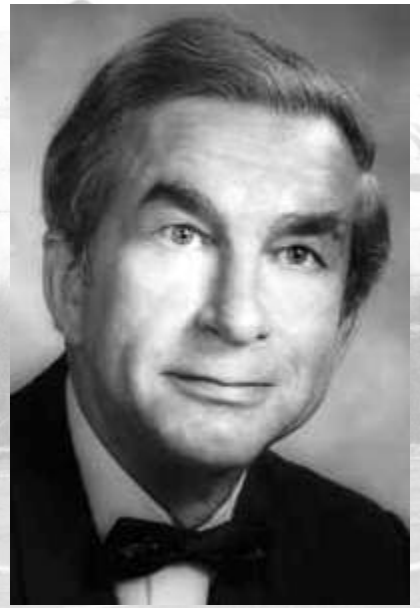
“Catastrophe is always just around the corner.”

https://aka.ms/csfail

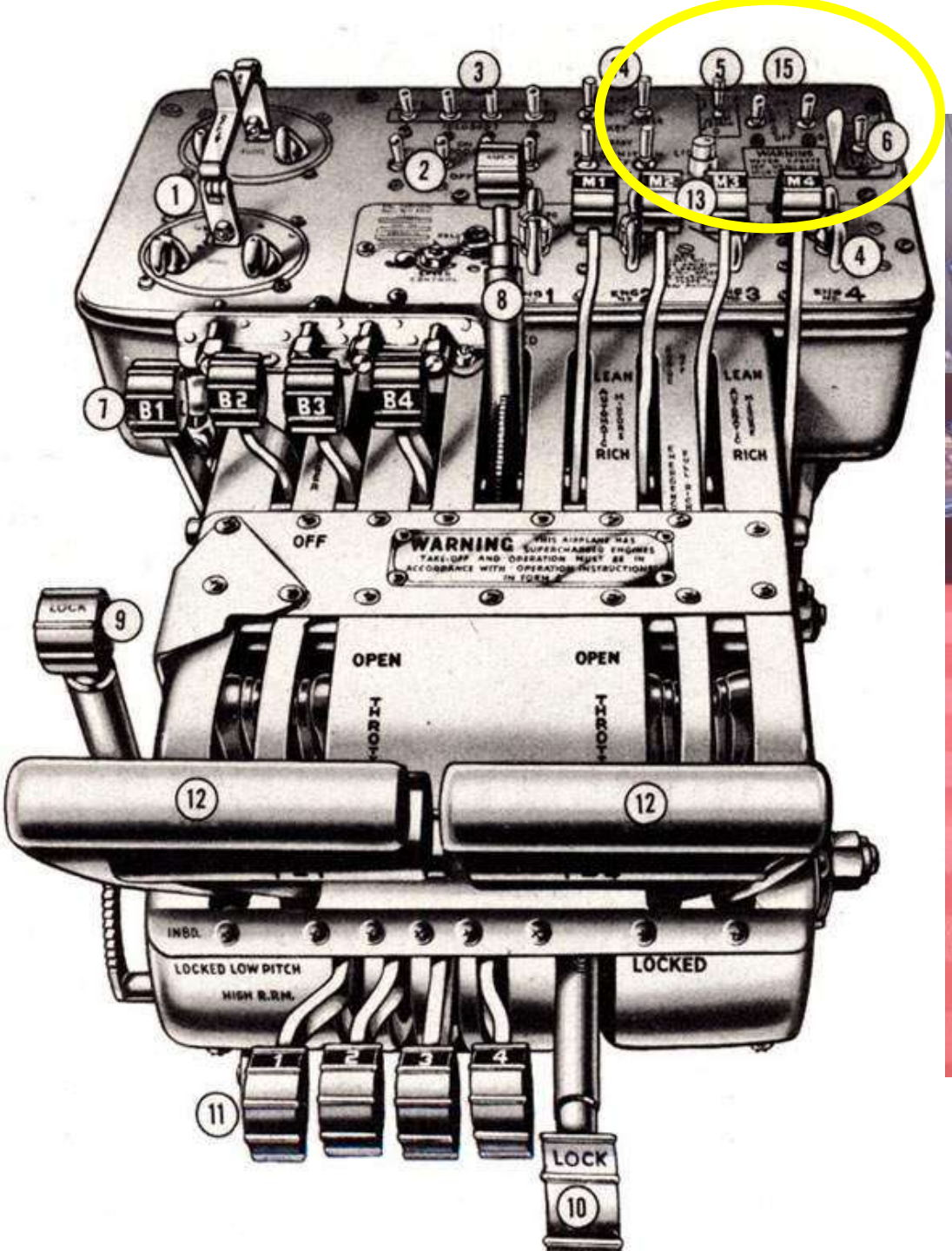
Prevent a catastrophe

Respond to a catastrophe


Language matters



Alphonse Chapanis


1
2
3
4
5
6
7
8
9
10
11
12
13
15
M1
M2
M3
M4
B1
B2
B3
B4
LEAN
RICH
OFF
WARNING
OPEN
THROT
LOCK
INBD.
LOCKED LOW PITCH
HIGH R.P.M.
LOCKED


FEATHERING MOTOR NOT FUSED. IF NO FEATHERING WITHIN 20 SECONDS OPEN THIS SWITCH BY PULLING.
PUSH
PROP FEATHER ENG. 1 ENG. 2
PROP FEATHER ENG. 3 ENG. 4
BLOWN TEST
FUEL PUMP
OFF
WARNING
COWL FLAPS
SPEED CONTROL
RIGHT ENGINES
ENG 1 ENG 2 ENG 3 ENG 4


UP
DOWN


PARKING BRK
SPEED BRAKE
COCKPIT DOOR

Trap #1: Attribution to “human error”

Photograph by Sheila Sund (https://flic.kr/p/FzZFb1)

Humans do make mistakes, but system design, organizational context, and personal context affect when, how and with what impact.

The human didn’t think they were making a mistake. What they did made sense to them at the time.

The problem

"Human error" is a label which causes us to stop investigating at precisely the moment we're about to discover something interesting about our system.

When we hear “human error”
we need to **look deeper**.

Trap #2: Counterfactual reasoning

Photograph by Einheit 00 (https://flic.kr/p/oQ5X91)

“should have,” “could have,” “would have,”
“failed to,” “did not.”

Counterfactual reasoning is telling a story about events that *did not happen*, in order to explain events that did.

The problem

Time spent talking about things that didn't happen is time not spent trying to understand *how what happened, happened*.

Move beyond

"The problem wasn't detected in Canary."

and get to

"How was it detected?"

"What systems or people were involved?"

"How effective is Canary usually when it comes to detecting this kind of problem?"

Trap #3: Normative language

Photograph by Nimish Gogri (https://flic.kr/p/8WXy8B)

👂

“inadequately,” “carelessly,” “hastily”

Normative language judges the decisions and actions of those responding to an incident with the benefit of hindsight.

The problem

Decisions are judged based on their outcomes.

The outcome is the one piece of information we *know* wasn't available to the person who made the decision.

Move beyond

"The relevant code was reviewed hastily…"

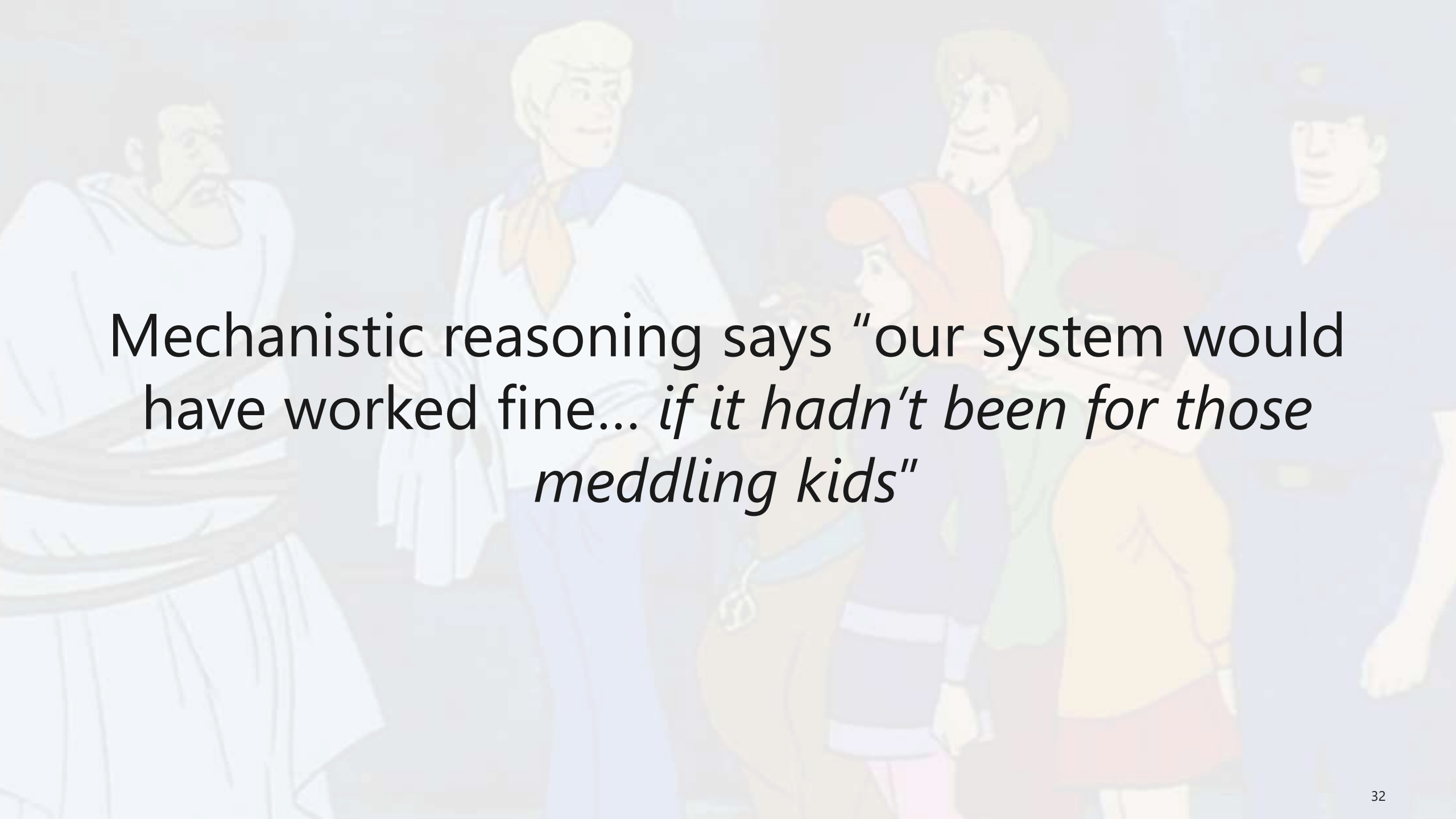
and get to

"What had the previous 48h been like for your team?"

"Tell us how rollouts like this normally work."

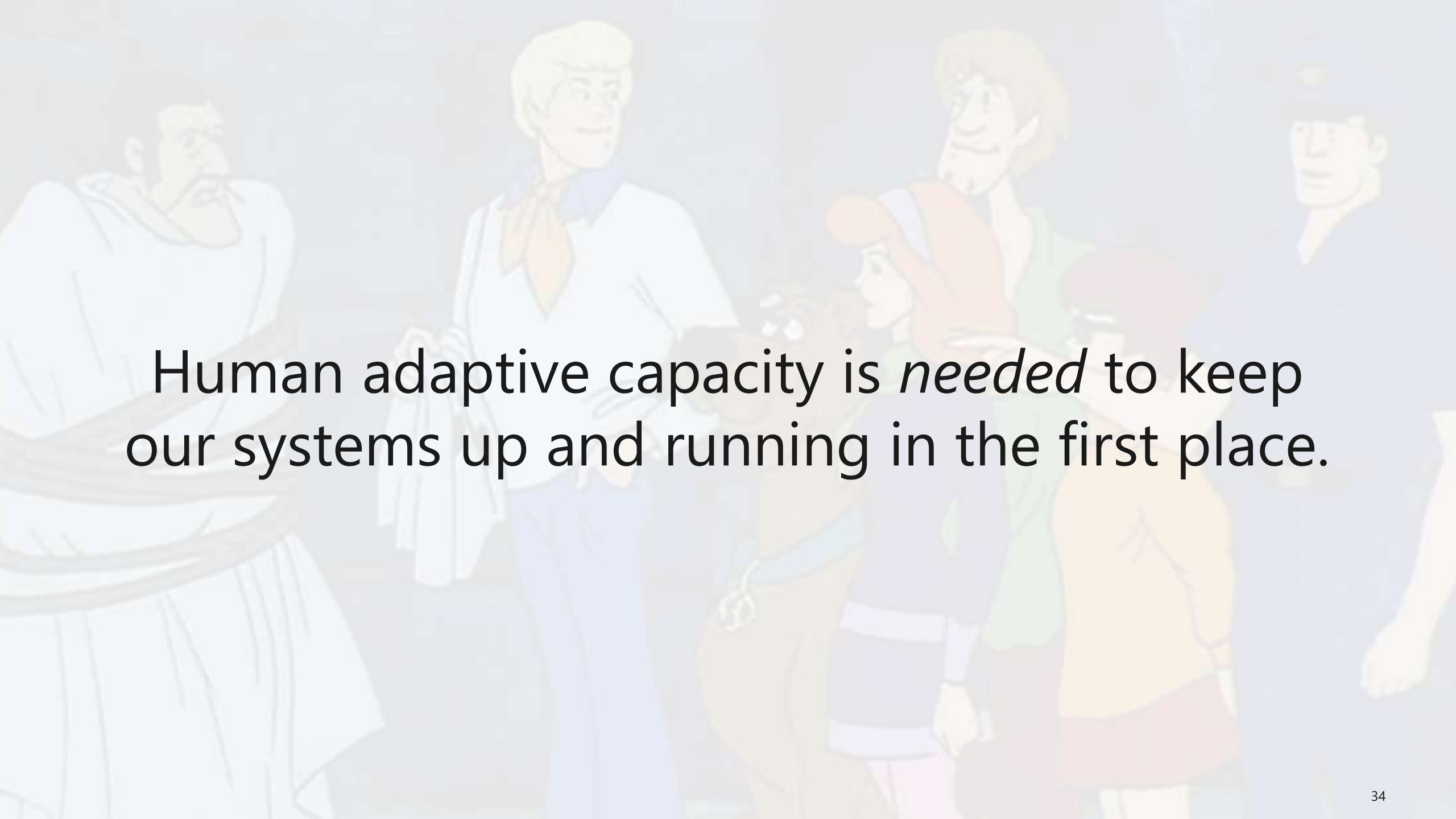
"What signals were you looking at when you made that decision?"

Trap #4: Mechanistic reasoning

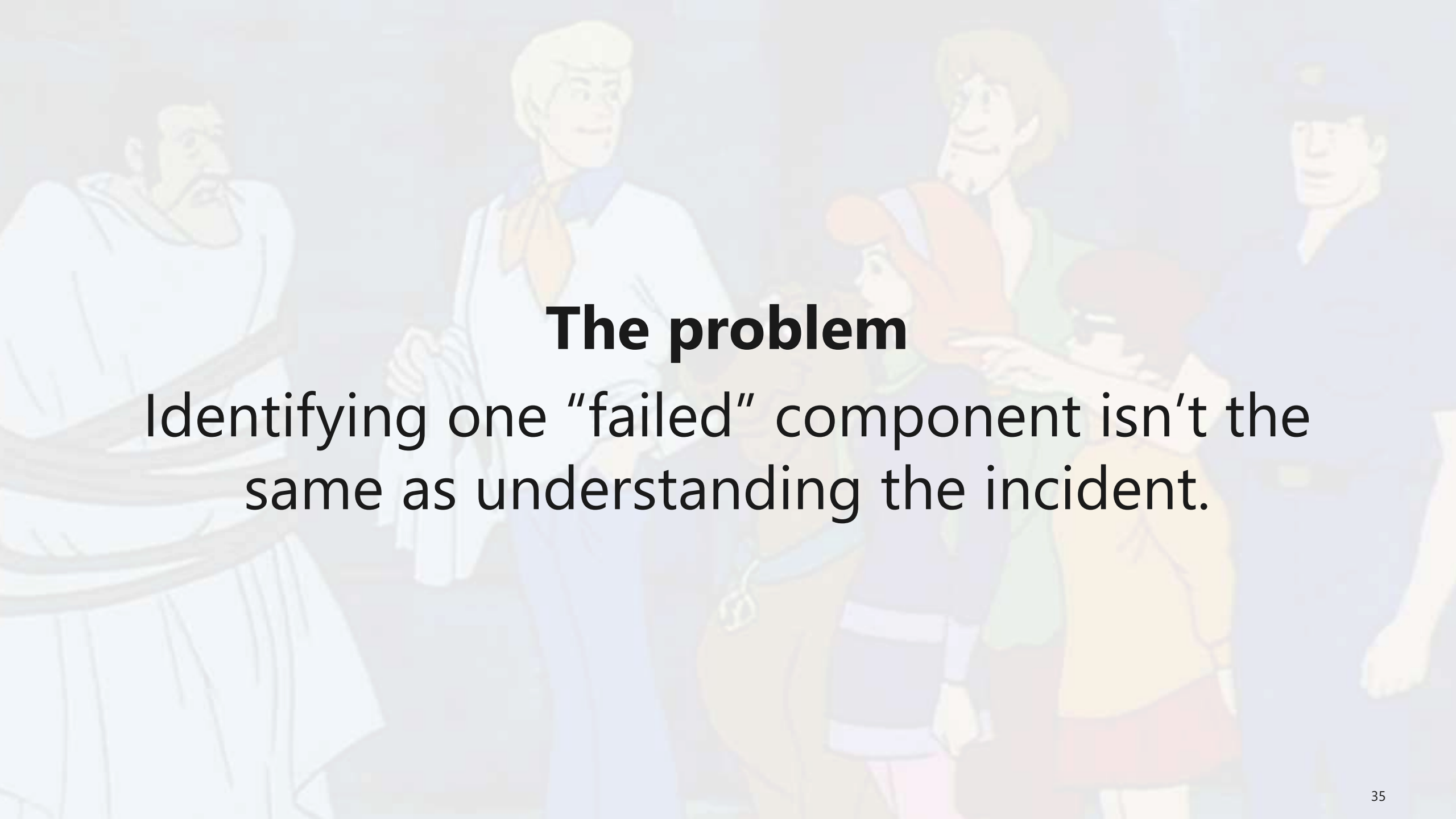


Mechanistic reasoning says “our system would have worked fine… *if it hadn’t been for those meddling kids*"

For how long could your service run without human intervention?



Human adaptive capacity is *needed* to keep our systems up and running in the first place.



The problem

Identifying one “failed” component isn’t the same as understanding the incident.

Agenda

1. Why learn from incidents?

2. Four common traps.

3. Four helpful practices.

1. Run a facilitated post-incident review

Photograph by Melody Ayres-Griffiths (https://flic.kr/p/bnT21X)

Hold a *facilitated* meeting to review what happened with incident participants.

No marathons. 60 to 90 minutes max.

Prepare with one-on-one interviews.

Facilitators should ideally *not* have been involved in response to the incident.

Think carefully before allowing management presence in the room.

Start slow and build confidence.

Pick *interesting* incidents, not big scary ones.

2. Ask better questions

Photograph by Barney Moss (https://flic.kr/p/gmMJ4K)

Language matters: prefer “how?” over “why?”

Each participant has a different experience of the incident. Your job is to help them share these experiences, not to determine “ground truth.”

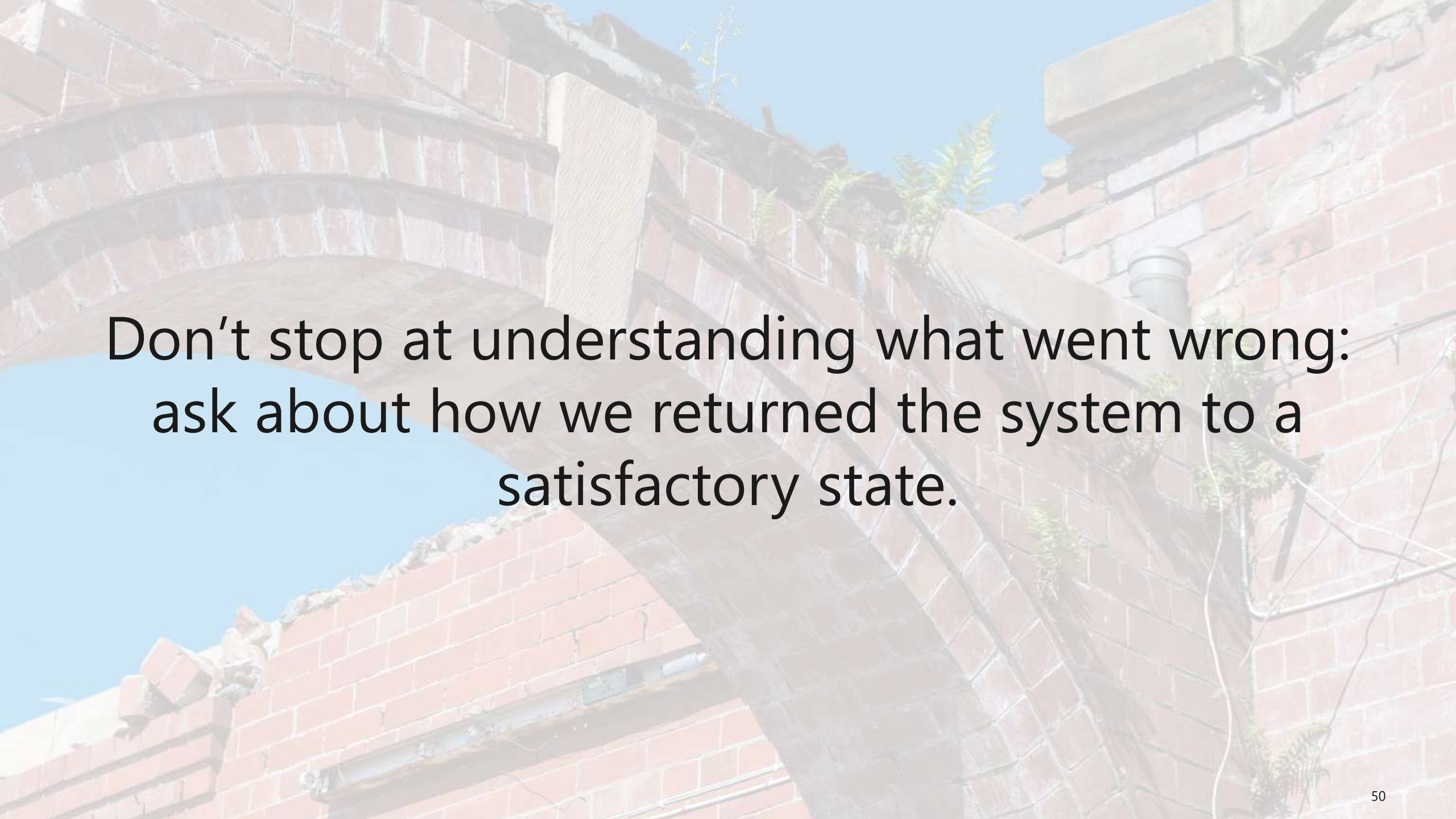
Ask about what normally happens, too!

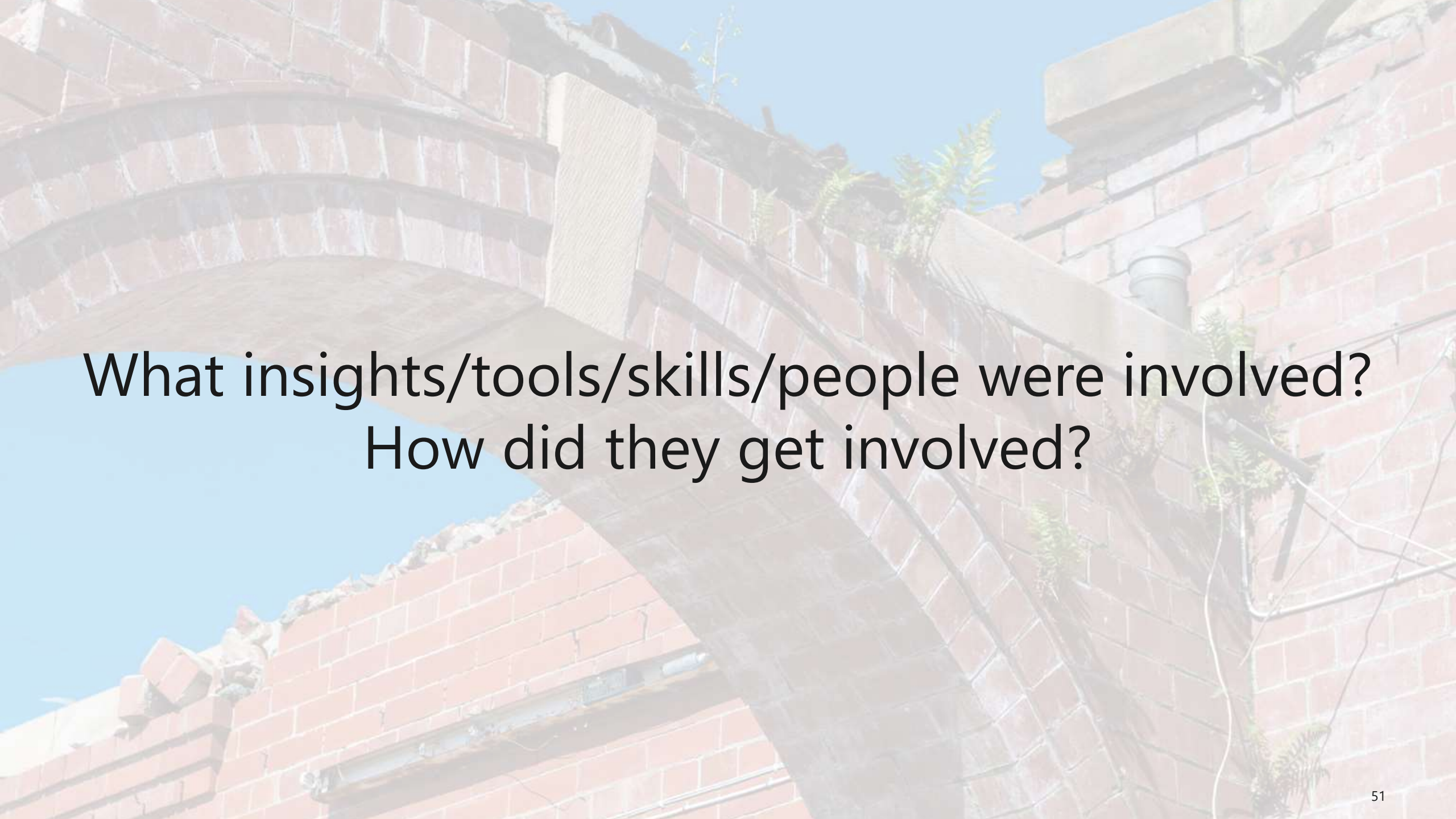
Learn how to be a better facilitator:

https://aka.ms/etsydebriefing

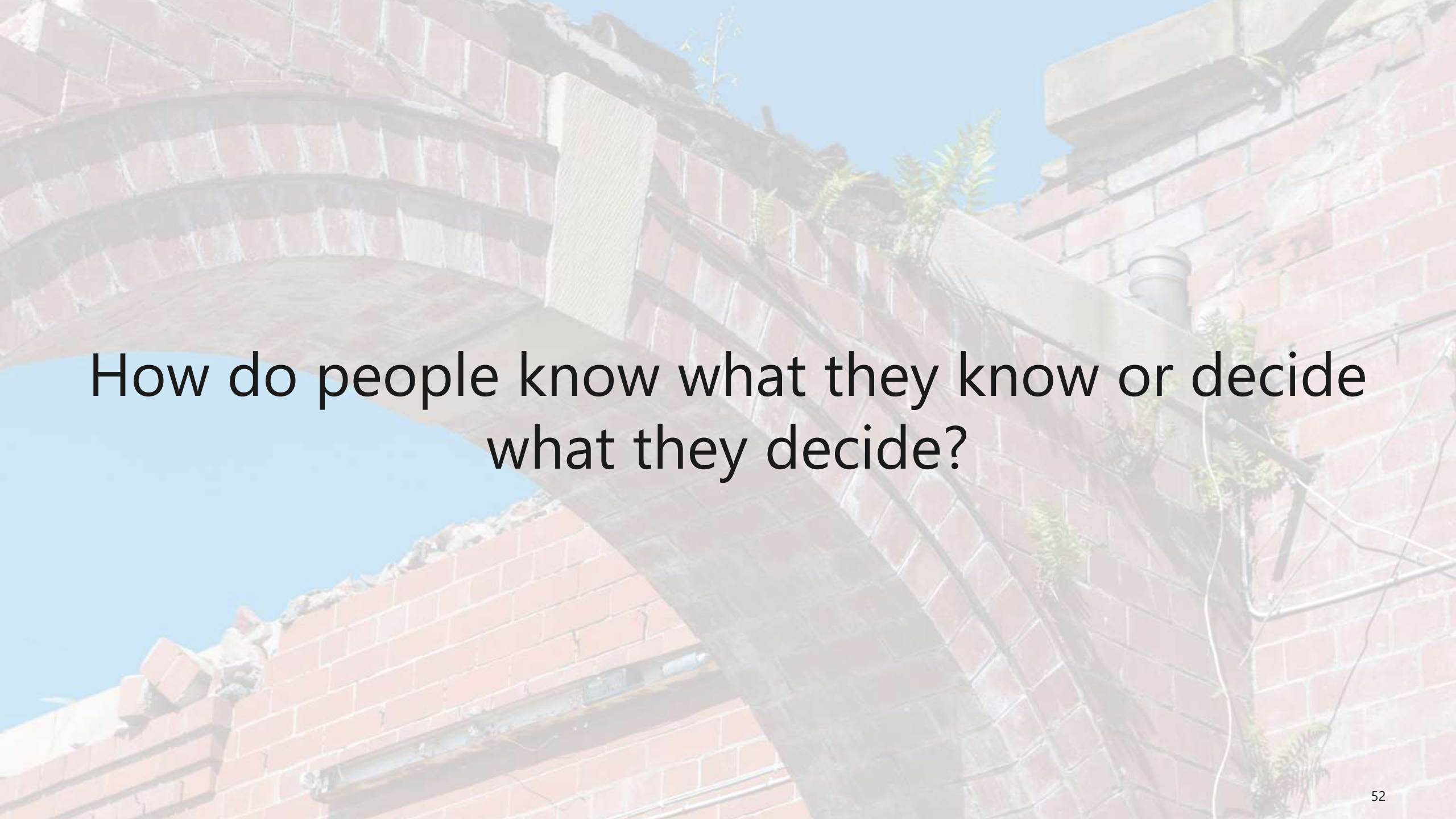
3. Ask how things went *right*

Photograph by Dave Bleasdale (https://flic.kr/p/H9ZgUq)


Don’t stop at understanding what went wrong: ask about how we returned the system to a satisfactory state.

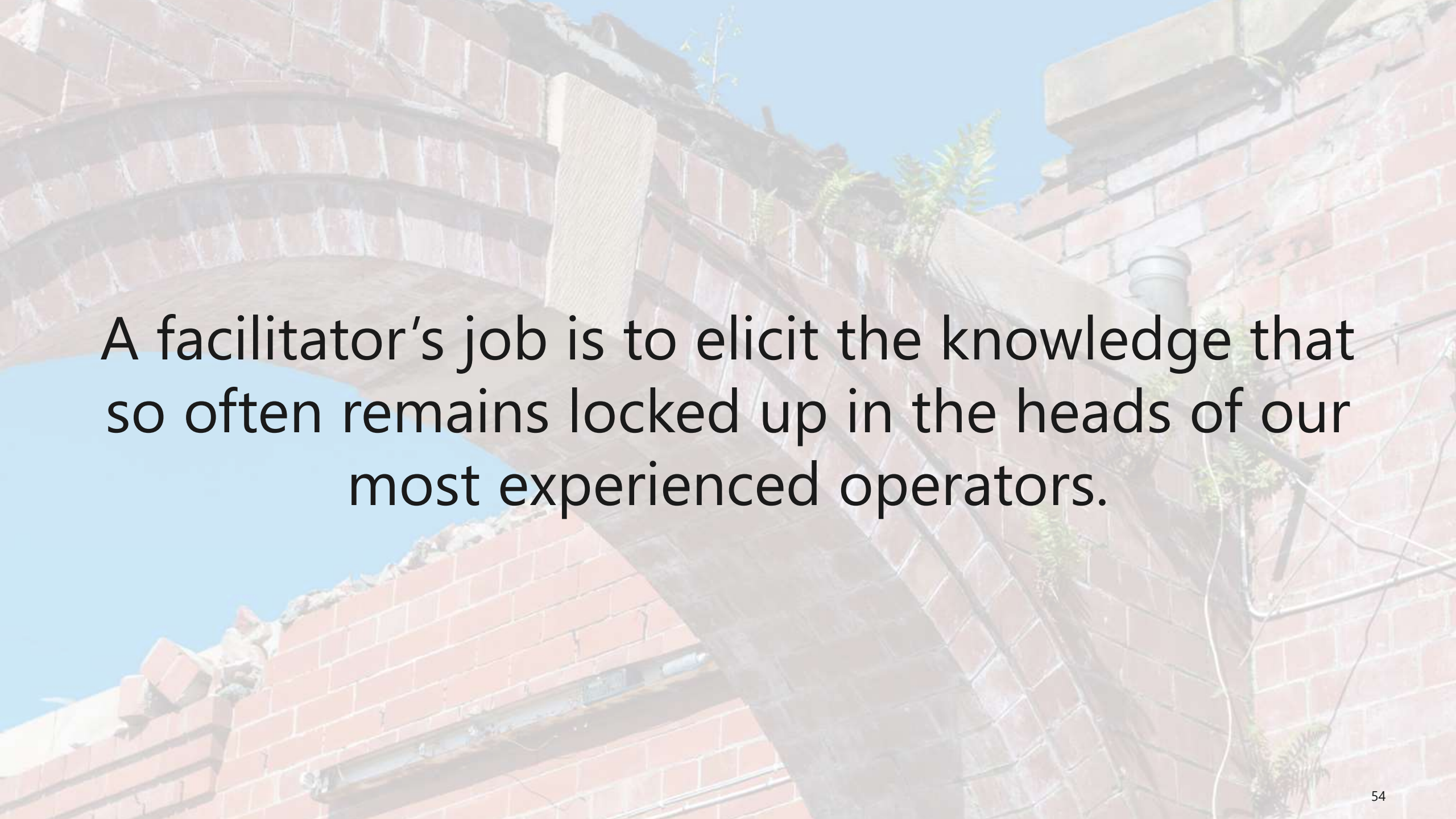


What insights/tools/skills/people were involved? How did they get involved?



How do people know what they know or decide what they decide?

Where did we get lucky?



A facilitator’s job is to elicit the knowledge that so often remains locked up in the heads of our most experienced operators.

4. Keep review and planning meetings separate

Photograph by Ian D Keating (https://flic.kr/p/2c9scLA)

Keep discussion of future mitigation out of the post-incident review.

Hold a separate, smaller planning meeting a day or two after your post-incident review.

Discussion of repair items will *easily* derail an attempt to understand what actually happened.

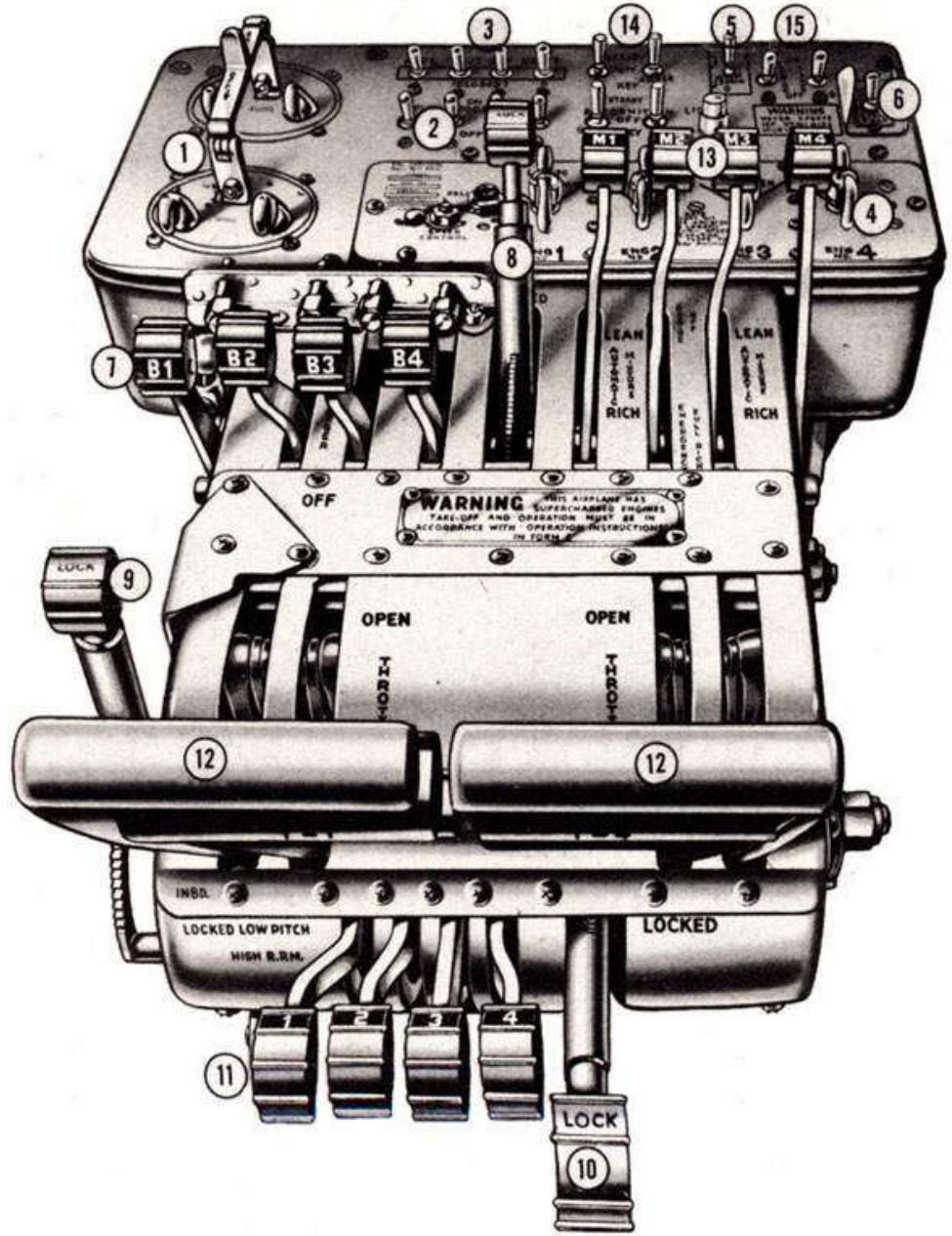
“Soak time” will allow you to identify better areas for improvement.

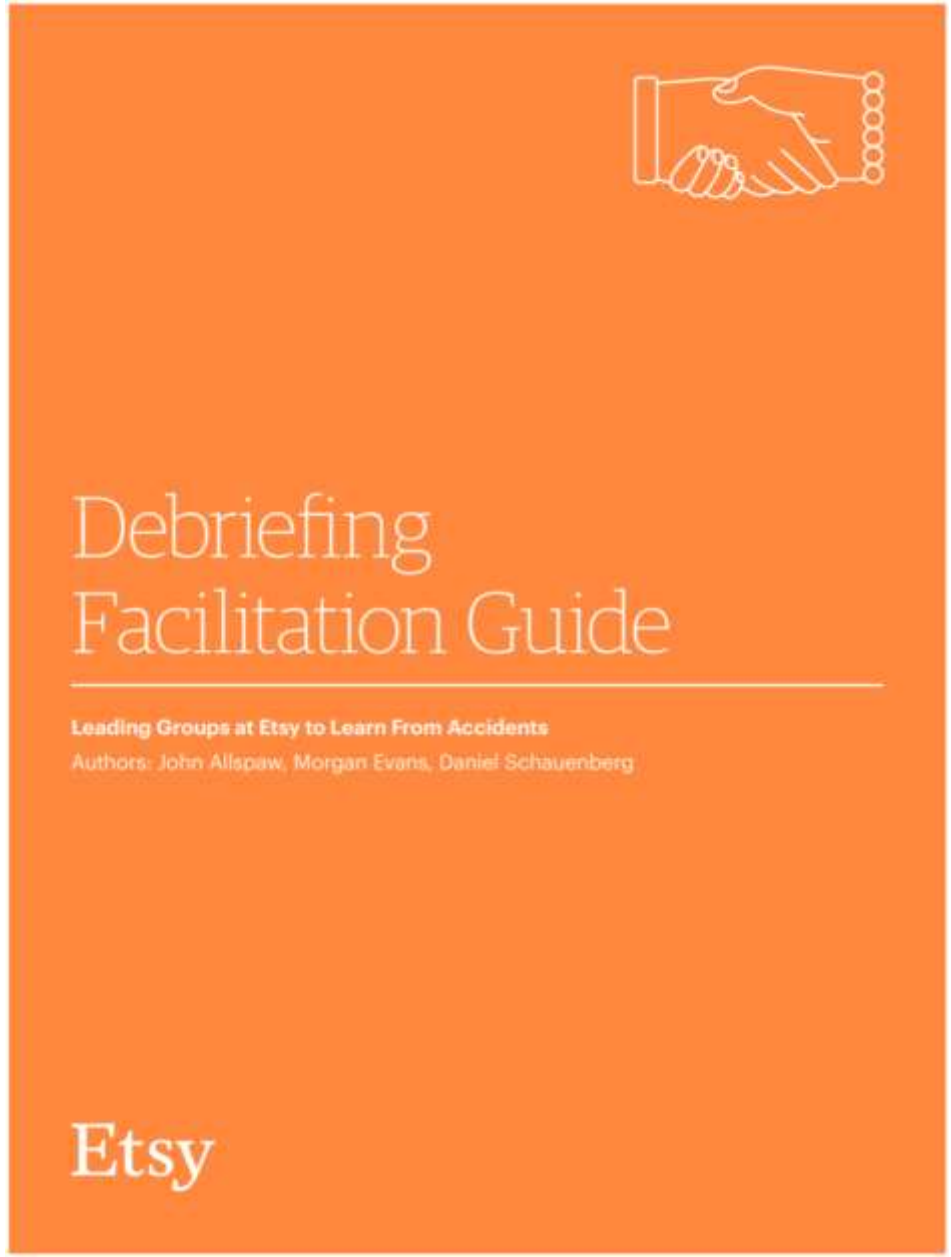
If management or executives need to feel involved, include them in the planning meeting.


100%

⇒


WEST MOUNT
FIRE DEPT.
ENGINE
4


1
2
3
14
5
15
6
4
13
8
7
B1
B2
B3
B4
LEAN
RICH
LEAN
RICH
OFF
WARNING
9
LOCK
OPEN
OPEN
12
12
INBD.
LOCKED LOW PITCH
HIGH R.P.M.
LOCKED
1
2
3
4
11
LOCK
10


Debriefing
Facilitation Guide
Leading Groups at Etsy to Learn From Accidents
Authors: John Allspaw, Morgan Evans, Daniel Schauenberg
Etsy

Thank you

https://aka.ms/leaddev/lfi

New “learning from incidents” community: @LFISoftware