catchpoint

The SRE Report 2024

Making IT better.

Sixth edition

Foreword

When I was studying Computer Science at UC Santa Barbara (go Gauchos!), I had a part time job as a SysAdmin for a small department on campus. I split my time between programming and administration: while I was provisioning and updating machines at work, I was implementing a compiler. I pulled ethernet cables in closets then tinkered with adding TCP features onto UDP. I kept a busy mailserver buzzing along for a multi-campus research program while I studied Lamport Clocks. SRE hadn't been invented yet, but little did I know that I was developing all the right knowledge and skills for it.

Today when I see people wringing their hands about overlapping roles and titles like SysAdmin, DevOps, SRE, or Platform Engineering, I get it. SREs are descended from both System Administrators (SysAdmins) and Programmers. While the Programmers wrote product code, SysAdmins were responsible for "everything else." It's hard to define all the "everything else" stuff and we tend to get caught up in it.

SREs consider "everything else" through a single, broad word: *reliability*. This gets confusing when you compare it with similar words like *robustness* or *resilience*, but for now let's just use reliability to simplify "keep the system running." It's not the same as "keep bad people out" (security) or "build things people will love" (product). As a team grows, each of these large areas becomes the realm of a different set of people; specialization emerges. A complex dance of people and responsibilities results in an often poorly-defined map of who does what. It can be tempting to silo these teams off at some point, to lock it in stone. "These people do X, not Y." But what if that doesn't work well, over time? Can you break up that silo and evolve? Too often, we imagine it is the titles and roles that keep us from evolving, so we invent and adopt new ones. But we have to remember to change the work, not just the name.

Constraints, like individual roles or a team's scope, can bring freedom. Knowing you must work only up to a certain line, then someone else takes over, frees you from having to do everything all the time. People like having a defined scope. But when a system evolves - but constraints and models don't keep up - you start to get misalignment and gaps. You can get overlaps and overhead. So, how can we prevent this? We can focus on end-to-end outcomes, not in-silo metrics. We can take measurements *across* those constraints, not *within* them. Take a look at the metrics you use today to evaluate your success. Do they span teams? Do they make sense to the end user?

A bad example of measuring within a silo is hyperfixation on efficiency or "goodness" of a single thing. Make sure you consider if a metric has meaning, not just if it is easy to measure. For example, MTBF and MTTR are popular and make sense for mechanical systems, but they don't actually work in the context of a complex digital system. In these, components don't just fail due to a few well-known wear patterns, but via a wide variety of subtle failure modes instead ("Plato's Dashboards", Fred Hebert). If we don't adjust our constraints and our metrics appropriately, we end up performing actions that no longer deliver value. Teams get caught between what they know makes sense and what their collective history had decided was "right" at one point in time. When we evaluate teams with the wrong metrics, we can wind up incentivizing and rewarding toil, burning people out, and causing real harm.

Don't try to do this just on your own. Learn from others. Surveys and reports like this one are a great resource, but don't forget to use your critical thinking. Know where you're getting your advice from. Don't just adopt models without understanding them. By keeping an eye out to what others are doing, by attending conferences, by subscribing to newsletters, and by reading and contributing to surveys and reports like this, you will better understand complex systems of humans and machines, improving their reliability.

Please contribute and help build this community of practice, as it needs many different voices to really succeed. After reading this report, I encourage you to share your thoughts with your teams, with your communities. How does this change your view of the past? What do you want to do in the future? Who will help? Who will help you? I genuinely thank Catchpoint for continuing to conduct this research every year, as our community grows and adapts, no matter what names we use.

Steve McGhee

SRE, Reliability Advocate, **Google**

Introduction

At Catchpoint, we are honored to once again present our annual SRE Report, a tradition that stands as a testament to our dedication to providing in-depth, independent assessments in the realm of reliability. This year marks the sixth edition of our report, which is shaped by the insights gathered from The SRE Survey. Our participants predominantly included individual contributors, but we were also pleased to welcome a substantial number of inputs from reliability leaders across various levels of management.

The significance of Site Reliability Engineering (SRE) in fostering enhanced customer satisfaction, robust resilience, and reduced incidents, among other benefits, is undisputed. Our SRE Report offers unique insights into the practical application of reliability, equipping you with valuable knowledge for enriching your business dialogues.

The production of this year's report, including the survey itself, saw an unprecedented level of community involvement and feedback. A common observation from our contributors was that interpretations of the data vary based on the reader's perspective. Consequently, we have chosen to minimize prescriptive advice in our current edition, opting instead to present the data in its purest form, enriched with additional insights from practitioners in the field. We trust you'll find this approach both refreshing and empowering.

Looking ahead to next year, we anticipate celebrating the appointment of numerous Chief Reliability or Resilience Officers in Fortune 2000 companies. With billions invested annually in observability, it's imperative for companies and their boards to integrate a resilience dimension, akin to their approach to security, given the equally significant risks and rewards.

Mehdi Daoudi

CEO, Catchpoint

The SRE Report 2024

Key Insights

This year's report falls into seven sections, yielding new insights into reliability practices and opinion on a wide range of topics. Areas explored range from hot trends like the perceived value of AI and the need to learn from incidents to SRE fundamentals like the amount of time spent on engineering activities.

Here's a sneak peek of the seven insights we'll be digging deeper into:

Insight I

Loss of Control Creates New Opportunities for Relationships and Learning

Insight II

SRE is not Platform Engineering, But They Both Develop Capabilities

Insight III

Learning from Incidents is a Universal Business Opportunity

Insight IV

AI is not Replacing Human Intelligence Anytime Soon

Insight V

When it Comes to Service Levels, Ignorance is Bliss (in Smaller Companies)

Insight VI

No Single Monitoring Tool Does It All (and in the darkness bind them)

Insight VII

Efficiency is the Enemy of Pride

Insight I

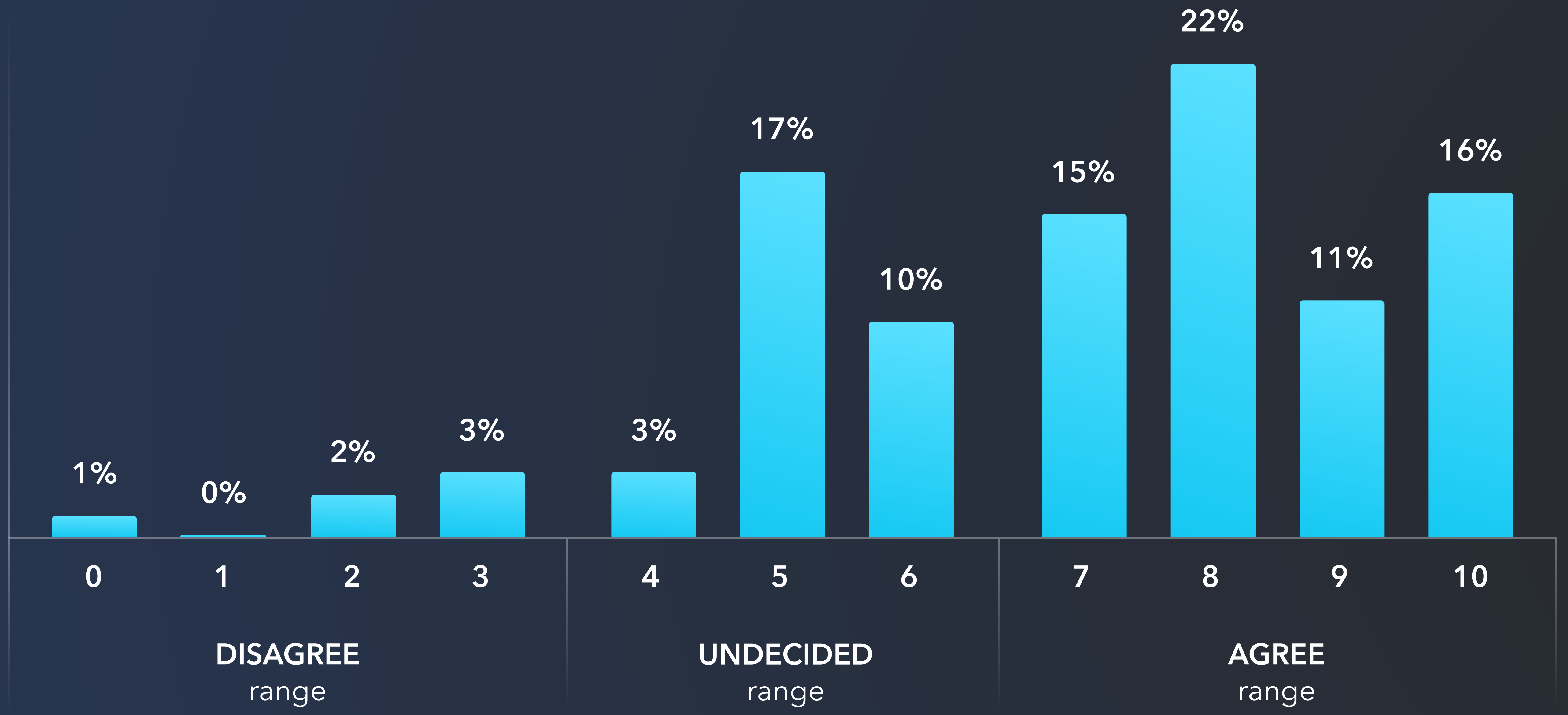
Loss of Control Creates New Opportunities for Relationships and Learning

64%

of organizations believe they should monitor productivity or experience-disrupting endpoints, even if they lie outside their physical control.

In a rare alignment, the majority of all organizational ranks sat in the 'agree' range around the need to tactically monitor productivity or experience-disrupting endpoints, even when outside their physical control.

Do you agree reliability practitioners should tactically monitor productivity or experience-impacting endpoints? Even if outside their physical control*?



Modern reliability practices will increasingly federate third-party services. This has implications from individual contributors all the way up to line of business owners. Consequences also resound in the wider ecosystem where software development, as well as computing resources, are being bought as-a-service.

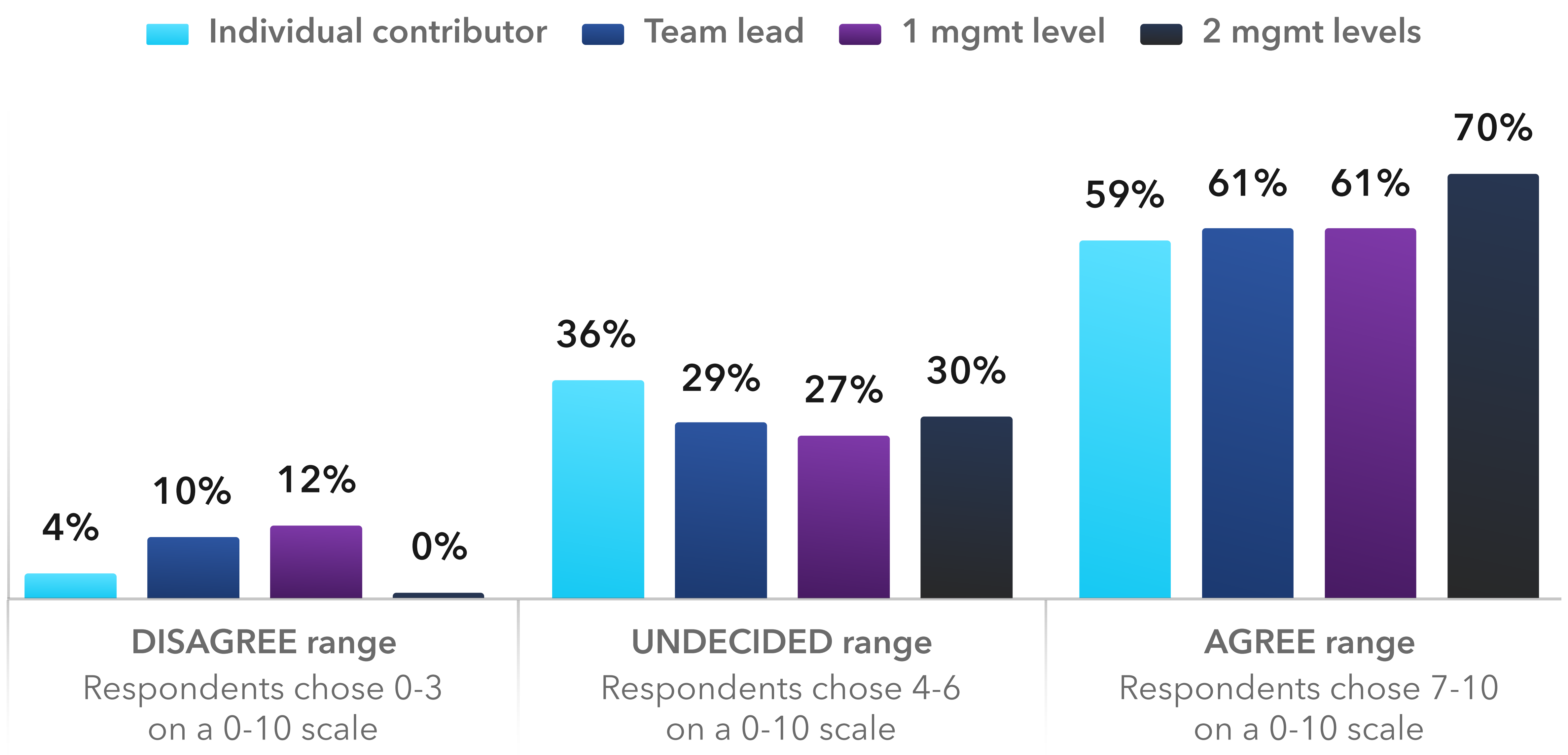
This data illustrates industry awareness of this dynamic and the need to incorporate third-party services within the scope of reliability. As such, reliability engineering must adapt to this multi-party way of building services.

**Additional survey context:*

There are many endpoints outside of our control which, when they go into incident, will have either external or internal impact (e.g., on productivity or customer experience). Examples are SaaS apps (e.g., Microsoft 365/Google Workspace), inter-cloud connectivity, or your home/residential ISP.

Reliability is a team sport.

Do you agree reliability practitioners should tactically monitor productivity or experience-impacting endpoints? Even if outside their physical control (by rank)?

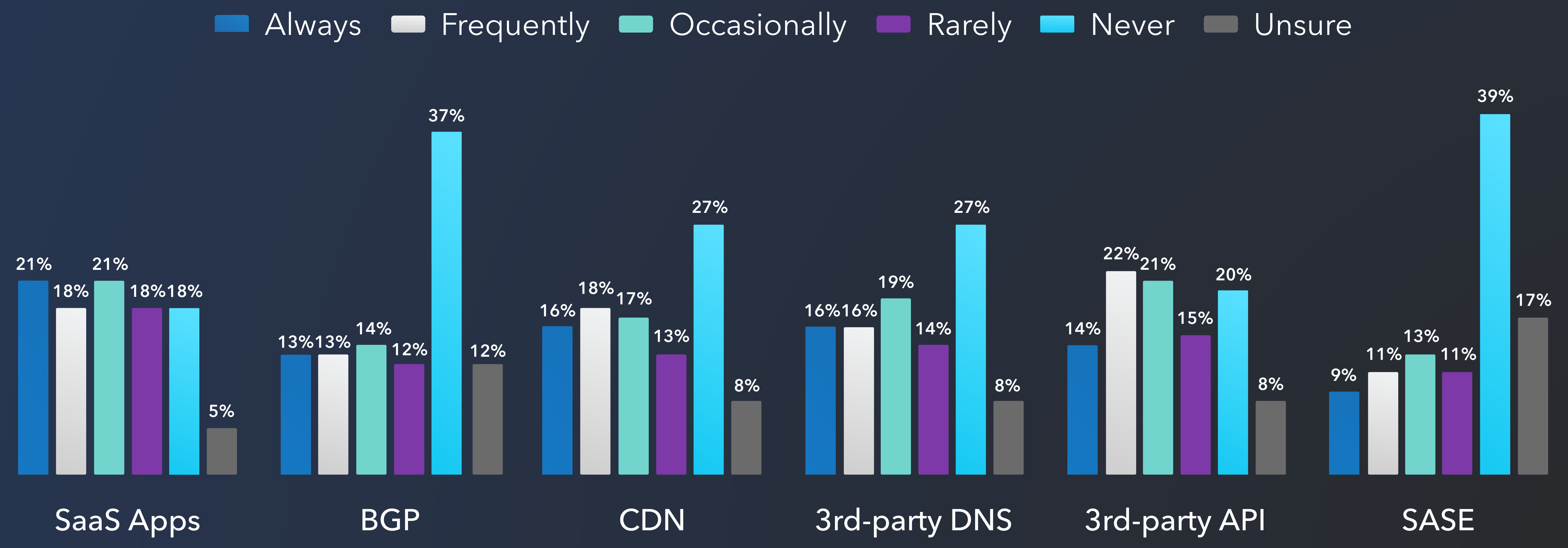


In a rare alignment, all ranks of the organization agreed on the need to tactically monitor certain endpoints, even those outside their physical control. Just as with last year's report, we broke down answers to various questions based on rank of the respondent. We decided to selectively do the same this year. However, for such breakdowns, we had to exclude management levels three, four, and five/five+ since there were not enough respondents in those categories for meaningful insight.

In the survey, one level of management was indicated as 'I have one level of direct reports' and two levels of management was indicated as 'I have up to two levels of direct reports'. In other words, two management levels was intended to mean 'a director with a manager reporting to them, and people reporting to that manager'.

Based on this intent, we can see an upward trend of agreement to this question as organizational rank increases.

Do you or your team monitor any of these?



SaaS Apps	60%**
3rd-party API	57%
CDN	51%
DNS	51%
BGP	40%
SASE	33%

** sum of occasionally, frequently, and always

The need to drive reliability and resilience will increasingly necessitate the inclusion of third-party vendors in monitoring strategies. After all, a productive approach to addressing and solving problems is for SREs to work collaboratively with third parties.

Any single point of failure in the complex web infrastructure that makes up the Internet, which we call the Internet Stack, can break a system. As the Internet becomes more complex, so too does what we have to monitor.

We see this area of reliability work as an opportunity to both build better relationships beyond the "four walls" of the company and to improve learning about those lesser-monitored areas of the Internet Stack, such as BGP and SASE. While it is impossible to list all possible endpoints in a 'do you monitor these' question, this insight should be used to deepen – or create - a conversation around incorporating third-party vendors into playbooks.

** Additional survey context:*

SaaS Apps implies "provided by somebody else"; the provider may be within the same organization. For example, if you are not responsible for maintaining or supporting Microsoft365/Google Workspace, do you still monitor them? SASE is Secure Access Service Edge.

View from the field

As services become increasingly advanced, there has been a marked shift toward outsourcing the maintenance and operation of critical components, often to third-party vendors. This loss of control leads to increasingly decentralized architectures. By shifting some responsibilities across organizational boundaries, teams can focus on core competencies while relying on the expertise of others in non-core areas. This dynamic is often necessary and helpful since it introduces new ways of working in complex sociotechnical systems. It also introduces new forms of failure.

The "multi-party dilemma" is a dynamic that describes challenges at the boundaries between interdependent parties, particularly during temporally compact, high-consequence events such as incidents. Much of the time, reliance on these vendors "just works." Much of the day-to-day operational burden is handled by one side of the boundary. The burden of coordinating across organizational boundaries to handle these shared systems is usually minimal. However, during anomaly response, all parties are immediately presented with the challenges of working across organizational boundaries. These challenges include increased cost of coordination, asymmetry concerns, and goal misalignment.

Successfully integrating third-party vendors involves recognizing and embracing new forms of failure and ways of working together. While research is still ongoing regarding the multi-party dilemma, initial findings point to activities that help manage this new dynamic which aim to increase reciprocity, establish shared frames of reference, and reduce asymmetry. Engaging in these activities before a critical event occurs can help organizations more successfully manage some of the challenges introduced by the multi-party dilemma.

Examples of these activities include:

- Preparing to have cross-boundary incidents through joint game days or tabletops
- Developing efficient means of bi-directional communication between organizations
- Conducting knowledge elicitation and subject-matter deep dives to discover areas such as loose vs tight couplings and layers of hidden transitive dependencies
- Sharing additional context on the configuration, health indicators (including monitoring), and use of systems as well as collaboration on shared processes (such as initial troubleshooting steps) through venues such as office hours

Sarah Butt

Director, Site Reliability Engineering

SentinelOne,

Alex Elman

Director, Site Reliability Engineering

Indeed,

See also:
https://youtu.be/CbSiKAtO7Fk
https://youtu.be/Veq7VUbPwWo
https://www.researchgate.net/publication/376354074_Handling_the_Multi-Party_Dilemma

Insight II

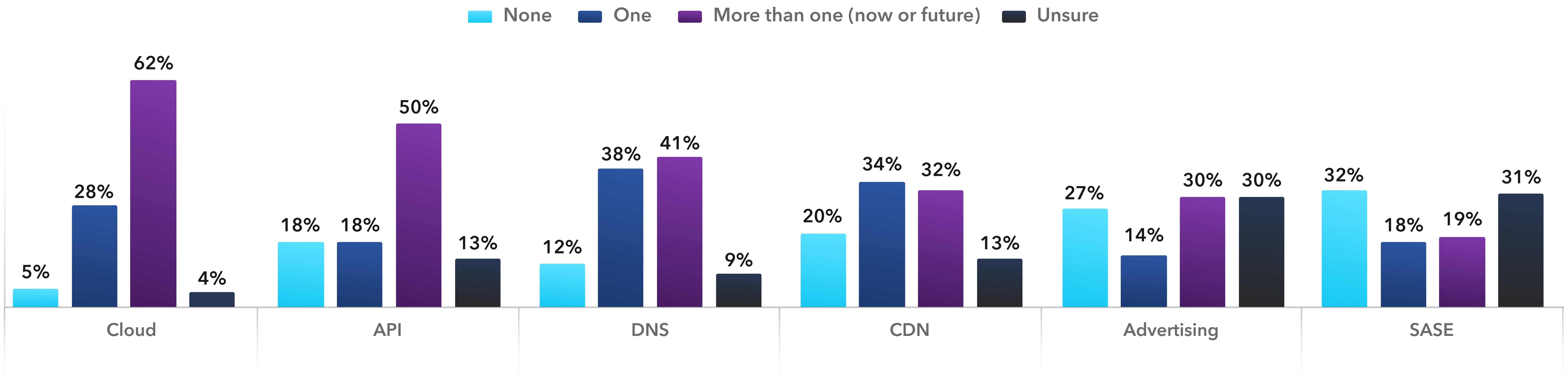
SRE is not Platform Engineering, But They Both Develop Capabilities

44%

of organizations said they employed a team structure ‘by platform or capability’.

The use of this team structure trends upward as company size increases.

How many third-party providers does your organization use for the following service categories?

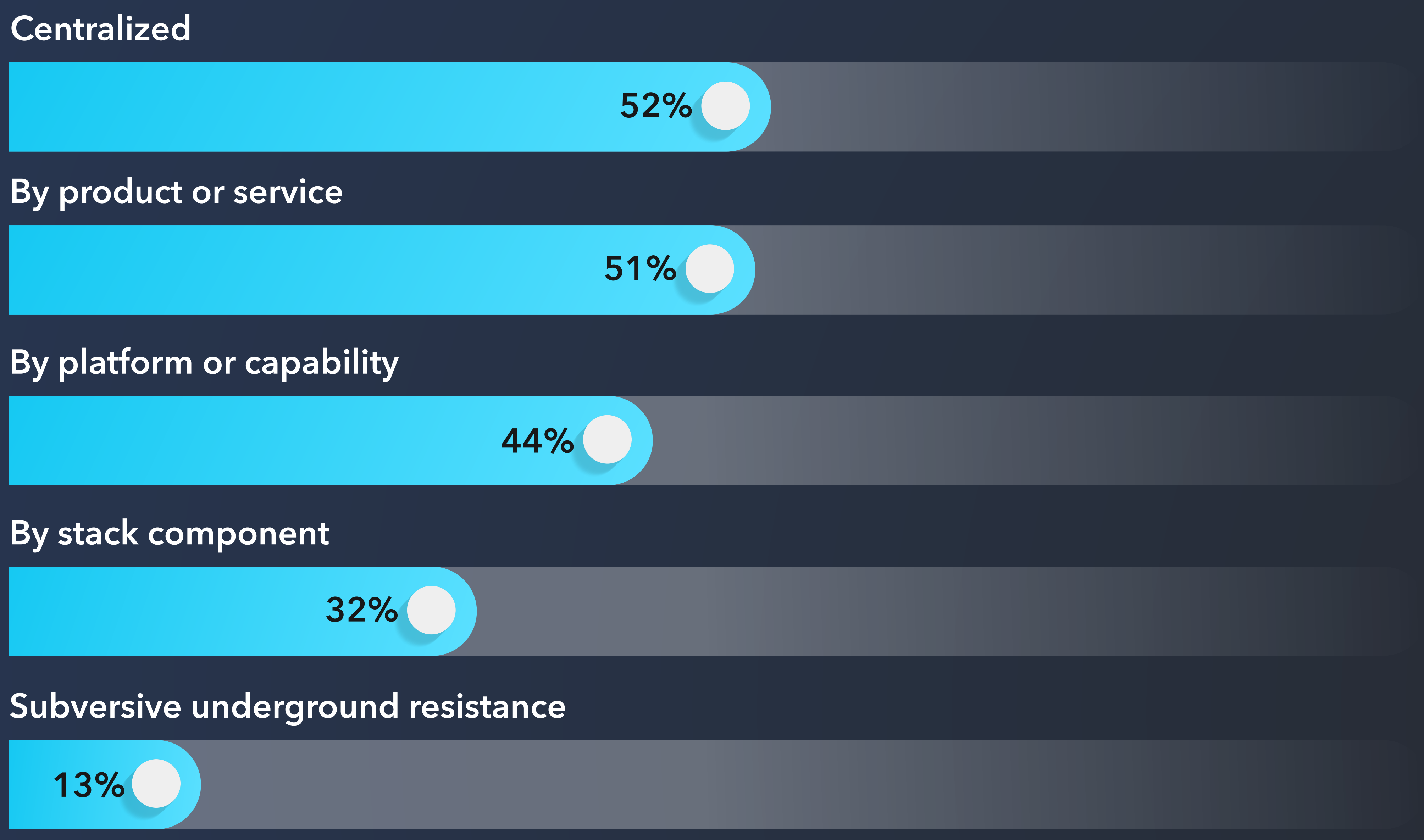


Engineering activities may include the development of tools, workflows, or capabilities. They may either be for the consumption of the person who developed them, offered to your internal workforce or team members, or offered to third-party augmented staff.

They may also (probably) include reliance on third parties. Building on the lack of control theme, we wanted to explore the use of third parties for portions of the Internet Stack, even though we did not explicitly ask who their consumers are.

Take, for example, the (potentially obvious) use of multiple cloud providers (62%). Now consider a platform capability that may engage with a third-party API (e.g., for authentication). Or consider a popular tool that may be fronted by a third-party provider for DNS resolution. As such, teams will want to consider the relationship and impact of using third-party providers. For example, how do organizations monitor these Internet Stack components, what happens when they go into incident, or will they provide the flexibility to adjust as business needs change.

Are any of these team structures used within your organization?*



*Additional Survey Context

Centralized
supporting various products/services, platforms, and/or stack components

By product or service
E.g., we have a dedicated team supporting our maps product

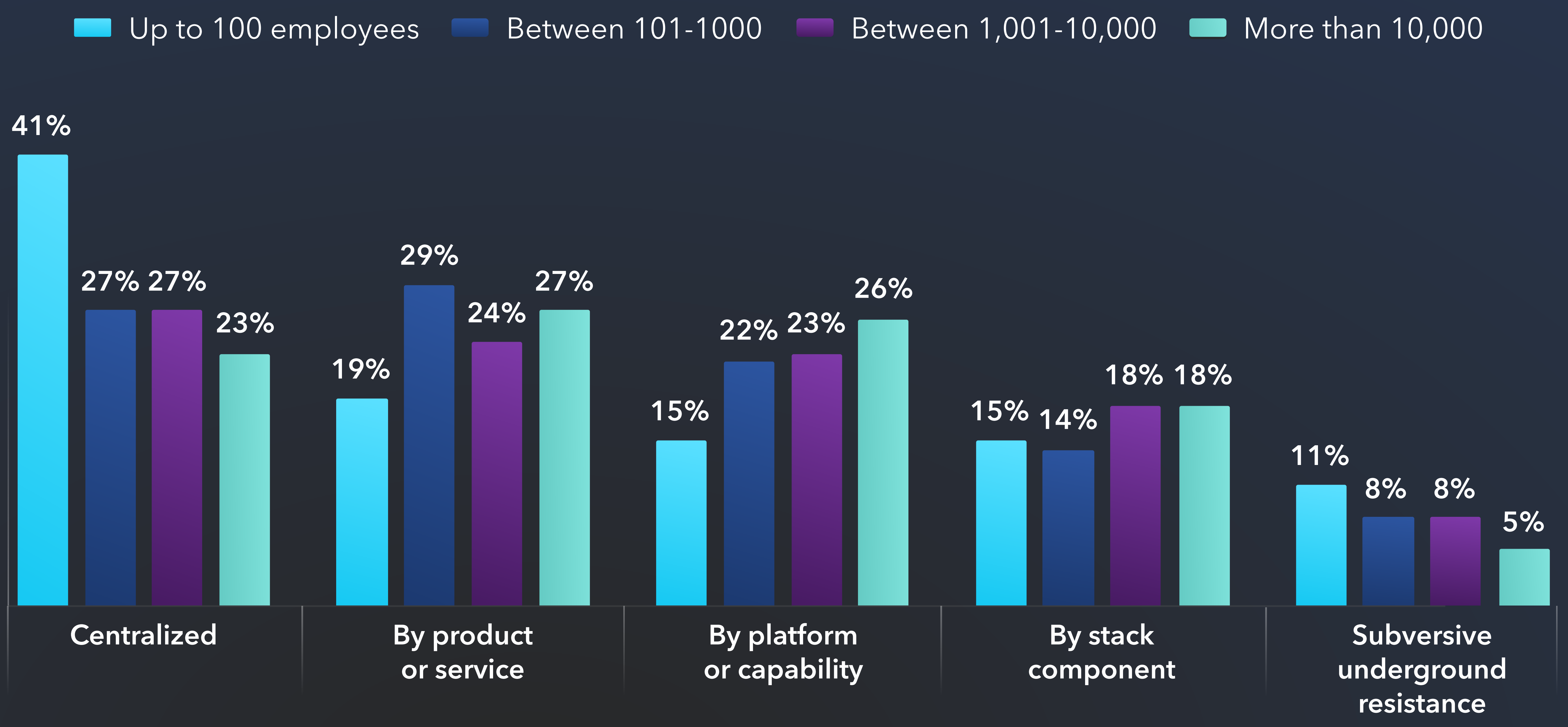
By platform or capability
E.g., we have a dedicated team building capabilities on top of our cloud platform(s)

By stack component
E.g., we have a dedicated team maintaining application or infrastructure stack(s)

SREs and general reliability practices have many layers or dimensions. We asked respondents about how different team structures are used within organizations and offered the opportunity to cite more than one example type. Just over half said they have centralized teams while another 51% (some of which may overlap) said they have teams aligned by business product or service. 44% cited teams being aligned by platform or capability offering. However, note these numbers directionally trend when broken down by company size (see next page).

In addition to team structure(s), culture or other contributing factors may also affect how a new problem or challenge is approached within different companies.

Are any of these team structures used within your organization (by company size)?



By their very nature, reliability practitioners perform both operational and engineering activities. They may respond to incidents and/or they may develop automation capabilities to make their job more efficient.

In this visual, the use of centralized team structures trends downward, and the use of 'By platform or capability' trends upward as company size increases ('By product or service' also trends upward as company size increases). We explicitly do not indicate that the use of platform or capability structures means Platform Engineering. However, we do intend this data to be used as an index if businesses are having a Platform Engineering conversation.

View from the field

Site Reliability Engineers will develop an array of assets to make their job easier. These assets may include tools, (what even is a tool, anyways?), scripts, documentation, workflows, or other types. When they develop for themselves, there may be less elegance because function is more important than form. But if an asset enables a better level of efficiency, consider whether it can enable capabilities beyond the originating team(s) and then make it available to other members in the organization.

This concept of offering capabilities to other members of the organization is not meant to distinguish between site reliability engineers, platform engineers, or aromatic flower engineers. Instead, it's meant to act as an opportunity to make the organization better. If tasks are assigned based on propensity to succeed, that is more important than any label, title, or moniker.

Ultimately, we are all trying to solve problems. Whomsoever is best suited to solve those problems should be part of the path creation. Titles should not matter whether they are SREs, DevOps engineers, product managers, or marketing evangelists. Instead, when the benefits of structuring a certain way manifest in measurables and tangibles, that is a win.

Kurt Andersen

Infrastructure Software Architect, **Clari**

Insight III

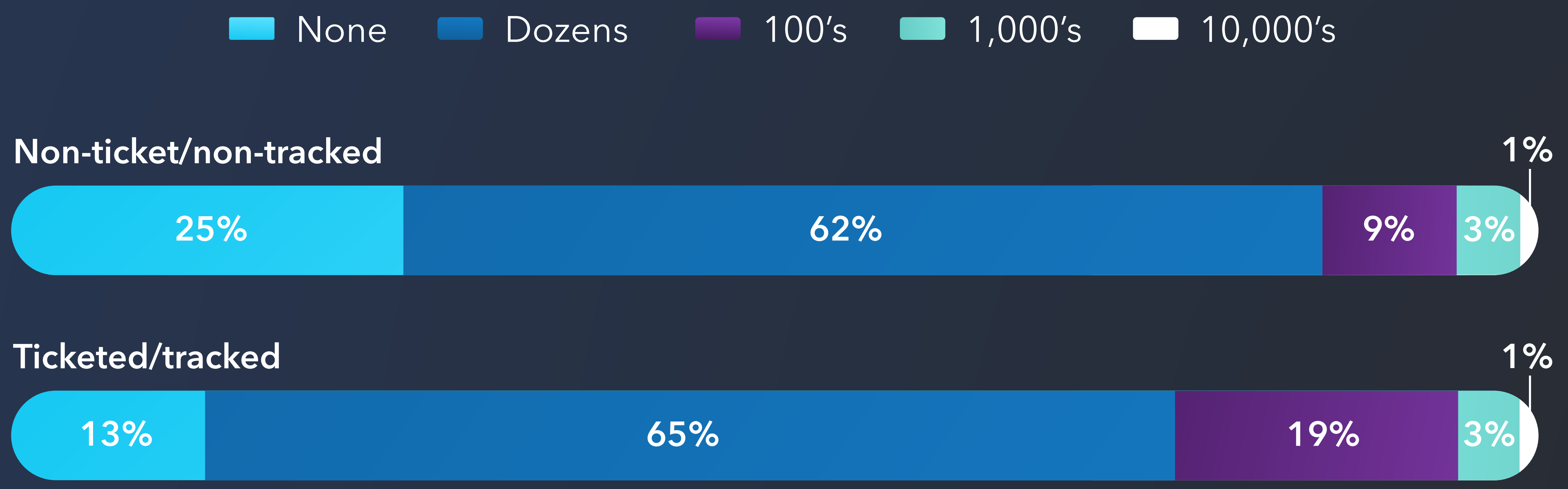
Learning from Incidents is a Universal Business Opportunity

47%

say learning from incidents (LFI) has the most room for improvement in overall incident management activities.

This number does not substantially trend up or down as company size increases.

How many incidents have you or your team responded to in the last 30 days?



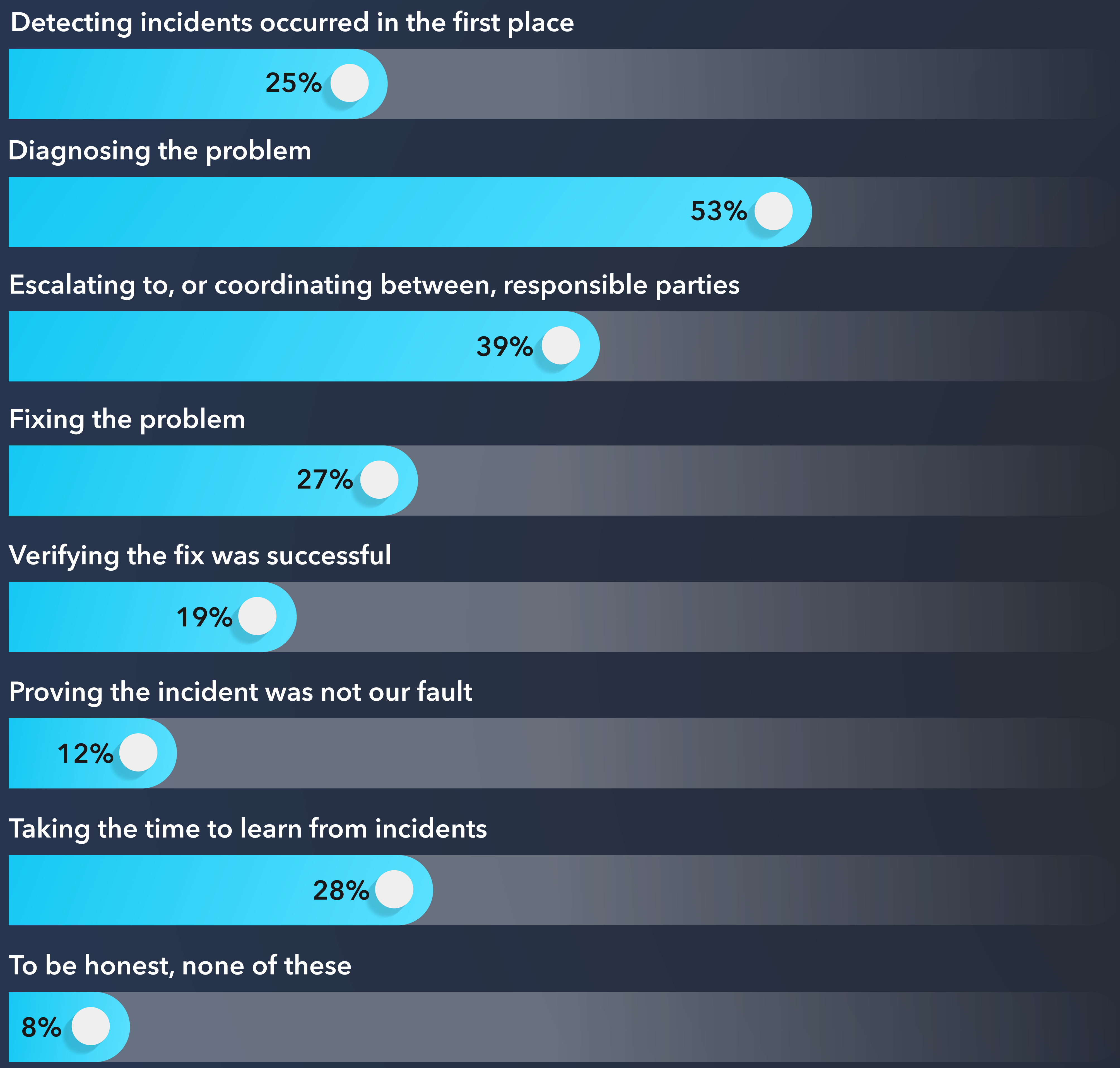
It's not a matter of **if** you will have to respond to an incident, it's a matter of **when**. To understand the importance of learning, we first wanted to get an idea of magnitude in terms of the volume of incident response work.

What we found was that 71% and 84% of SREs respond to dozens or hundreds of non-ticketed and ticketed incidents a month, respectively.

The width of the survey buckets e.g., dozens versus hundred versus thousands, were intentionally large. Note this means a delta in the percentage is larger than it appears.

Since non-ticketed incidents is as large as it is (for the dozens and hundred series), it appears that reliability practitioners may not be getting full credit for their work. Given that (as we will examine later) a sense of pride in one's work is what motivates most SREs, this demands digging into at an organizational level.

Which parts of recent incidents were most difficult?

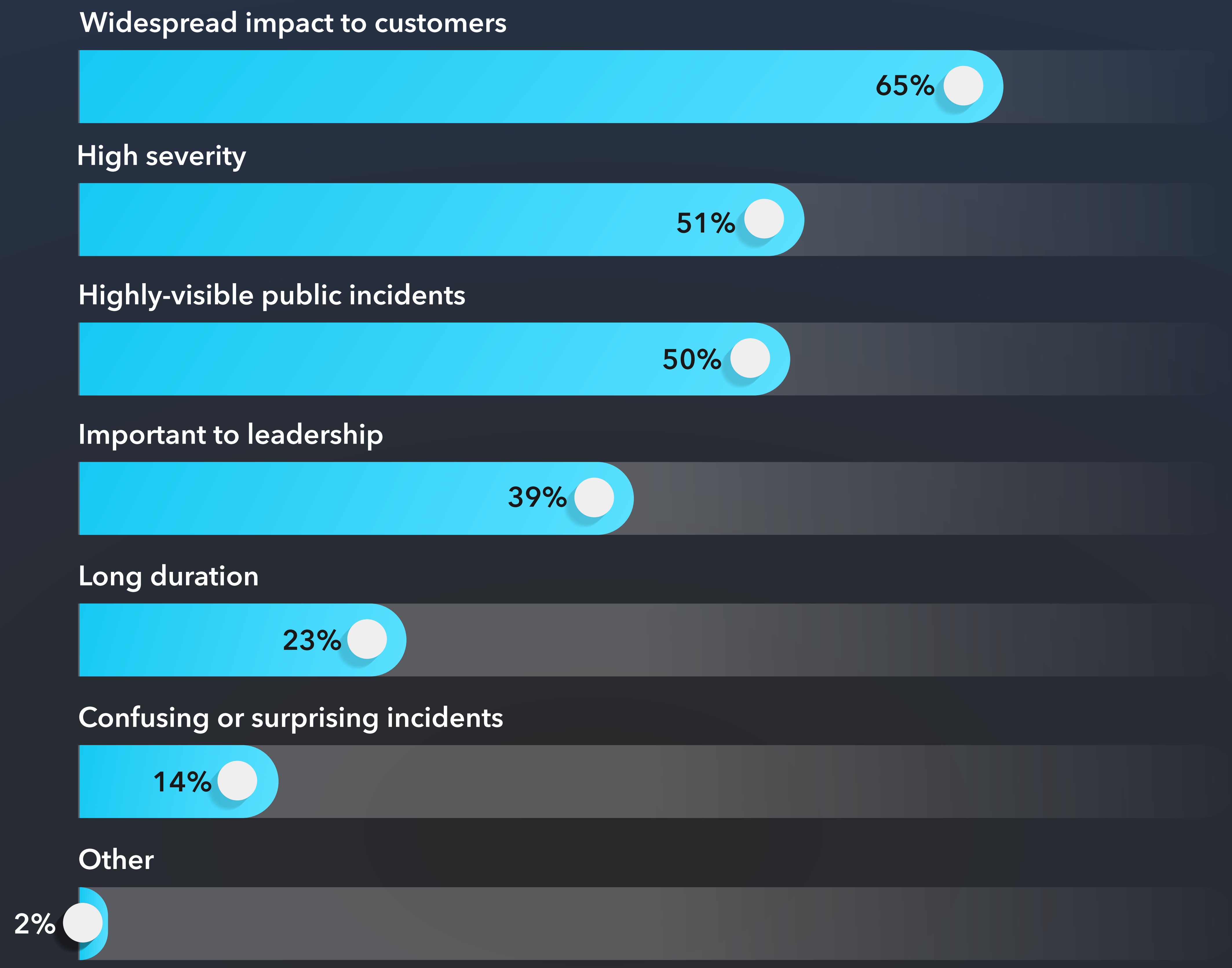


While reviewing the answers to this question, it's important to remember that the parts involved in managing incidents are far from linear.

However, we did try to provide a construct for the sequence that a typical incident management response would follow.

That said, by far the part deemed most difficult was accurate diagnosis of the problem, followed by escalating to, or coordinating between, responsible parties. Interestingly, fixing the actual problem fell fourth.

Which incidents get the most attention in your organization?

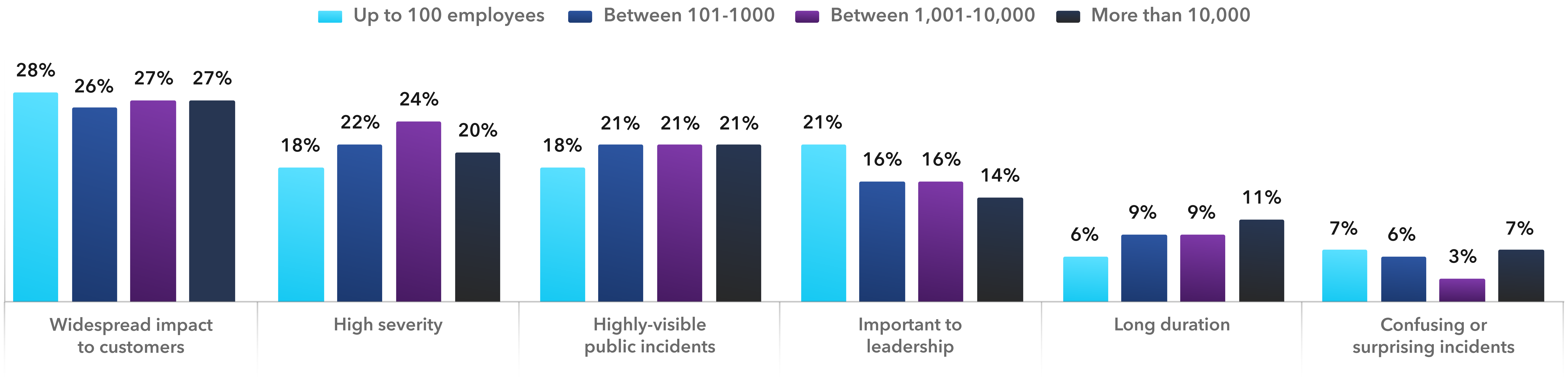


It's no surprise that the incidents which have the greatest impact on customers are those that get most examined. However, in our view, confusing or surprising incidents should be higher on the org priority list.

Why? The anomalies are often the incidents from which we have the most to learn because they represent gaps in the organization's knowledge.

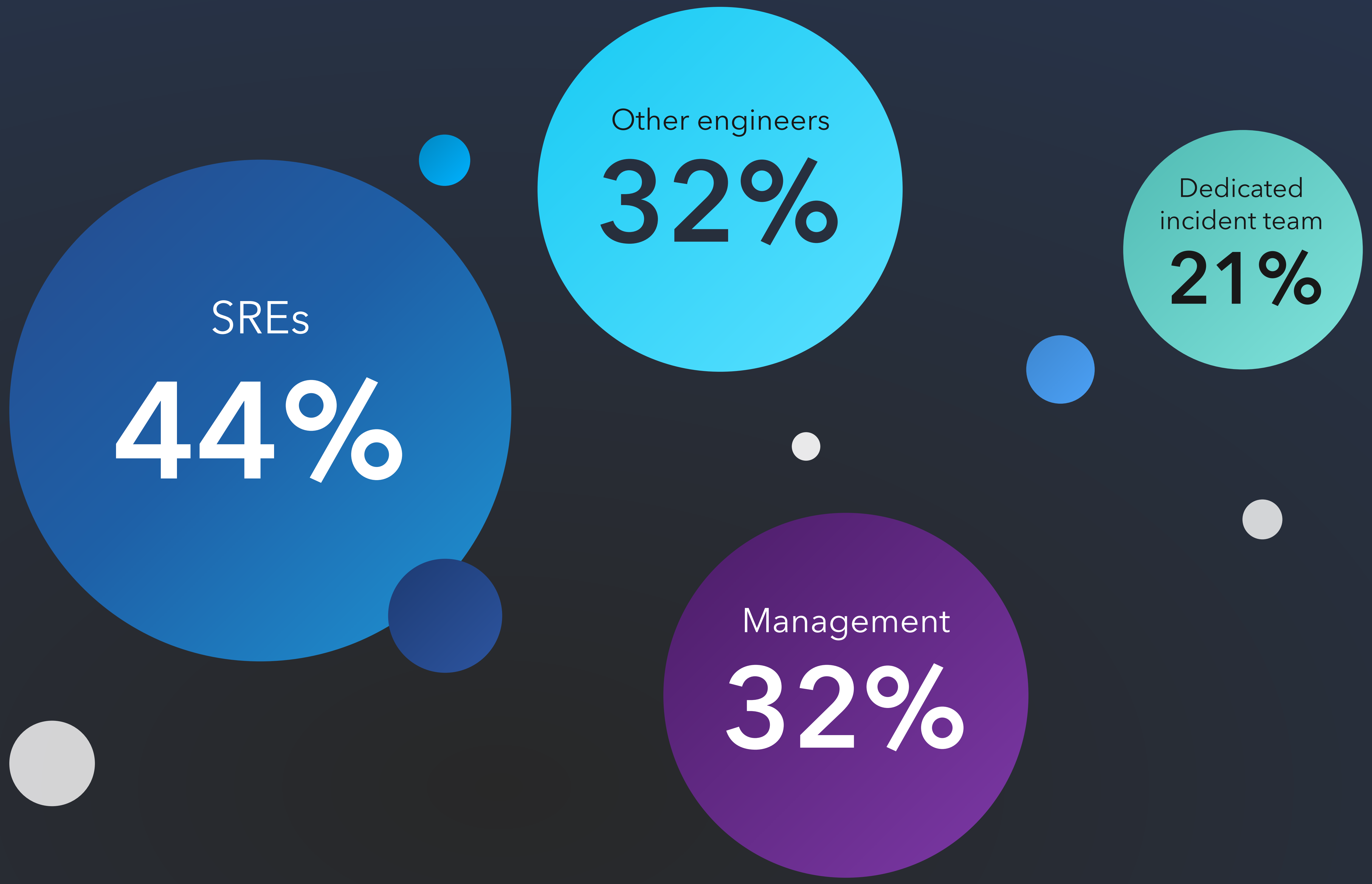
Resilient teams recognize that it is important to investigate the kinds of incidents that surprise SREs or keep responders up at night. They do this even if the incident didn't have a lot of customer impact or was not ranked high severity. This is because while they may not have been high impact this time, they may be early warning for the next major outage.

Which incidents get the most attention in your organization (by company size)?



As with other questions, we correlated replies for this one with company size. Interestingly, smaller companies said incidents that are important to leadership get a decent share of attention. By contrast, this trends downward as organizations get larger. Does this mean comms between SREs and executive leadership is breaking down in companies of a greater size? Or is it due to leadership being further removed from the incident level and therefore paying less attention? This could either be because they don't have to worry about incidents as a result of effective reliability work, or perhaps that they are too far abstracted from the day-to-day.

Who leads or drives the post-incident work in your organization?

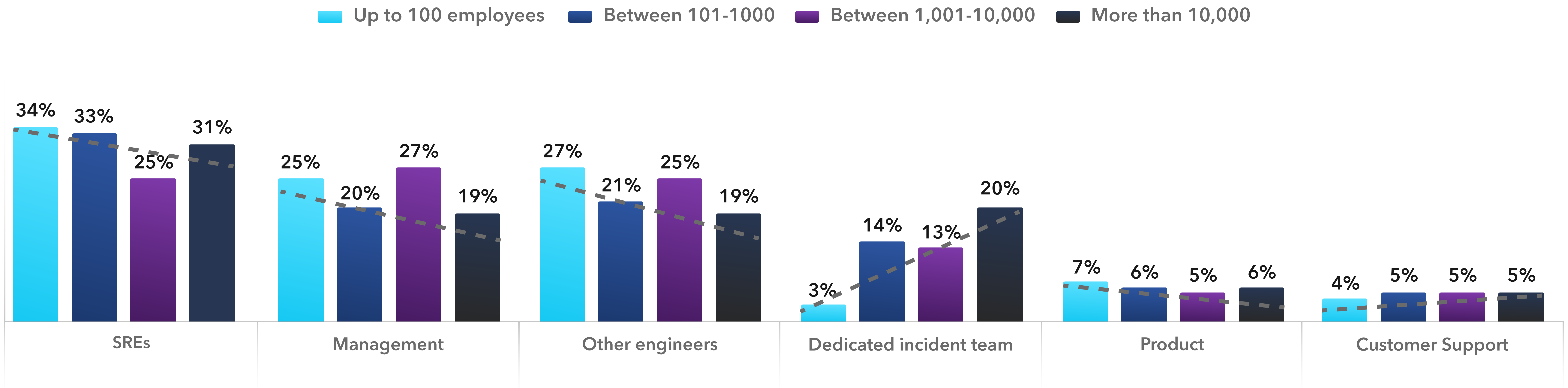


Why is knowing who leads post-incident work – and how widespread it is within SRE work – important? Largely for resourcing reasons.

The fact that SREs lead post-incident initiatives across company sizes demonstrates their need for protected time to participate in post-incident work. That other groups also take the lead in post-incident work from time to time, meanwhile, shows the value in interdisciplinary engagement. The leader(s) of the post-incident work can influence the nature and focus of the learning.

SREs	44%
Management	32%
Other engineers	32%
Dedicated incident team	21%
Product	9%
Customer Support	7%
Other	6%
Business analysts	3%
Marketing	2%
Legal	1%

Who leads or drives the post-incident work in your organization (by company size)?

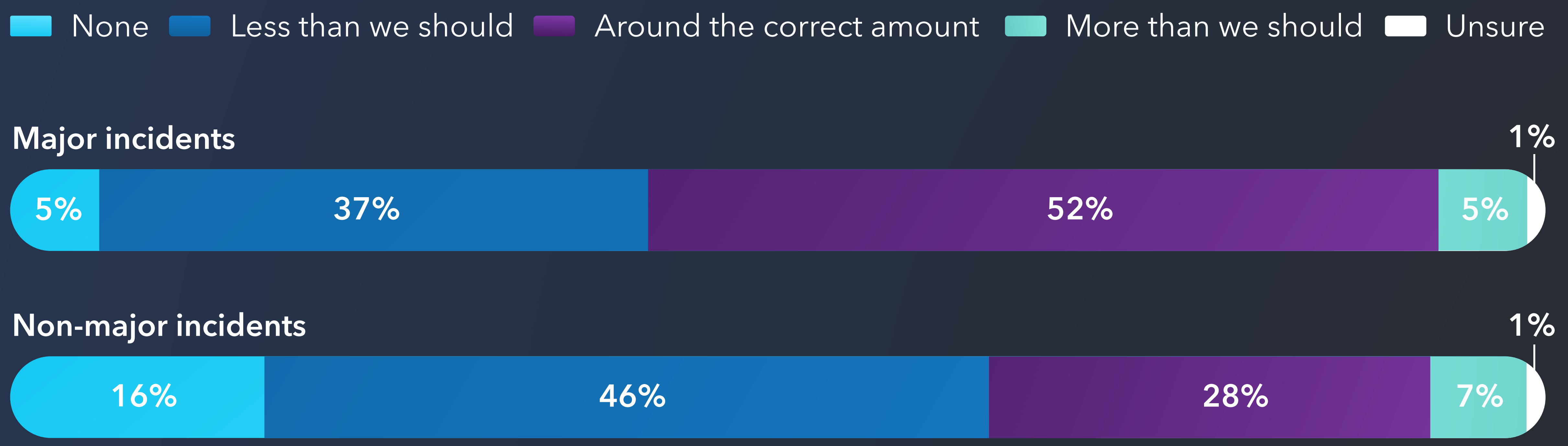


When we correlate company size here, we can see that the use of dedicated incident teams trends upward as companies grow in size. More resources likely mean more opportunity to dedicate a particular unit to incident response work. Only 3% of small organizations have this kind of dedicated initiative in place compared to one third of companies with more than 1,000 employees.

Adopting a "center of excellence" or even a "community of practice" approach can have radical benefits such as: greater clarity of objective, higher effectiveness, clearer focus, and in-depth knowledge of the issues at hand.

Further, in today's fast-paced environment, sharing best practice information gleaned from post-incident work as widely, clearly, and succinctly as possible (something a dedicated incident team will have the time and remit to do) will benefit the entire org.

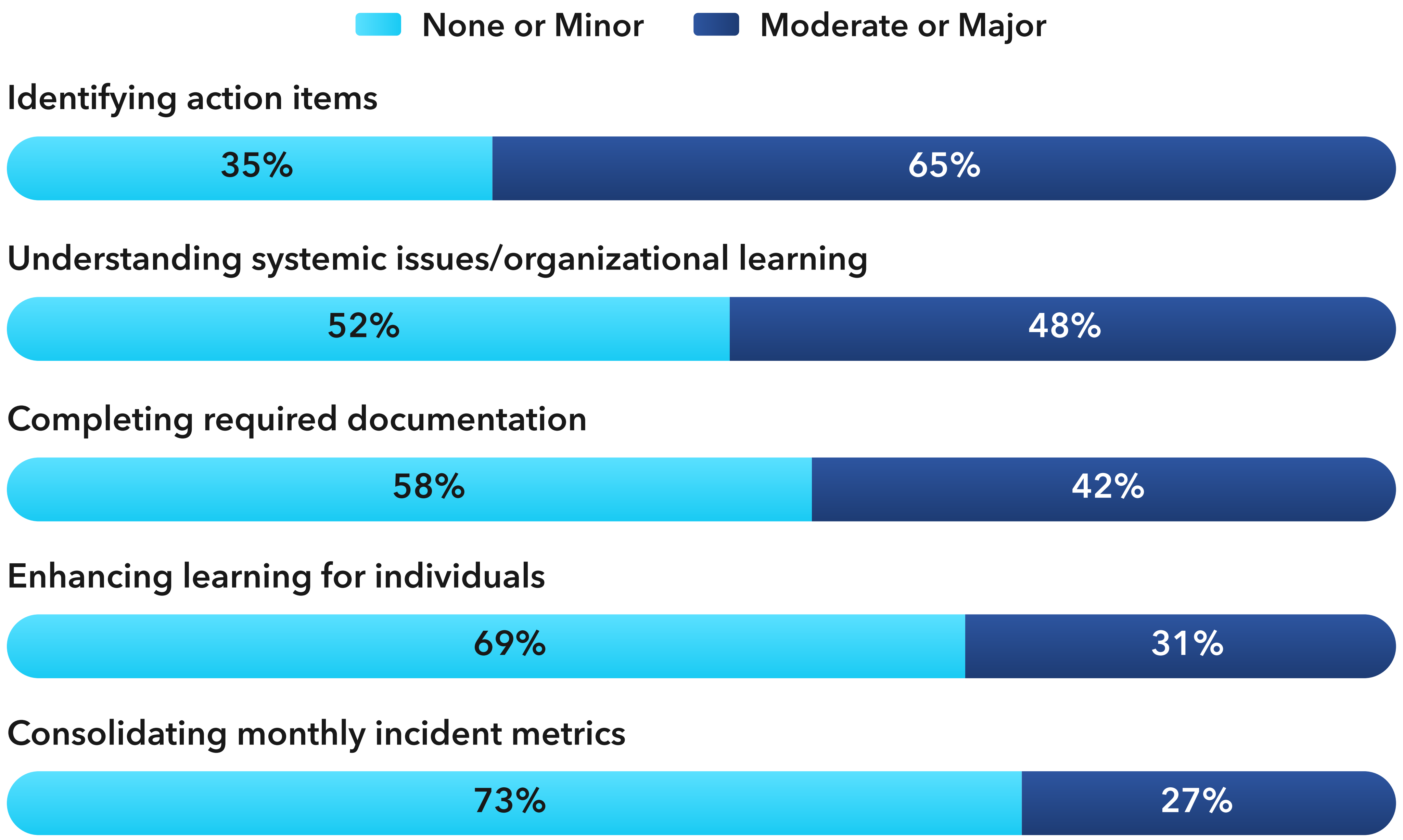
How much time does your organization spend learning from incidents?



We can see the time an organization spends learning from major incidents is higher than for non-major ones ('Around the correct amount' and 'More than we should'). It's understandable that the big incidents get priority in terms of tackling them in the present moment, but we see the general under-investment in learning from non-major incidents as a missed opportunity.

Time spent learning from non-major incidents should effectively prepare teams for when they must deal with the major ones. Incidents involving less stress that take place outside the glare of executive or customer attention can be far more conducive to learning (remember, learning takes place during the incident as well as before and after it). That's why we believe the high values for 'None' and 'Less than we should' for non-major incidents represents a tremendous opportunity for reliability practitioners.

How much of you or your team's post-incident work typically focuses on these?

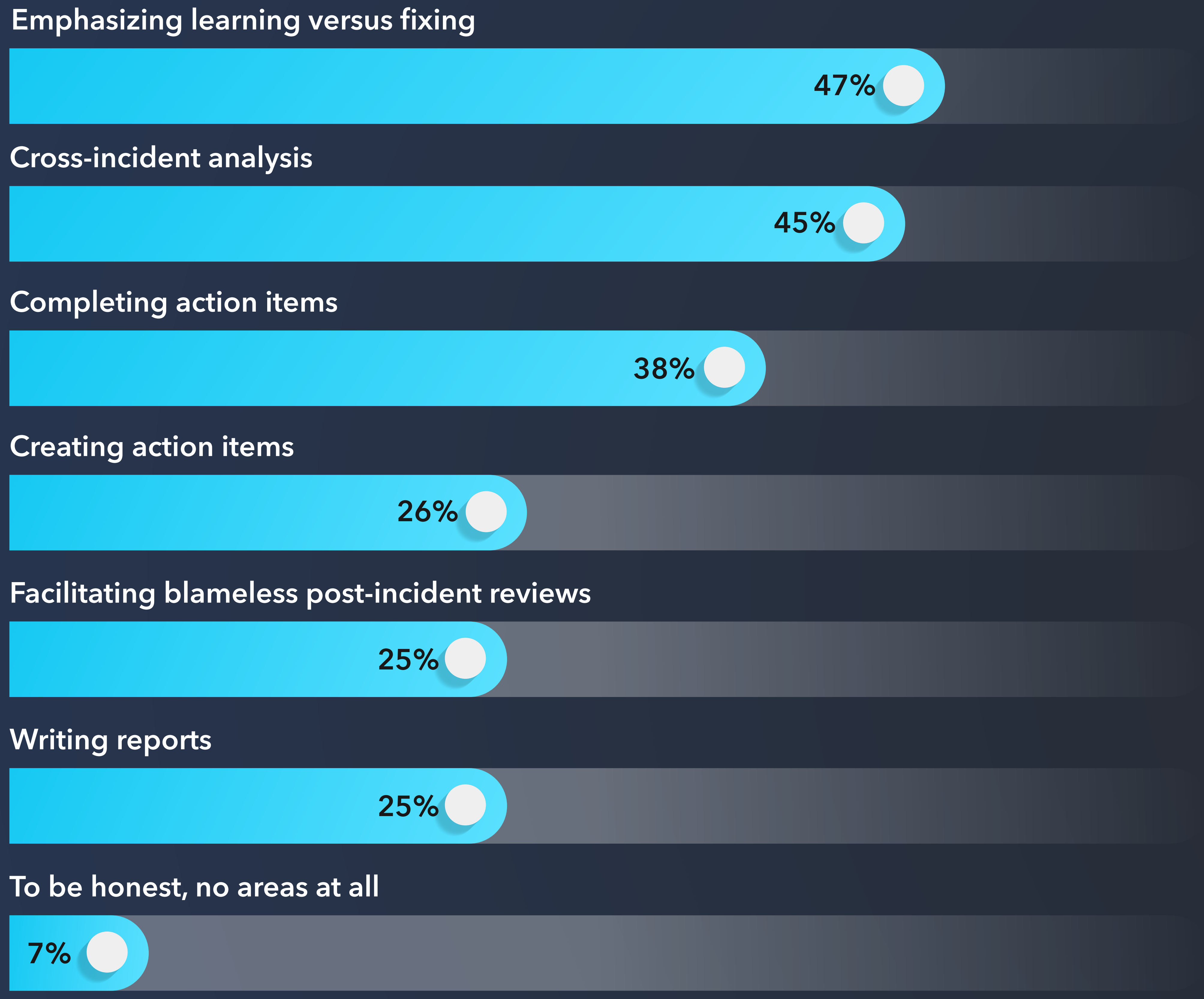


When it comes to the focus of post-incident work, we find it worrisome that ‘enhancing learning for individuals’ (almost 70% ‘none or minor’) came second to last while ‘identifying action items’ came in first (65% ‘moderate or major’). Why?

Studies have found that focusing on “what to do” tends to shut down the reflective work that relates to learning. Companies that focus on the org benefits over the engineers and individual responders during post-incident response may be inadvertently diminishing resilience if those org-focused activities do not meaningfully impact the needs at the individual engineer level.

The benefits of post-incident work that focuses on documentation, metrics, or action items, for instance, are typically for people farther 'from the keyboard', meaning they have broader organizational benefit but diminished resilience potential since they don't contribute to individual learning or identifying systemic barriers to performance.

In which incident areas could your organization most improve upon?

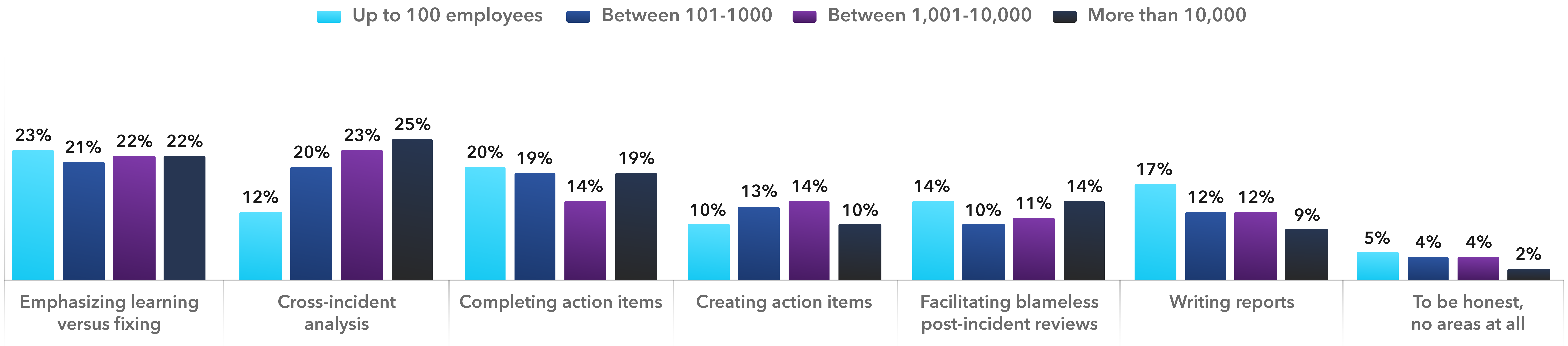


Reliability practitioners already appear to realize there is an issue with learning at an individual level from incidents.

The top answer in response to what organizations need to better prioritize was 'emphasizing learning versus fixing'. **How can teams change the mindset at an organizational level to enable greater learning for the individual (and long-term resilience for the company) from incidents?**

Oddly enough, completing action items fell higher on the list than creating them. We can create all the busy work we want ('creating action items') but if we don't do them ('completing action items'), any list is pointless. If we want to avoid a 'can't say nothing, so we have to say something scenario,' we need to make sure the action items have value, can be achieved, and are prioritized accordingly.

In which incident areas could your organization most improve upon (by company size)?



Looking back at the responses to the incident areas that organizations could most improve on, 'Emphasizing learning versus fixing' was the leading response from reliability practitioners regardless of company size. As you can see, there is no significant change as company size grows. **It's clear from these results that LFI is a universal business opportunity.**

Amongst the lower ranked items, it is interesting that larger organizations noted a larger gap for cross-incident analysis and a smaller gap for "writing reports" than smaller organizations. Perhaps this reflects the larger number of incidents with larger organizations and the greater tendency of having to complete "paperwork".

View from the field

Incidents are a part of everyday life in complex modern software environments so it's not a question of if an incident will happen, it's when. Being able to quickly and effectively learn from incidents is therefore an increasingly essential part of SRE work and it's a hot topic for good reason. Effective post-incident activities can lead to less surprises and more effective handling of future incidents by increasing responder's knowledge about how the system responds under different conditions - how it works and how it fails. These insights drive future resilient responses by ensuring the response team has the necessary expertise to flexibly handle difficult and unexpected incidents.

Dedicating time, tools, and effort to create the conditions for your teams to more effectively learn from incidents is a signal that your organization recognizes how stressful and technically demanding being on-call can be. It shows the company is committed to ensuring responders are well prepared and supported to handle challenging events.

Increasingly in today's ecosystem, managing incidents is about effectively working across organizational boundaries to secure needed system access and bring together both internal and external expertise. This is partly due to an increased reliance on third parties. This means that efficient incident resolution is not only about skillfully handling technical problems but engaging in effective coordination across all the necessary parties. Incident reviews that consider technical and coordinative aspects of incident handling help improve how teams work together under conditions of stress, uncertainty, and time pressure. They improve understanding of the sources of technical failure and expected system performance, increase understanding between team members as they will have increased knowledge of one another's skills and knowledge, and can lead to better coordination as team members can recruit the correct person into future incidents to work together more effectively.

As software systems increase in speed, scale, and complexity the need for SREs to have deep technical knowledge that can quickly and flexibly be applied to novel and unstructured problems will increase. Building your organization's capacity to extract meaningful learnings from its incidents is an investment in future reliability and higher retention of key personnel because it makes on-call a more manageable and less stressful aspect of Site Reliability Engineering work.

Dr. Laura Maguire

Principal, **Trace Cognitive Engineering**

Insight IV

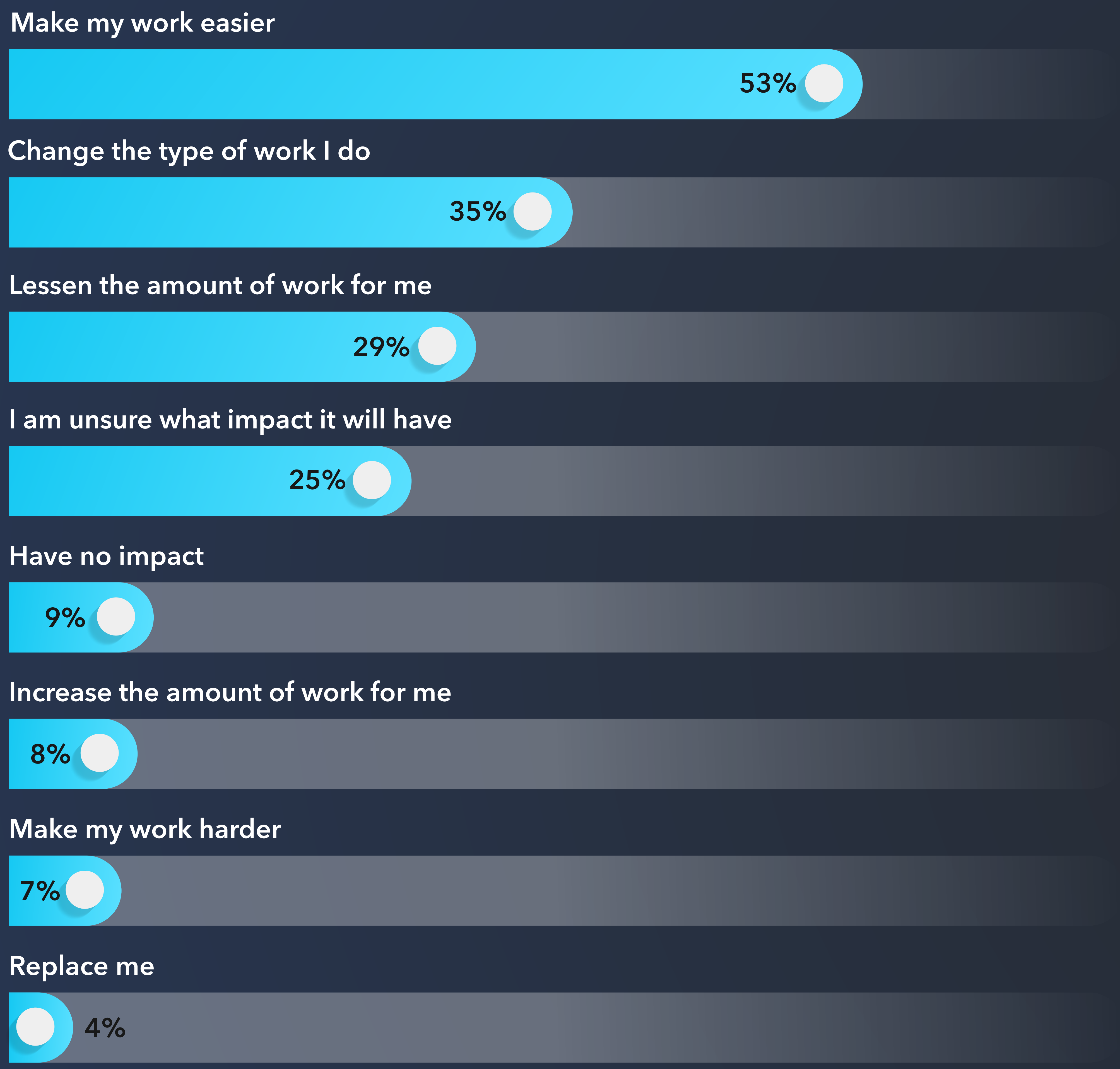
AI is not Replacing Human Intelligence Anytime Soon

4% of respondents felt AI will replace them.

53% of respondents said AI will 'make my work easier'.

The perception of AI's impact varies by organizational rank.

What personal impact will AI have on your work or role within the next two years?

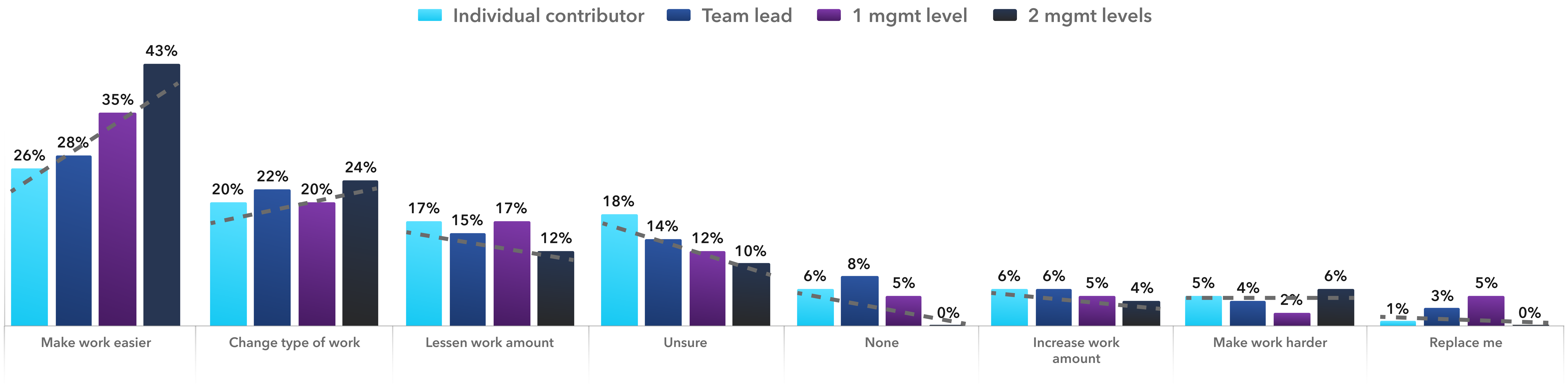


We've seen a deluge of coverage on AI since the last report. However, the community currently believes AI will not replace human intelligence within the next two years.

Completely negative sentiment was low: only 4% of respondents selected 'AI will replace me' and 15% said AI will either increase workload or make it harder. In other words, AI is now perceived as poised to help get shit done.

That said, it's also worth noting 25% answered they were unsure of its possible impact. There are a lot of continued unknowns, and unknown unknowns here.

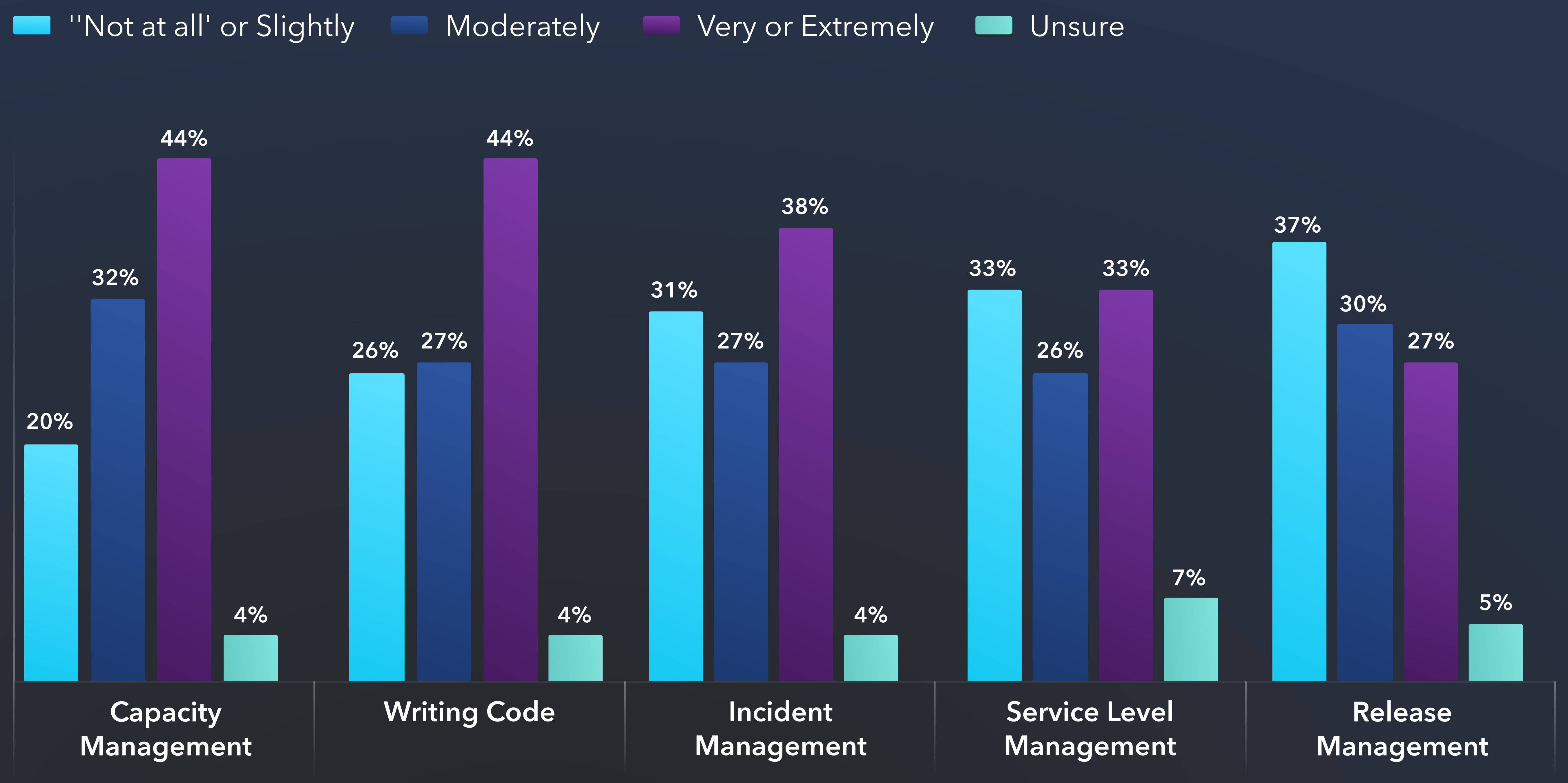
What personal impact will AI have on your work or role within the next two years (by rank)?



In another example of rank dissension, the perceived impact of AI within the next two years trends differently by organizational rank. For example, higher management degrees increasingly believe their work will be made easier thanks to AI (43%) and are less uncertain about its impact (10%) while individual contributors were less convinced (26% & 18% respectively).

There will be no shortage of decisions to be made when it comes to larger, business strategy approaches to AI. Factors like bias, the influence of social media and marketing, and the deluge of data (which may not translate to information with value) may also cloud or delay decisions. For existing business models with well understood parameters, incrementalism is probably easier to succeed with, particularly in the production domain. However, the reality is we are dealing with technologies we only partially understand, and there's a lot that's fundamentally unpredictable about where true sustainable value is going to come from.

How useful will Artificial Intelligence ("AI") be in the following activities within the next two years?



In the SRE survey used to generate The SRE Report 2023, we asked respondents to rate the value received from AIOps (the skew was toward no or low received value). For The SRE Report 2024, however, we decided to broaden the question to AI in general and add 'within the next two years' qualifier.

'Writing Code' was tied with 'Capacity Management' for 'Very or Extremely' useful at 44% each. Given the huge boom in generative AI, it's not surprising that 'Writing Code' is one of the most popular responses.

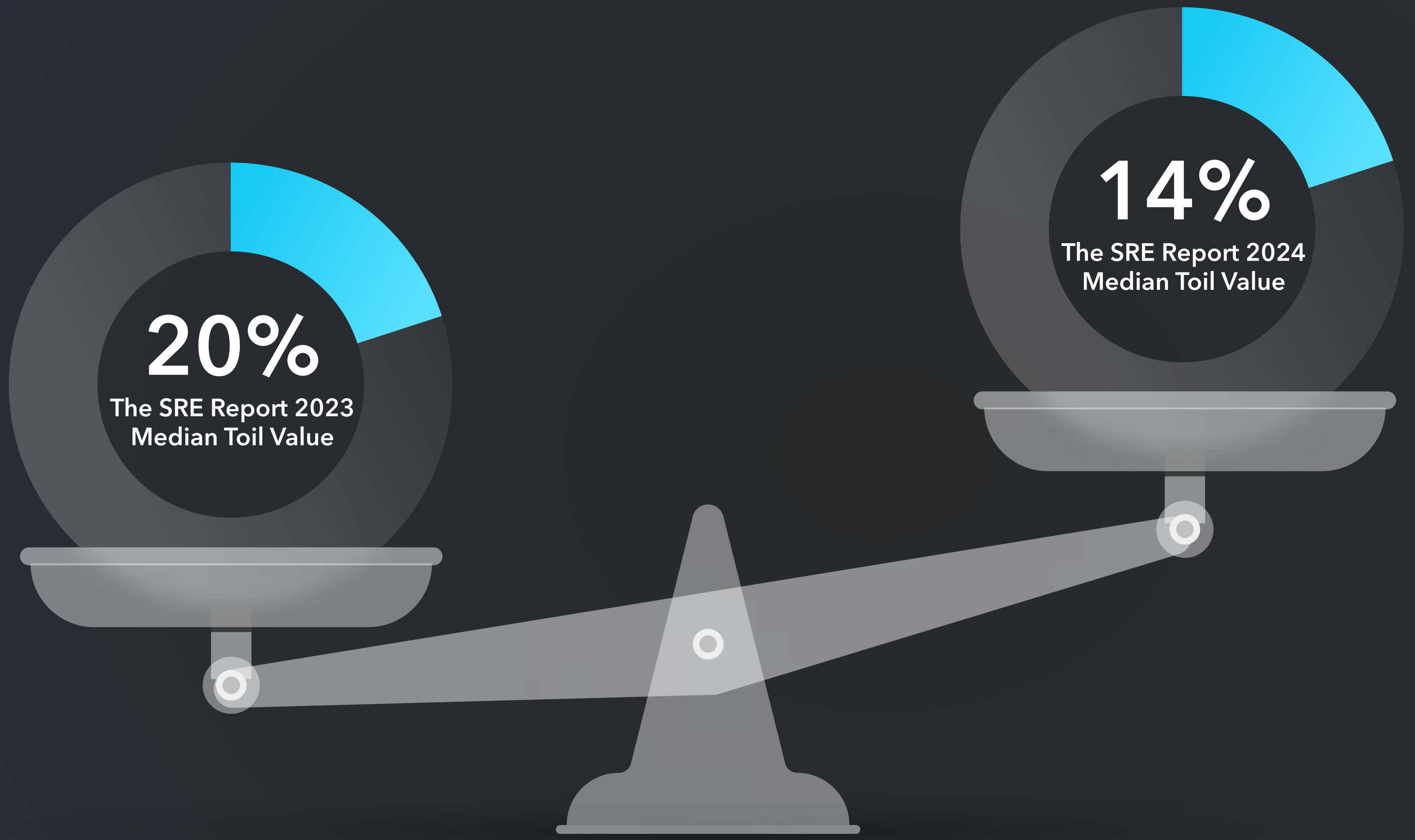
Merit was seen across all given categories (Incident Management at 38%, Service Level Management at 33%, and Release Management at 27%). The middle 'Moderately' dip for Incident Management and Service Level Management, though, may suggest polarized views for these use cases.

What percent of your work, on average, is toil?

Google defines toil as *the kind of work tied to running a production service that tends to be manual, repetitive, automatable, tactical, devoid of enduring value, and that scales linearly as a service grows*. We do not suggest there is a cause and effect relationship between the responses to the previous AI impact question and this toil question. However, we included it here because of the AI-to-automation relationship.

Editor's note: We do not think this median toil % change is because of ChatGPT or other GenAI assets. The survey data used to produce this report was done around eight months after the release of ChatGPT.

In the survey used to generate the insights for this report, we open-endedly asked, 'What top obstacle prevents automating more of your manual, repetitive workload?' In the spirit of this report section, we then asked a GenAI asset to classify the answers (see next page).



Bing chat response November 2023 to prompt:

Can you cluster these answers into categories? Can you give me the percentage of each category? Can you give me the answers for each category?

"I have clustered the answers into categories based on their content and format. Here are the categories, the percentage of each category, and the answers for each category:

Resource constraints: 38.3%. These are answers that mention the lack of time, money, staff, tools, or capacity to automate everything. They usually indicate that automation is not a priority or a feasible option for the current situation.

Some examples of resource constraint answers are:
Lack of time and prioritization from leadership, Budget, Lack of engineering capacity to devote to automating the toil away, Lack of time finding proper solutions, Backlog and capacity

Technical challenges: 34.0%. These are answers that mention the complexity, variability, dependency, or legacy of the systems, processes, or data that need to be automated. They usually indicate that automation requires a lot of engineering effort, expertise, or collaboration to overcome the technical difficulties.

Some examples of technical challenge answers are:
Complex and dynamic data sources, Dependency hell, Inconsistent code bases we have to provide workarounds for, Lack of functional API on dependencies, Legacy infrastructure

Human factors: 14.9%. These are answers that mention the fear, resistance, culture, or knowledge of the people involved in automation. They usually indicate that automation requires a change in mindset, behavior, or communication to gain trust, acceptance, or support.

Some examples of human factor answers are:
Fear of changing long running system in place, Lack of buy in from the team, Culture of fear, Lack of organizational communication and collaboration, Lack of knowledge in how to automate it

Other: 12.8%. These are answers that do not fit into any of the above categories, or are too vague, ambiguous, or irrelevant to be classified.

Some examples of other answers are:
Everything can't be automatable!, It varies too much, ITGC, Inventory use, Exception driven use cases

I hope this helps you understand the common themes and challenges of automation. 😊"

View from the field

It's hard to know whether this is another AI hype cycle or an intensification of the previous one, but it feels like there is something genuinely different between the (rather short on detail) promotion of AIOps, and what's happening with GenAI.

AIOps was (is?) in some sense narrowly focused on anomaly discovery and analysis, very much within a command-and-control framework. Business-as-usual, with go-faster stripes. But GenAI can do things entirely outside that framework, and as a result it doesn't seem quite appropriate to treat it as just another fixed scope tool. It's more like dealing with a very early-stage co-worker, who needs training and investment and constant review, but can occasionally be really valuable.

Conceptually, we'll start to use this to improve existing things we're familiar with in some particular way since incrementalism makes the sheer complexity easier to handle - but it is certain that larger approaches can and will emerge.

Niall Murphy

CEO, **Stanza Systems**

Insight V

When it Comes to Service Levels, Ignorance is Bliss

(in Smaller Companies)

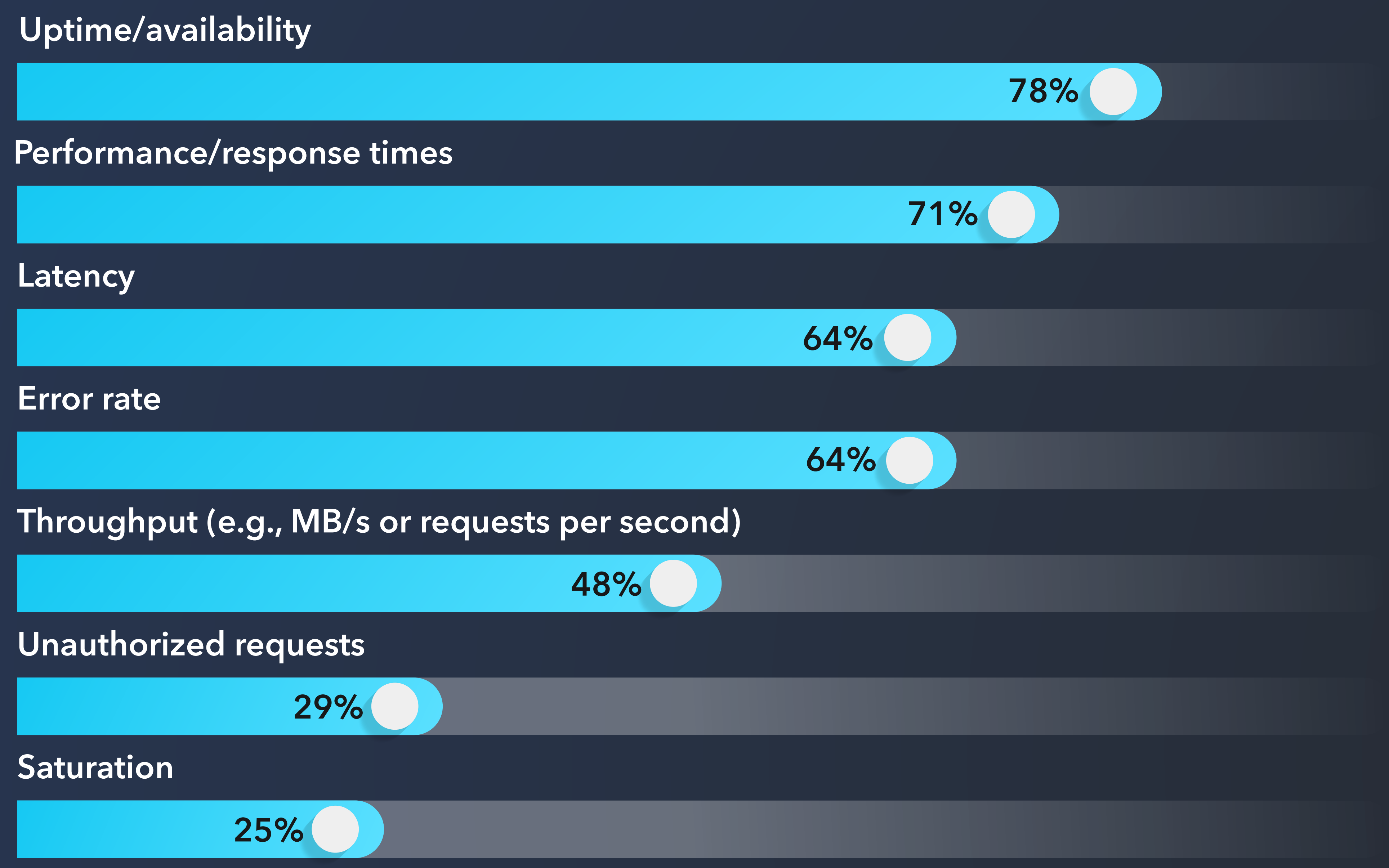
24%

of organizations have breached a contractual service level agreement in the last 12 months*.

The number of service level breaches trends upward with company size.

**From survey data June/July 2023*

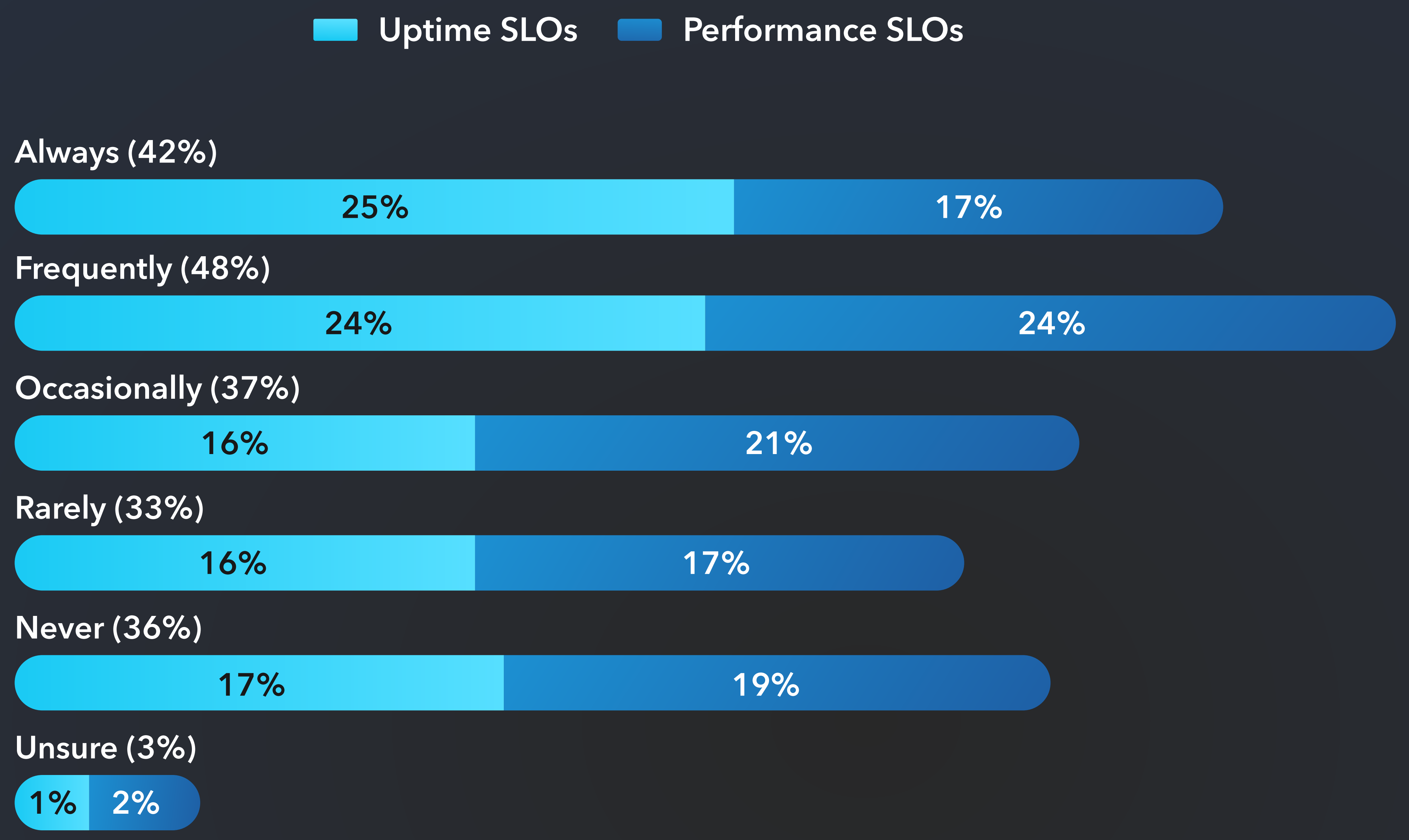
What indicators do you or your team monitor or measure today?



Service level indicators are a foundational component for being able to determine proper service level objectives. These objectives are also not to be decided within a silo. Instead, they need to be set with input from multiple stakeholders.

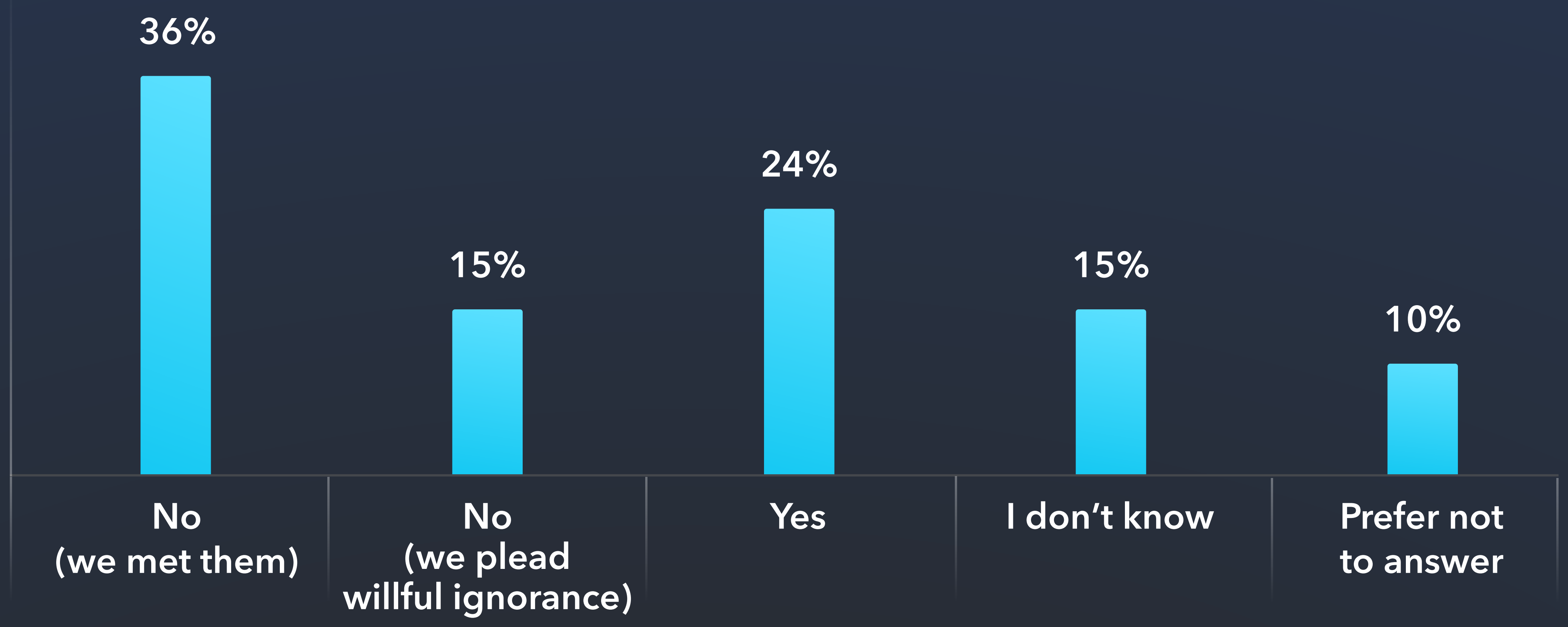
As such, we decided to ask what indicators are monitored or measured. The top three monitored or measured indicators have not changed since we asked this question in 2022. Uptime still leads as the indicator most widely monitored with performance/response times following closely behind.

Are your monitored or measured indicators tied to any service level objectives ("SLOs")?

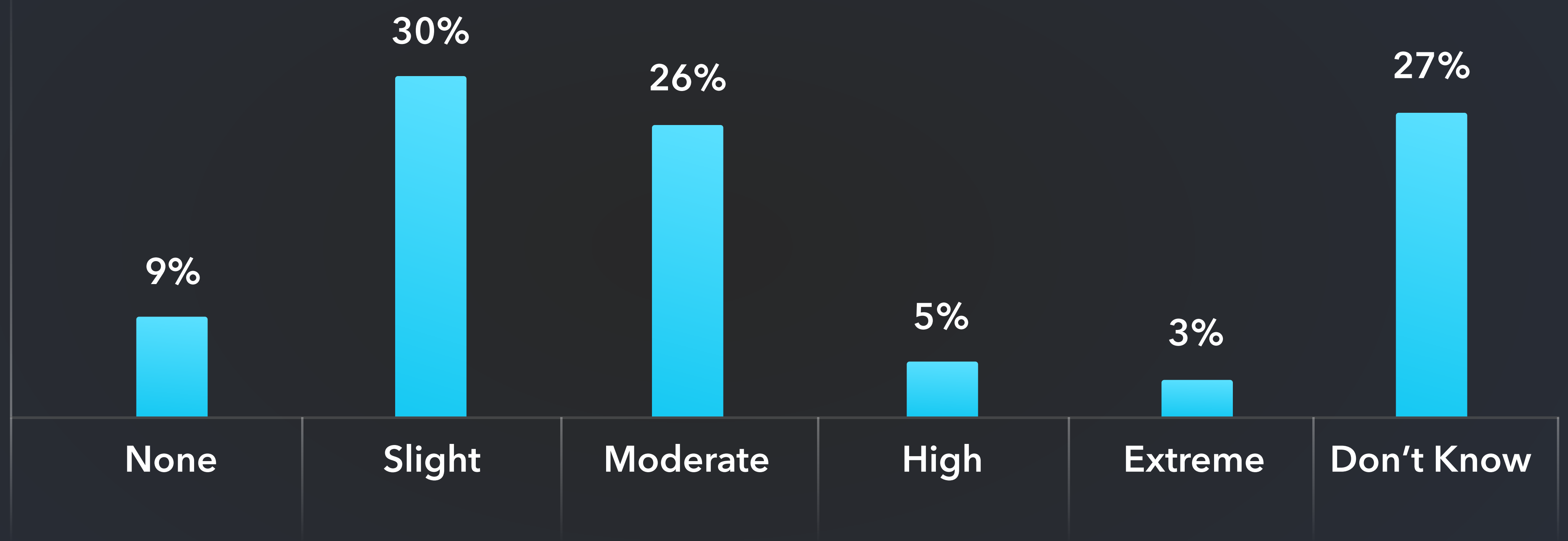


It's powerful to see that SLO usage tied to performance is so high (with usage of performance SLOs rated 'frequently' and 'always' head and shoulders above 'never' or 'rarely') – and is largely equivalent to uptime SLOs. If slow is the new down, then reliability practitioners are changing their reliability tactics accordingly.

Has your organization breached any contractual SLAs in the last 12 months?



What financial impact did the SLA breach cause to your organization?



We left quite a lot of wiggle room for respondents to state how their organizations were meeting contractual SLAs. Note that while the leading response was 'No (we met them)', 15% selected 'No (we plead willful ignorance)'.

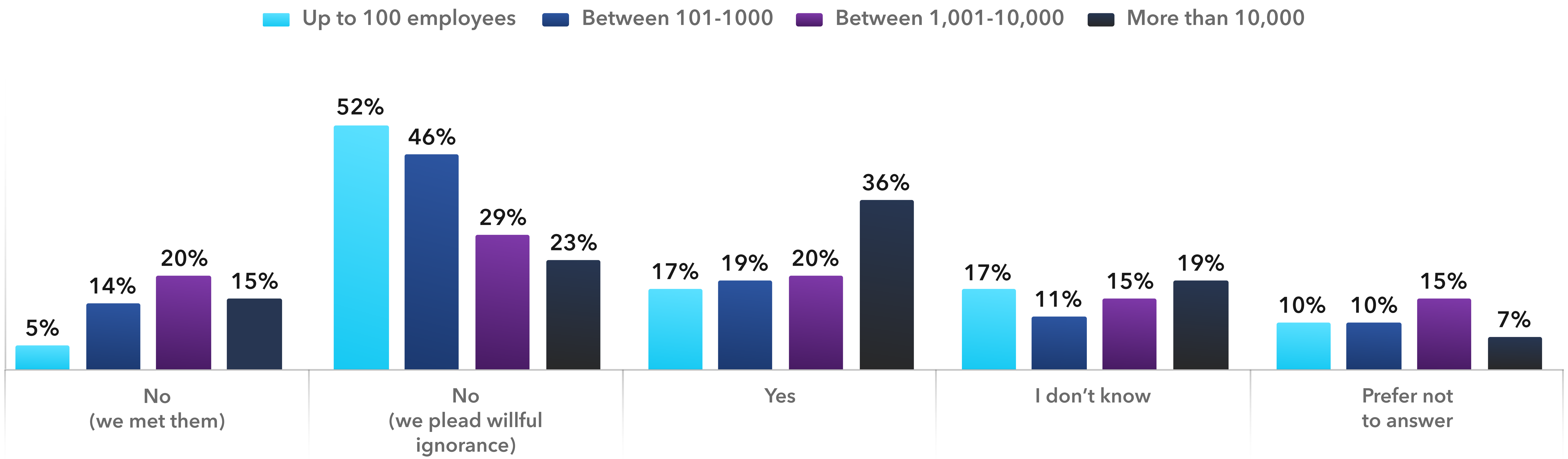
We were in fact surprised by the number of organizations who outright say they have breached contractual SLAs over the last year. We also wonder whether the 10% who said, 'Prefer not to answer' actually mean 'yes'?

Regardless, service levels indicators, objectives, and contractual agreements are crucial instruments for interacting with the business (hence why we included them in this section of the report.)

Regarding business interaction, the high percentage of respondents who did not know the financial impact resulting from SLA breach was alarmingly high. This lack of knowledge may be part of the willful ignorance, but it may also represent an SRE focus on objectives versus a business focus on contract language.

This was the only breach impact question and does not consider other dimensions like impact to brand, reputation, trust, or time spent reconciling.

Has your organization breached any contractual SLAs in the last 12 months (by company size)?



Let's analyze SLA breach with respect to company size. It's comical that as company size gets larger, the ability to plead willful ignorance gets (considerably) smaller. The up trend for 'Yes' as company size increases comes at the cost of ability to please willful ignorance.

Service level indicators, service level objectives, and service level agreements have a relationship and dependency with each other. While businesses will be more concerned with terms in an agreement, those same businesses will want reliability practitioners to invest in monitoring or measuring service level indicators against established objectives.

View from the field

Capabilities are the gateway between monitoring and measuring indicators, which lie at one end of the spectrum – to reliability and business outcomes, which lie at the other. The ability to manage service levels and answer questions like, "Do we have a healthy balance of velocity and reliability?" are such an example.

With around 25% of organizations breaching a contractual SLA in the last 12 months (at time of survey), a finer point around using service level indicators and service level objectives is important to make here. While many talk about core uptime and performance as indicators to monitor or measure, what we need to measure actually starts with the experience that your customers – including your workforce – should have.

We all face no shortage of challenges or constraints. Deciding which tools to invest in or having only 24 hours in a day are examples. It can be difficult to justify an investment today when the beneficial value may not be realized until a much later date – the classic leading versus lagging indicator conversation.

At the end of the day, service level indicators and objectives are critical components of reliability practices. They enable budgets for spending and allow us to focus on important goals. While service level agreements are filled with word salad, it's important to remember that they do work in two directions. In one direction, providers or vendors use them to set expectations for their customers. This ensures those providers and vendors know whether expectations are met. In the other direction, customers use them to hold vendors and providers accountable.

Alex Hidalgo

Principal Reliability Advocate at **Nobl9**
and Author of **Implementing Service Level Objectives**

Insight VI

No Single Monitoring Tool Does It All

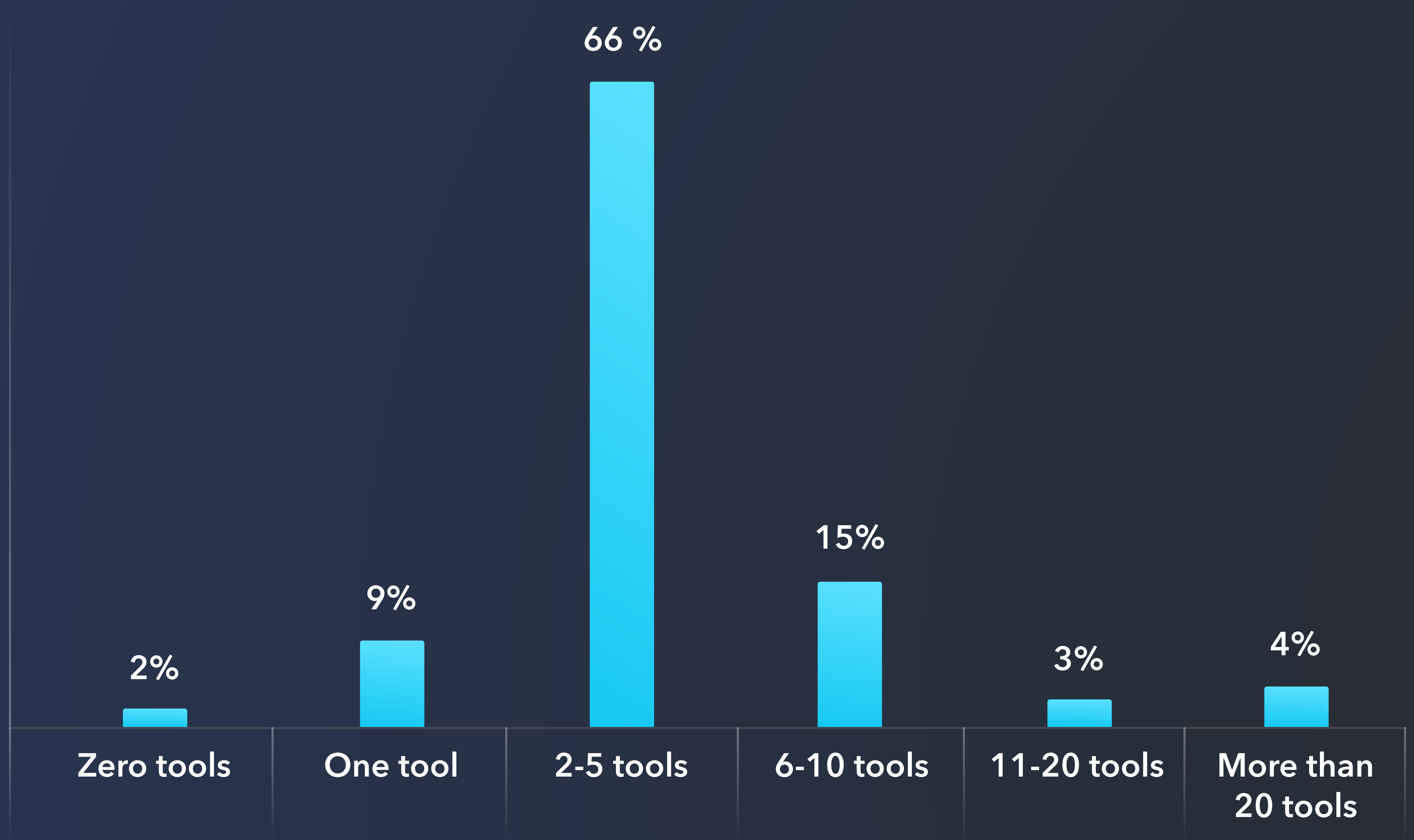
(and in the darkness bind them)

66%

of organizations use between 2-5 monitoring or observability tools.

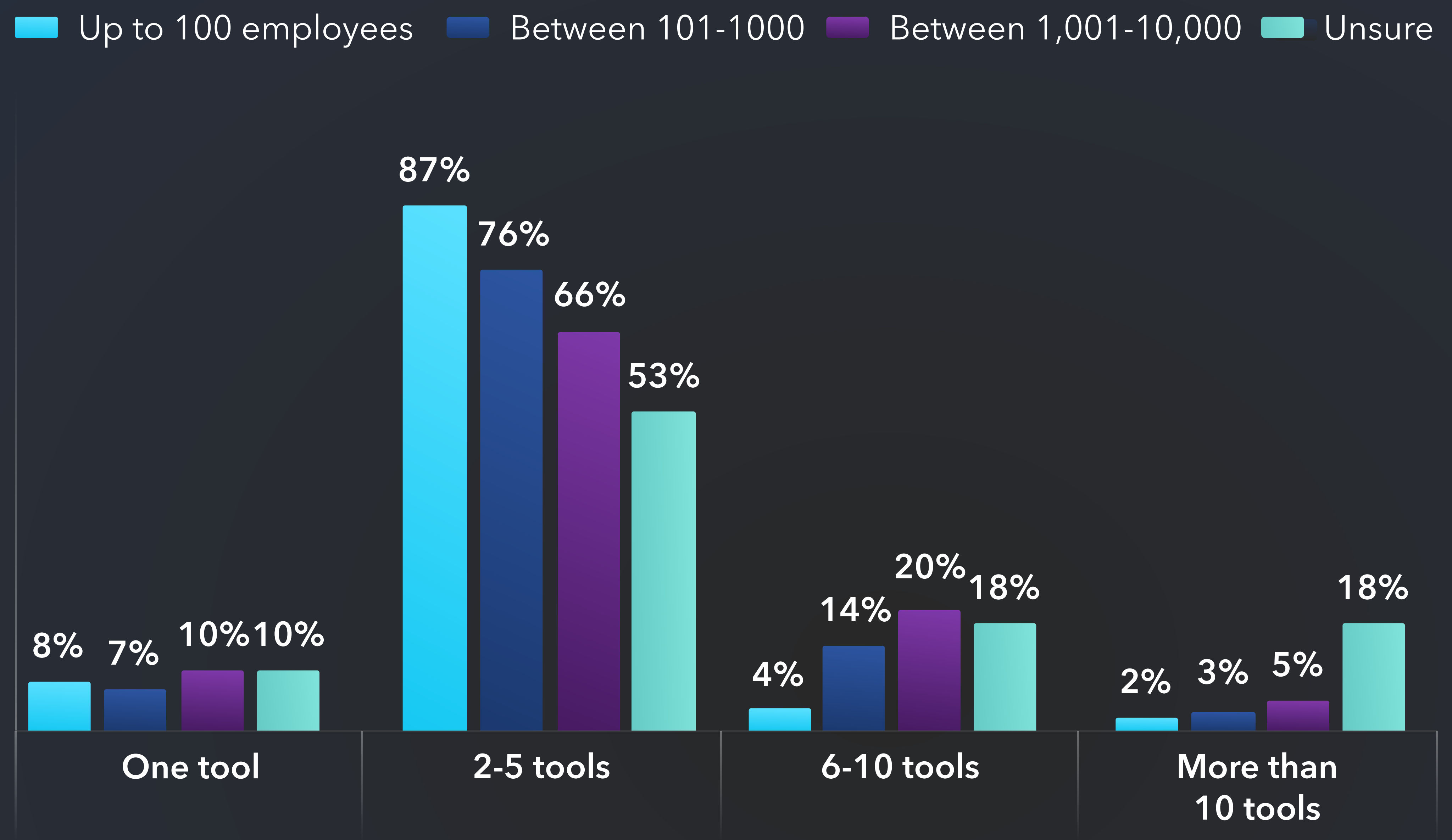
The number of monitoring tools in use trends upward as company size increases.

How many monitoring or observability tools does your organization use?



With so many organizations using multiple tools, it can be surmised there is value in such. In other words, don't let the conversation steer toward 'there are too many tools in the stack'. Instead, steer toward whether 'the value received from the tools in the stack is greater than the cost of those tools (where cost can take many forms)'.

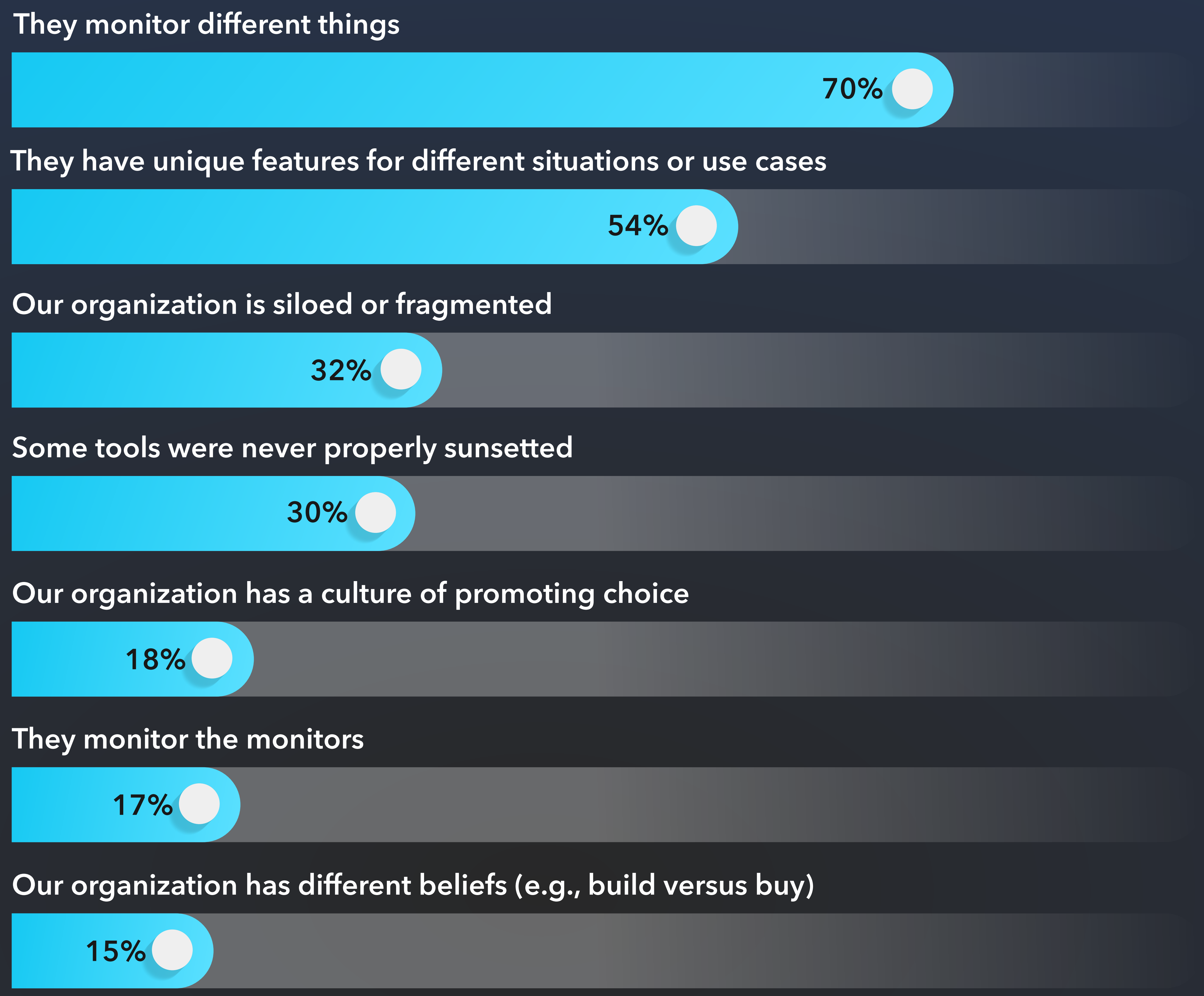
How many monitoring or observability tools does your organization use (by company size)*?



As the size of companies increases, 2-5 tools trends downward – 6 or more tools trends upward. Since the top reasons for using multiple monitoring tools were functional (see the next page), we suggest a 'diversity breeds resiliency' tactic as part of your conversations around defending tool value.

**Zero tools excluded from breakdown due to small sample size.*

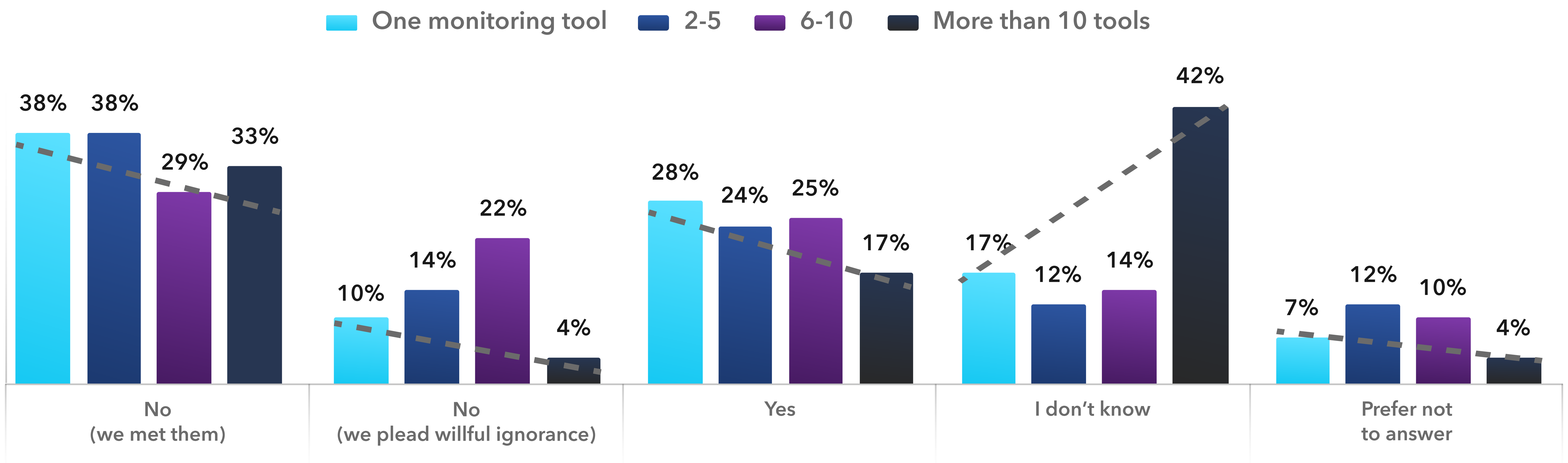
Why does your organization use multiple monitoring or observability tools?



We have seen other research ask the 'How many monitoring tools?' question, but we wanted to additionally ask why. The top two responses of 'They monitor different things' (70%) and 'They have unique features for different use cases' (54%) indicate that no one tool can monitor it all. Based on these findings, we suggest a leery eye for any vendor rhetoric around tool consolidation. Specifically, the conversation should not be around consolidating tools merely for the sake of consolidating them.

Further, when discussing value, ensure debt – technical or otherwise – is part of the conversation. While 'Some tools were never properly sunsetted' was only the fourth-ranked reason (30%) for why organizations use multiple monitoring tools, it indicates an accrual of debt. Said another way, legacy tools may not be adding value. Or worse, they may be subtracting from it in the form of inter-team communication problems, forgotten maintenance contracts, people-hours maintenance, and/or contributing to alert fatigue.

Has your organization breached any contractual SLAs in the last 12 months (by #monitoring tools)?

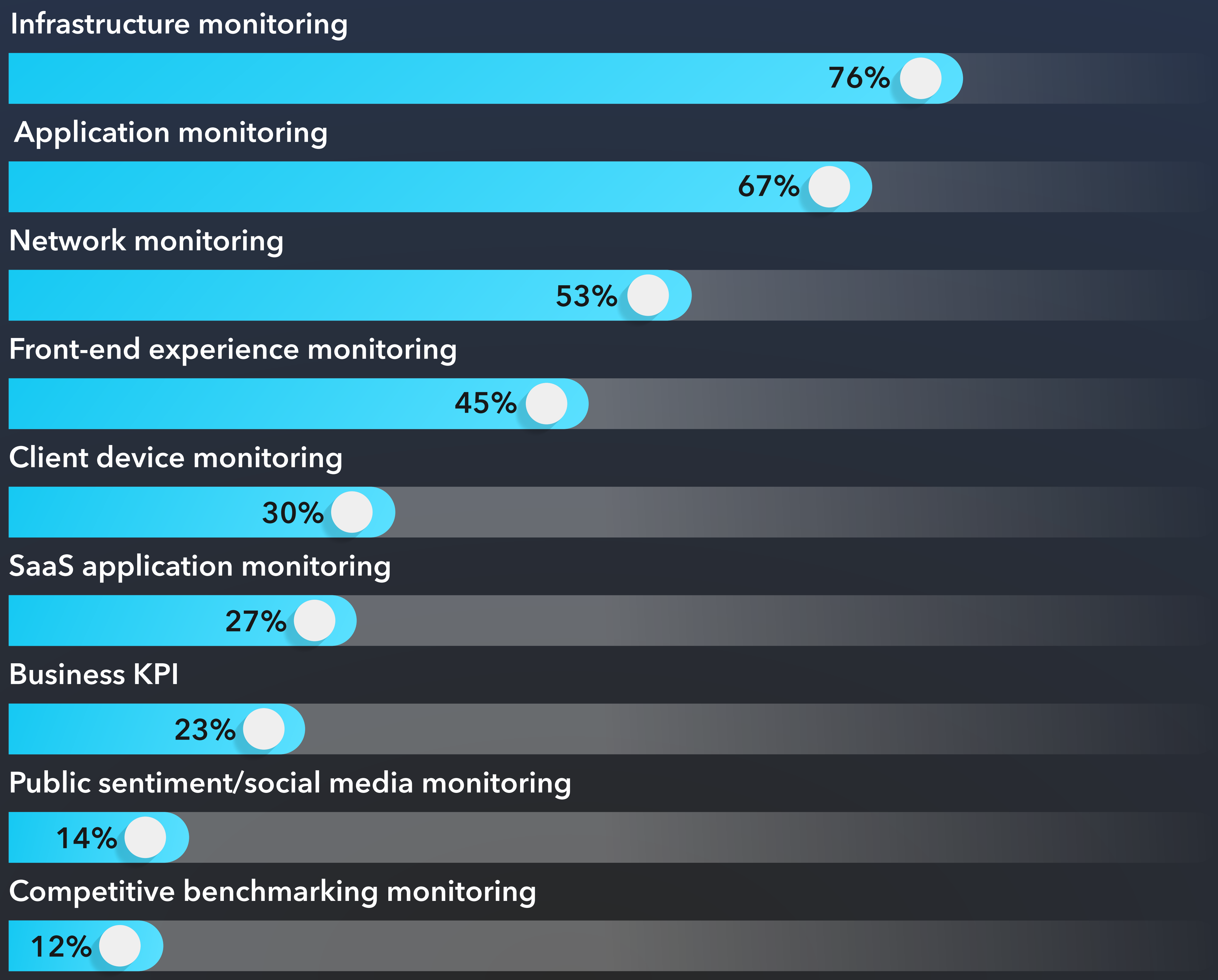


Reliability practitioners may find it difficult to convey business value when defending tool purchases (or resource investments when they wish to build their own). On the one hand, they know what their established uptime and performance objectives are. They also know, however, that no single tool can monitor everything, so the need for multiple tools will (probably) always exist. On the other hand, line of business owners will inevitably always be looking at bottom line dollar amounts.

We performed this investigation into the number of monitoring tools versus the number of SLA breaches to give an example of how a new or better business conversation could be had. In this set, **as the number of monitoring or observability tools trend upward, the 'Yes, we breached a contractual SLA' answer trended downward.** Based on this, it can be easier to place a hard dollar amount on what an SLA breach would cost, and then compare it to the cost of tools.

Keep in mind, this is just an example. Some organizations do not have formal service level agreements. In which case, the conversation will be about revenue-generating applications; nonetheless, conversations should align themselves around solving problems and value, first and foremost.

Which telemetry feeds your monitoring or observability framework(s)*?



In addition to the number of monitoring tools, we wanted to explore the **types.** From the previous question, infrastructure (76%) and application (67%) were the two most selected. These, along with network and front-end experience led the way last year, too. We also added two new categories: SaaS application monitoring and business KPIs. Each of which would likely benefit from a different monitoring tool.

*Additional survey context

Infrastructure
e.g., utilization metrics of server, database, hypervisors, or storage

Application
e.g., installing agents to discover components or perform tracing

Network
e.g., flow, SNMP, or packet capture

Front-end experience
e.g., synthetic transactions or probers, or placing a real user tag

Client device
e.g., the type of data you get from opening task manager or activity monitor

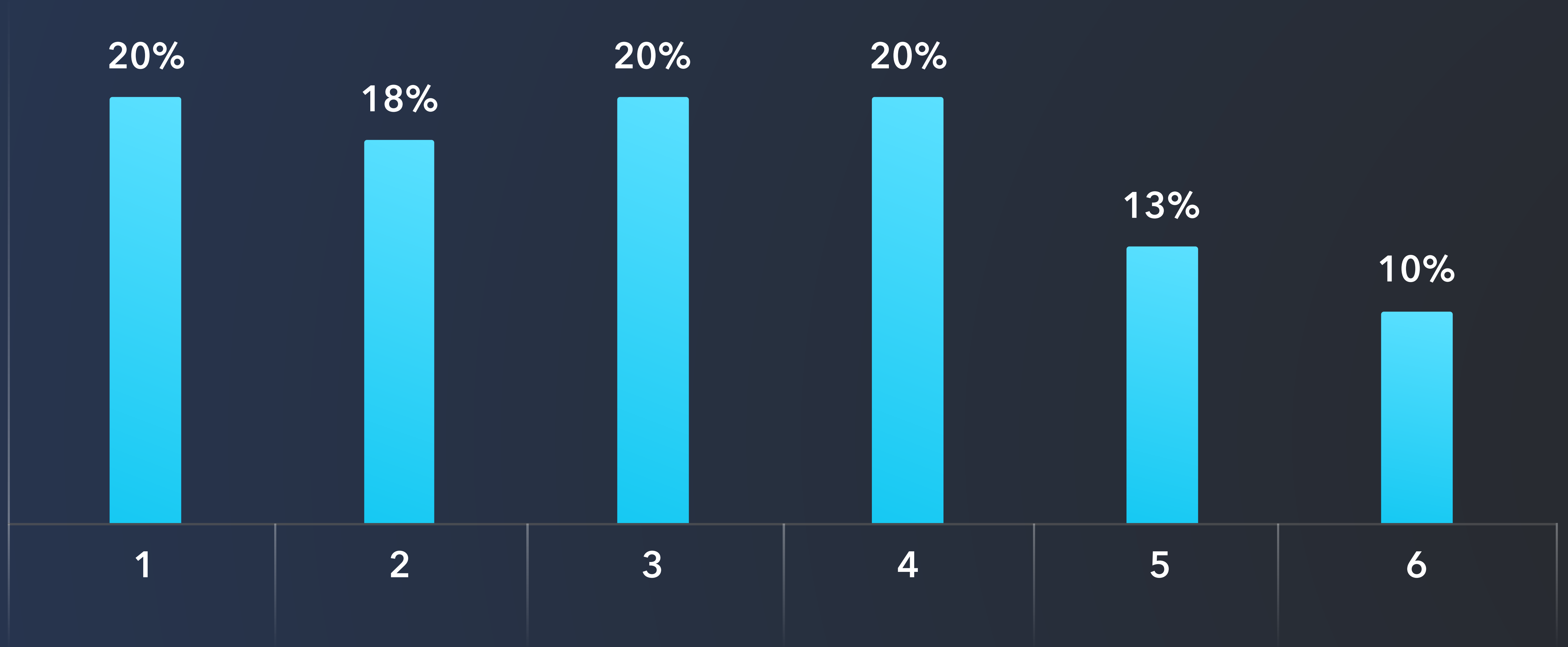
SaaS application monitoring
e.g., if you are not Microsoft/Google, do you still monitor M365/Workspace

Business KPI
e.g., site conversions or NPS scores

Public sentiment/social media
e.g., tweets or callouts on social platforms

Competitive benchmarking
e.g., comparing website performance of your page(s) versus competitor page(s)

[Derived] For organizations using multiple monitoring tools, how many different types of telemetry feed your monitoring or observability frameworks?



From the previous question of 'Which telemetry feeds your monitoring or observability framework(s)', we derived the frequency of how many ***types*** of telemetry feed those frameworks. For example, if a respondent selected 'Application monitoring' and 'Front-end experience monitoring', then that's represented by a tick in the two bin (in this visual).

As you can see, 81% of organizations have more than one telemetry type feeding their monitoring or observability frameworks! This visual reinforces 'they monitor different things' as the number one answer to 'Why does your organization use multiple monitoring or observability tools'.

View from the field

Clearly the value of monitoring and observability capabilities, and by extension data-driven decision making, is not in question. As the need for these capabilities grows to encompass the application stack, the infrastructure stack, the Internet stack, or their et cetera stack; there will be an increasing need to address the problem of needing to *'monitor different things'.*

Having multiple tools will inevitably lead to conversations of redundancy and ROI. When these conversations do occur, steer it toward understanding whether received value is greater than the cost. Or, at a minimum, understand how capabilities factor into the conversation. For example, if a [monitoring] tool or platform natively supports service level objective tracking, then one does not have to build and maintain it themselves.

Considering the earlier insights around control and third parties, it will be interesting to see how OpenTelemetry will shape this conversation – especially as it pertains to third-party support. If open standards such as this enable better ownership, stewardship, or federated collaboration, that may also affect received value. Either way, they will need to be included in your telemetry pipelines.

In reference to having multiple monitoring tools:

> **“Am I alone in this?”**

Not at all. As our landscape changes so does our process and practices for navigating it. If using multiple tools provides you with additional context, which in turn leads to faster identification and remediation of incidents, perhaps it’s time to embrace a multipronged approach.

Leon Johnson
SRE Manager, **Betterment**

Insight VII

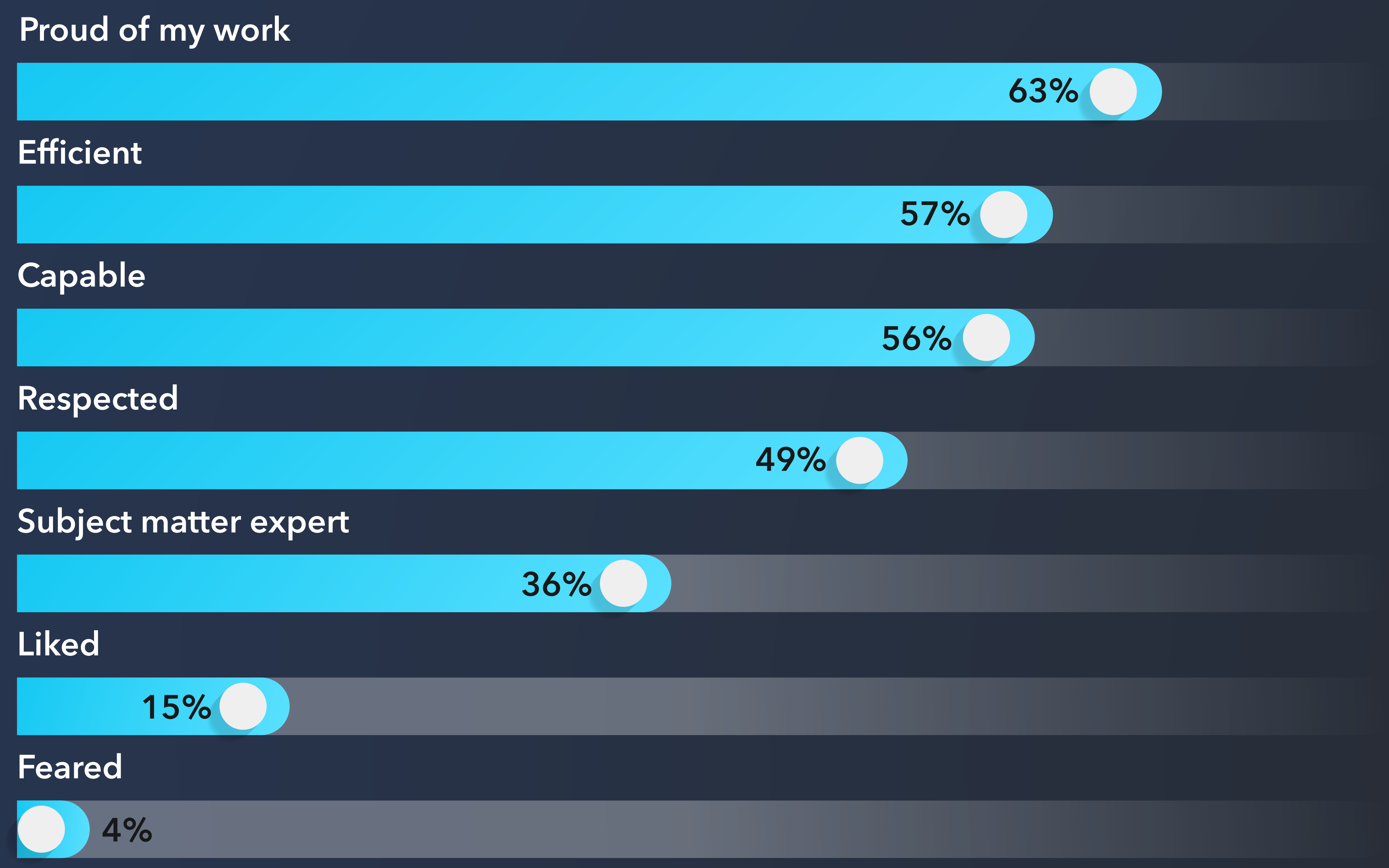
Efficiency is the Enemy of Pride

63%

of ***people*** said being proud of their work was most important to them.

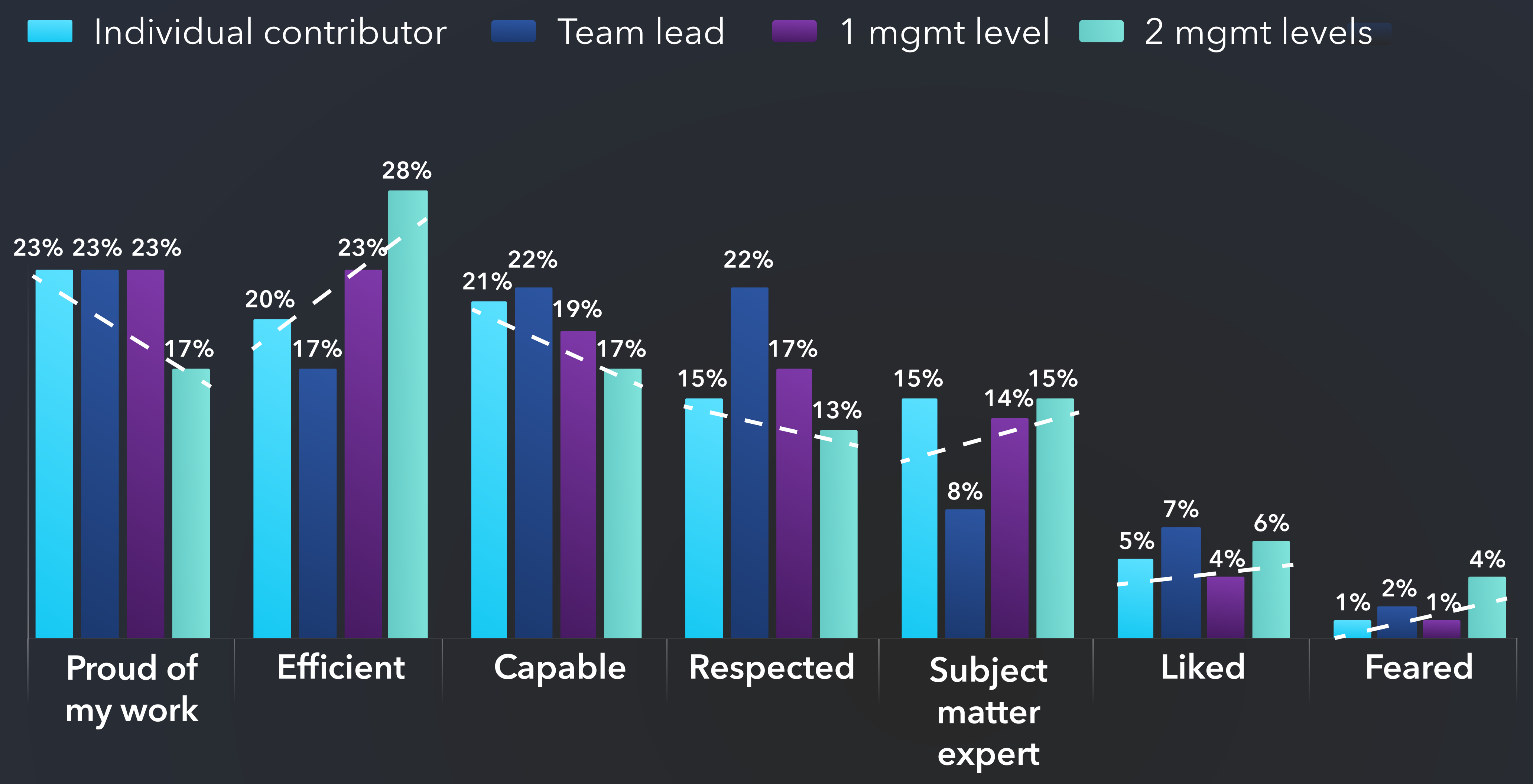
Unfortunately, this number trended downward as organizational rank of respondents increased.

Being which of these are most important to you?



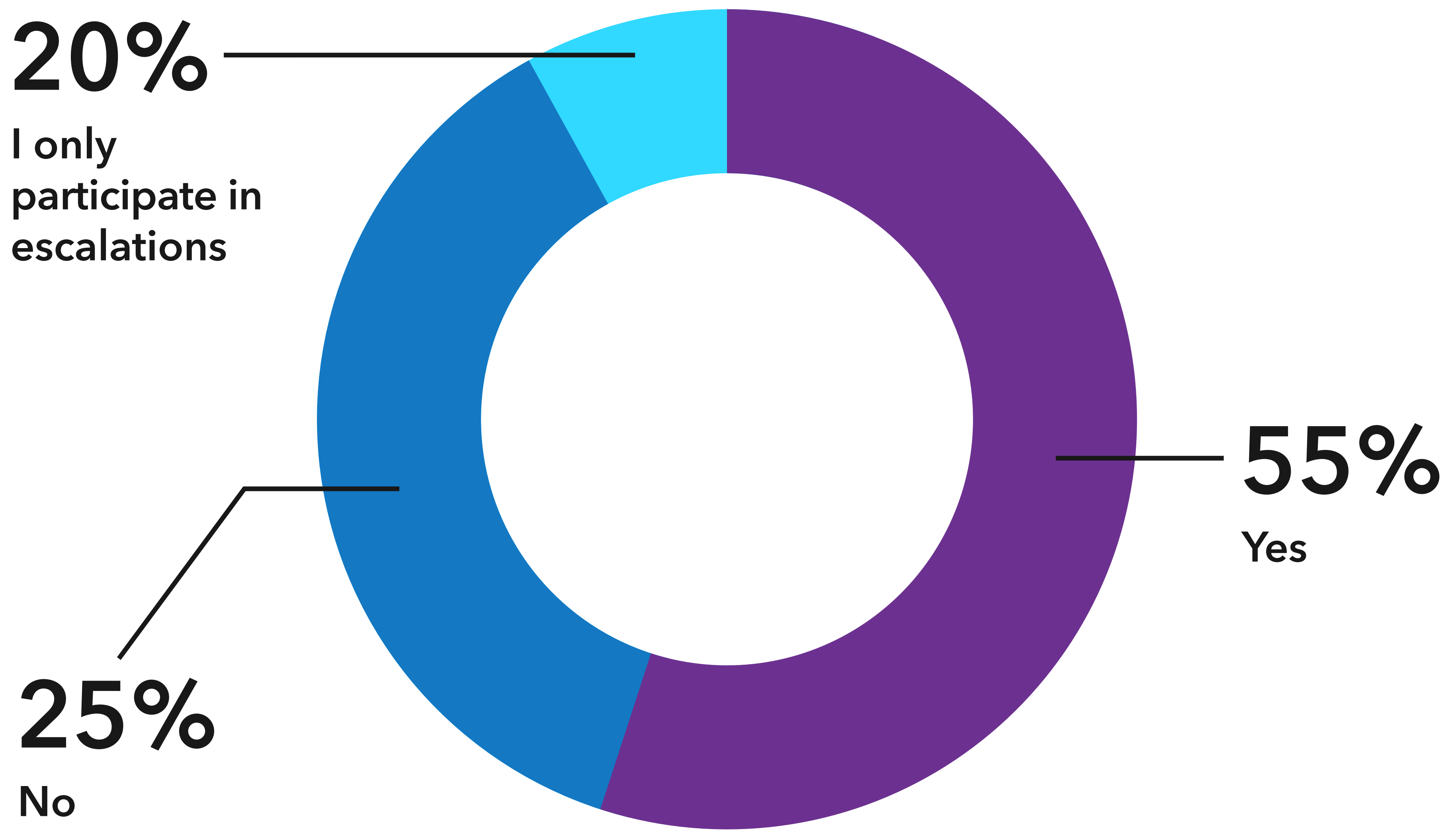
What's important to the business is often discussed in marketing research. However, we feel it is also important to discuss what's important at the personal level because a business cannot exist without people. While we are glad to see efficiency ranked as high as it did, we are also glad that 'Proud of my work' took the top spot. The overwhelming majority said being proud of their work - followed closely by being efficient – was most important to them.

Being which of these are most important to you (by rank)?



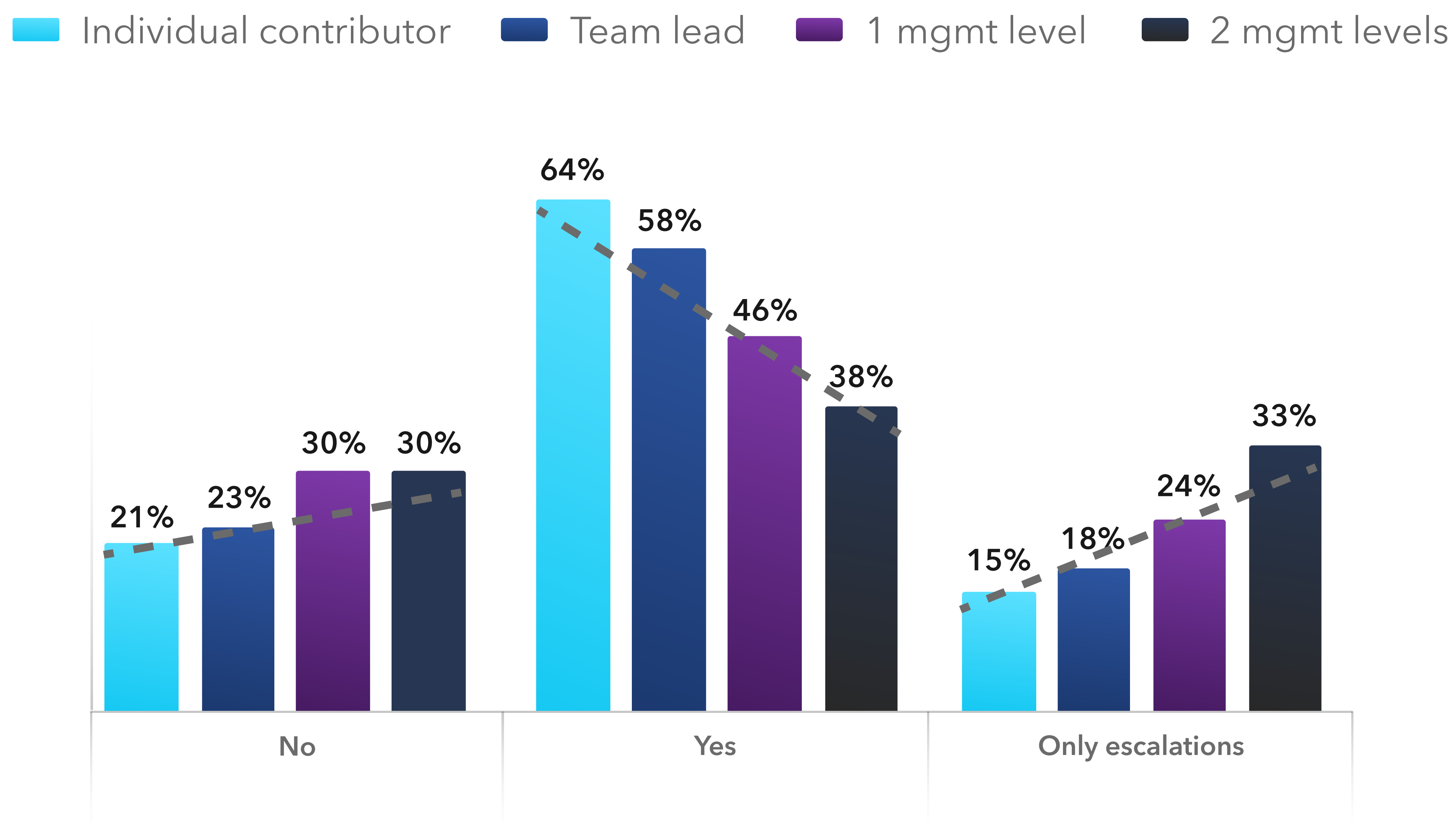
While the aggregate majority said being proud of their work was most important, this sentiment varied by organizational rank. As an indicator of higher rank suggests more pressure around overall business performance, they will want to take note not to diminish the personal importance of pride, ability, and respect in their day-to-day work.

Do you participate in an on-call rotation?



As per each previous year, here is our benchmarking data which reveals how SRE time is invested. As always, we recommend you compare it to Google's seminal recommendations on balancing operational versus engineering activities. As Google reminds us, the importance of keeping SRE ops in check is so that no SRE org or subteam inadvertently devolves into an Ops team. Knowing how you spend time and comparing it to others in the field can be invaluable in shifting individuals and teams to an 'investment of time' mindset.

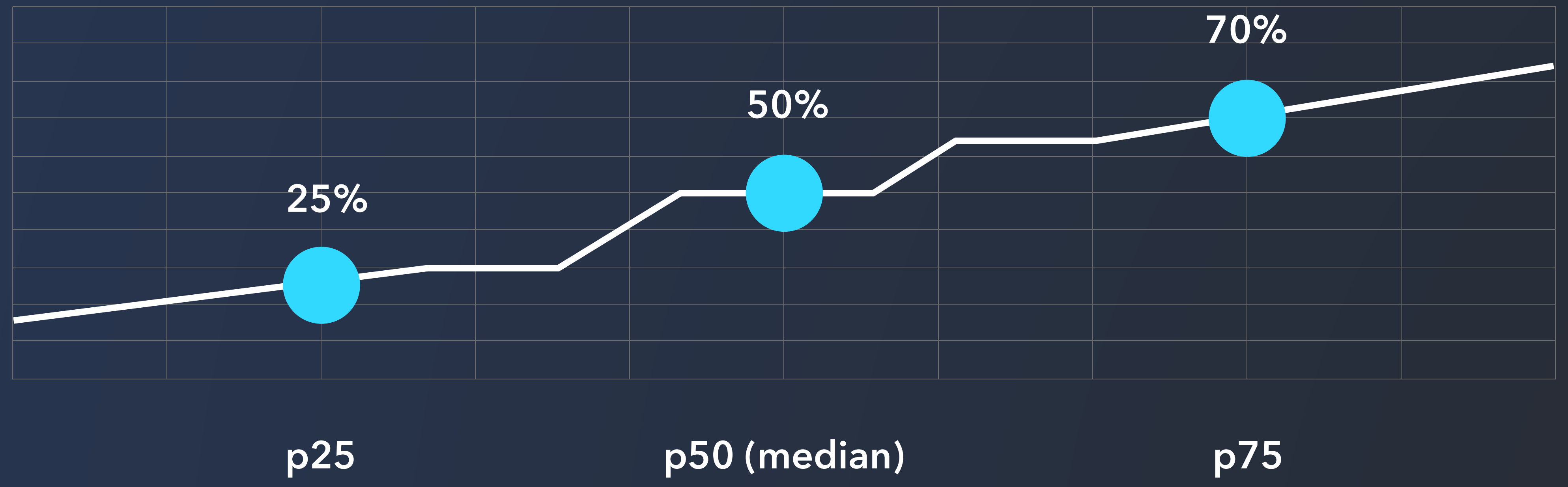
Do you participate in an on-call rotation (by rank)?



As organizational rank increases, participation in on-call (distinguished from 'I only participate in escalations') trends downward. This makes sense to us.

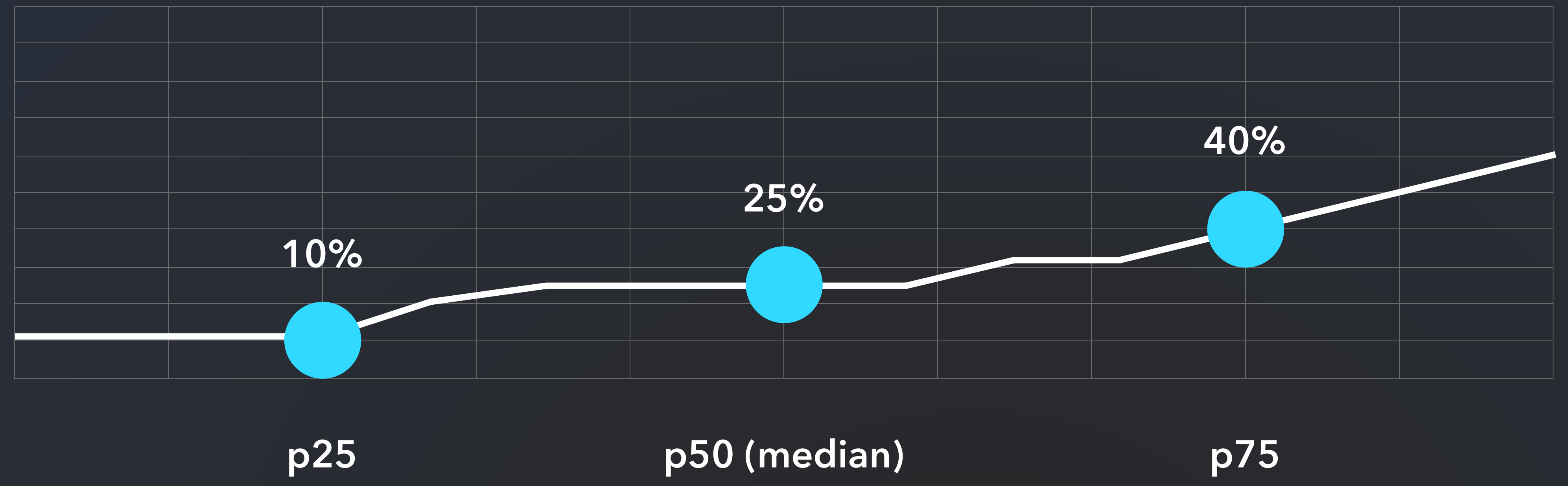
Tradeoffs of higher ranks might include technical skill atrophy but at the benefit of better business acumen. Regardless, remember that we don't spend our time, we invest it.

In a typical week, when you are not on call, around what percent of your time is spent on engineering activities?



	p50 (median)
Aggregate	50%
Individual Contributor	50%
Team Lead	50%
1 Management Level	40%
2 Management Levels	45%

In a typical week, when you are not on call, around what percent of your time is spent on operations activities?



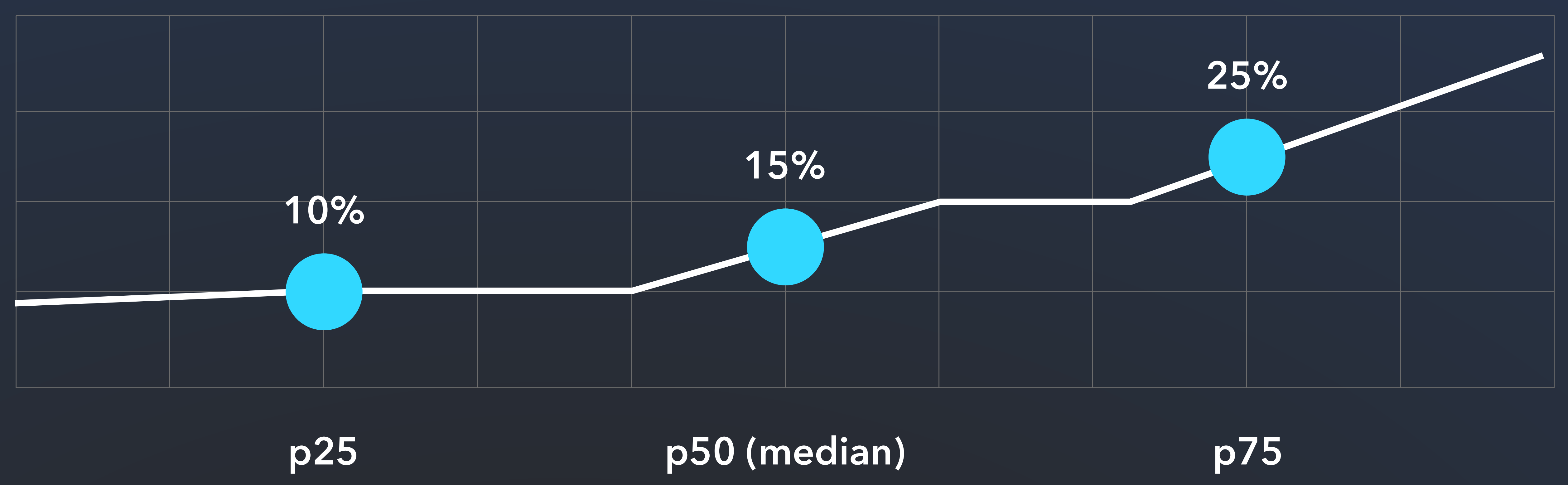
	p50 (median)
Aggregate	25%
Individual Contributor	25%
Team Lead	20%
1 Management Level	20%
2 Management Levels	25%

The median values across all survey respondents for time spent on engineering activities versus operations activities were 50% and 25%, respectively.

Example engineering activities given in the survey were ‘Writing code’ and ‘Developing an auto-provisioning capability’. Example operations activities given in the survey were ‘Responding to pages’ and ‘Manually spinning up new hosts/ instances’.

We typically present this data as a distribution because we believe in the power and truth of not lying with averages.

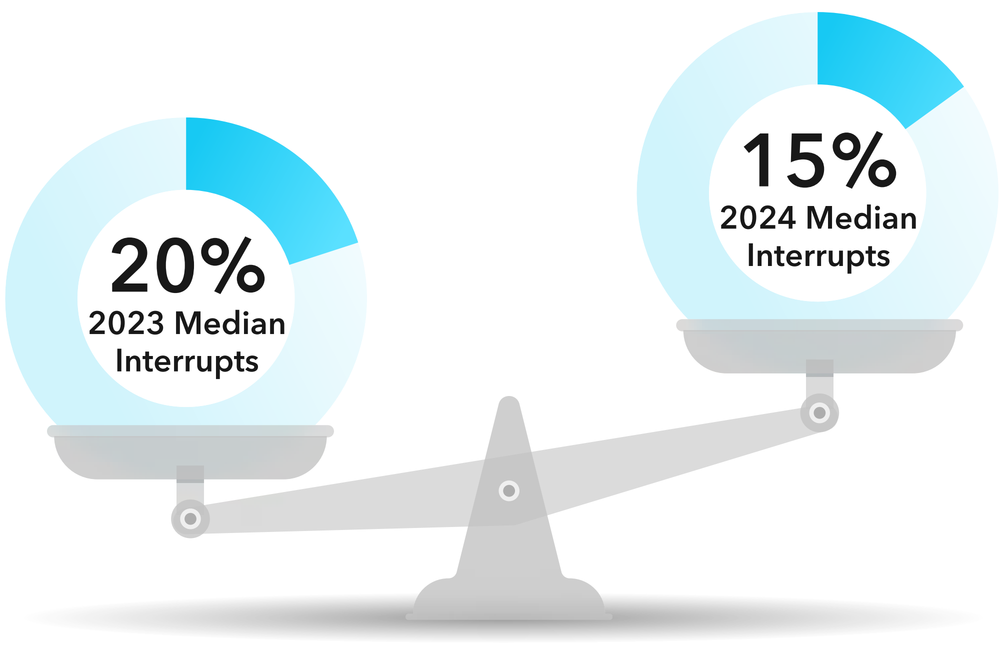
In a typical week, when you are not on call, around what percent of your time is consumed responding to interrupts?



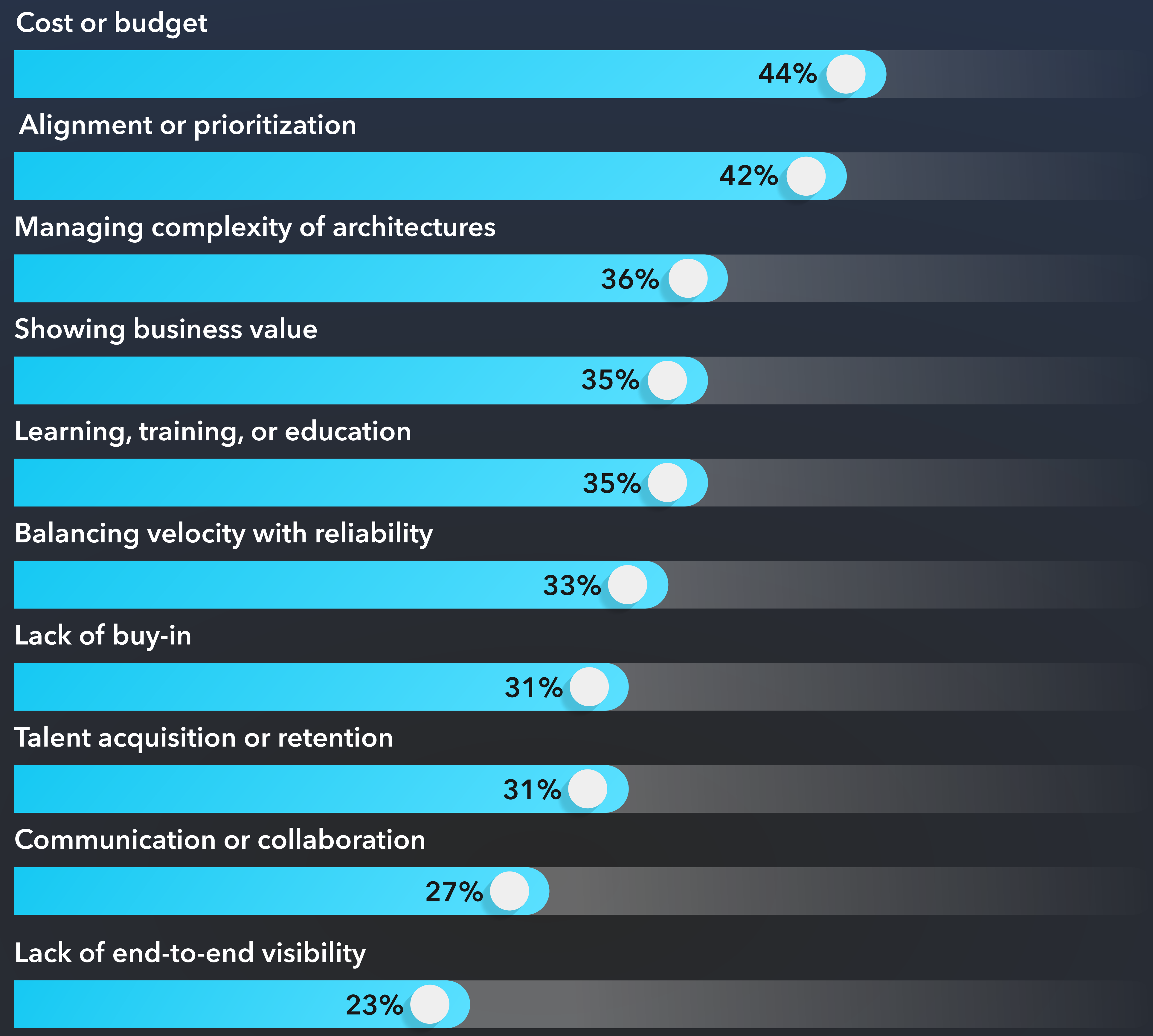
	p50 (median)
Aggregate	15%
Individual Contributor	15%
Team Lead	10%
1 Management Level	20%
2 Management Levels	14%

Inevitably, we will all face a shoulder tap or notification starting with 'Hey, do you have a minute to chat?' or 'Hey, can you hop on this call?' To make matters worse, these interruptions may make us completely switch contexts, increasing the time of the interruption.

This is the second year we asked this interrupt question. Median time spent on interrupts was 15% (with higher organizational ranks self-reporting more interrupts), down from 20% in The SRE Report 2023. Given the relatively large volume of non-credited work on incidents, it might be worth questioning if this self-reported figure is actually higher in reality.



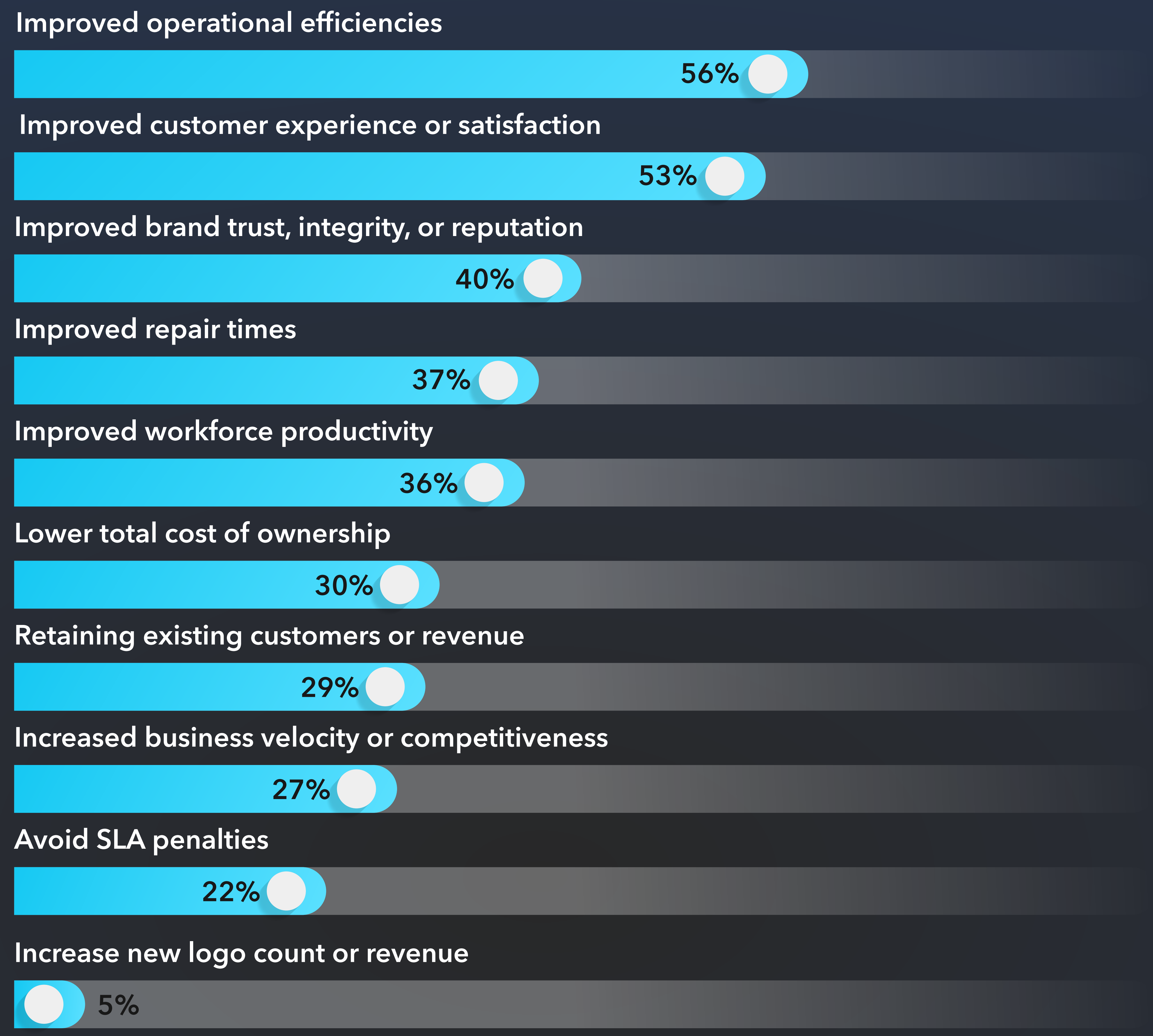
What are the top challenges hindering successful reliability implementations?



Reliability practitioners see cost or budget as the biggest challenge to achieving successful reliability implementations. This is also why, as we wrote in Insight VI, we need to have a different kind of conversation around SRE tooling and business value.

Closely following is 'alignment or prioritization', which equally demonstrates the key need for the type of value-based conversation SREs need to drive with management to better achieve their reliability priorities.

What are the top ways in which reliability engineering practices add business value?



Businesses need people to create or preserve revenue. Conversely (most) people need healthy businesses to pay their rents and mortgages. That is why it is important to consider outcome-based conversations through a business lens.

'Improved operational efficiencies' topped the list of ways in which reliability engineering practices add business value, closely followed by 'improved customer experience or satisfaction'.

When we asked this as an open-ended question last year, the leading responses were similar with 'lower cost' and 'improving customer experience, sentiment or satisfaction' tying for first.

We hope all ranks within the reliability field can use data and context from this report to have data-driven conversations to guide aligned decision-making.

Insights I II III IV V VI VII

View from the field

SREs are the invisible force behind site and service reliability. Individual success is often only seen in a surfeit of metrics. However, individual pride in one's work often serves as the single source of motivation. We all seek personal fulfillment and a sense of purpose. High quality, innovative contributions to projects or complex issues give us that deep sense of accomplishment. It also challenges us and pushes us on the path of continuous growth.

For managers, the challenge lies in balancing the time between team member motivation and development, strategic efforts, operational issues and administrative tasks. This drives the need to minimize the amount of time, effort and resources. But, in which area? If being proud of one's work as an SRE manager equates to pride in team success, should one be efficient when focusing on team motivation and development? Especially, if being 'proud of their work' is most important to SREs?

Organizations would greatly benefit by minimizing administrative tasks and allowing more time for training in improved communication, coaching skills and strategic thinking skills for engineers and managers. This would shift the mindset from efficiency to effectiveness which would lead to increased value of the team and increased pride across the individual team members.

Susan Shimamura

Leadership and Team Coach/Facilitator
Why Not? Coaching, LLC

The SRE Report 2024

Until Next Year

When I think about *control*, I think about my time on the production operations team at Ask.com, formerly Ask Jeeves. We had a relationship with Google and the Google ad call was one of our most monitored endpoints. Why? Because when there was a problem with it (that is, when there was a problem with one of our revenue streams), our CEO was in **our** office(s), not Google's, asking when it was going to be fixed. When it comes to federating and relying on *stuff outside our control*, I see the battle lines being drawn when they go into incident. On one side, I see the empathetic: the people who understand it's amazing any of this works at all. But I fear that empathy may be abused, which is probably why the other side exists: the apathetic. This group may intentionally or unintentionally be this way. As with most, these opposing forces should help us find balance (at least I hope).

There are many other dimensions – or layers - to consider (Yes, reliability practitioners and practices are onions). This report also does a nice job of breaking down by some of them like organizational rank or company size. But there are many other dimensions as well. Take the time dimension as an example. If you surveyed 100 people, then you would get a set of insights. If you surveyed those same 100 people a year later, then you might get a different set of insights. So when reading (and hopefully re-reading) this report, consider these and other situational contexts to appreciate the words on the paper.

I am proud to contribute this report to the larger community. But if you asked me the "Being which of these is most important to you [when working on this report]?" question, I would have answered as *grateful*. I would have answered this way because it is the community that makes this report possible, and I am grateful for being a part of it.

I look forward to The SRE Report 2025. It will undoubtedly contain fresh insights, stories, and surprises. Bring it on!

Leo Vasiliou

Author of The SRE Report and Catchpoint PMM

Afterword

We want to do *just enough* of this reliability work to make our customers happy, no more. We want to avoid "optimization traps" and focus on what's actually holding us back. But how do we know how to do this? We evolve, adapt, and focus on the outcomes. Don't hyperfixate on things like overlapping tools and teams going “outside their lane". Expect your complex system of humans and machines to produce emergent behavior and then be ready to respond to it. Don't place anything in stone. Write down your rationales for decisions, even if the rationale was "we just guessed." Make sure people know what others knew (or didn't know) in the past when decisions were made. Empower teams so that they can change the system. New hires and old-timers have different reasons for complaisance or compliance; make sure you address both. Let teams discover those rationales and question if they're still relevant (or if they ever were!). Budget-holders, let your teams know how much they're spending. Let them know what other teams are spending. Give them the data, incentives, and clearly articulated outcomes so they can make good system-wide decisions.

Leaders, make it safe to report incidents. Don't incentivize teams to hide them by demanding fewer incidents. The goal is to understand the nature of the system that is affected. We "learn from incidents" not "about" them – learning about the system through the lens of its observable problems. Can we ever actually understand our systems in complete detail? They run at gigahertz speed, spread across the globe, how could we? Incidents are a way of learning new things about something you thought you already understood; don't miss that opportunity!

Recognize emergent behavior and understand that while it isn't predictable, you can adapt to it. Being able to adapt to new failure modes is at the heart of SRE, these new stewards of complex, society-impacting systems. This is a lot of responsibility but get ready because more is coming: AI. Those building products alongside new AI systems will face new challenges and they will require a new mindset: not expecting predefined best practices or taxonomies of known failure modes, instead the willingness to probe, sense, and respond to these complex systems, forever and ever. These systems will not replace SRE, but SRE will certainly adapt to them.

You already know not to "waterfall" SRE. Resist the temptation to define ideal reliability for your organization up-front. Instead, take what you've learned here and elsewhere, learn about your system from incidents, and prepare your teams to become more adaptable.

Steve McGhee

with special thanks to Jessica DeVita

Demographics

The SRE Survey, used to generate insights for this report, was open for six weeks during June and July of 2023. The survey received 423 responses from all across the world, and from all types of reliability roles. Catchpoint made donations in the amount of $2,115 to the International Committee of The Red Cross.

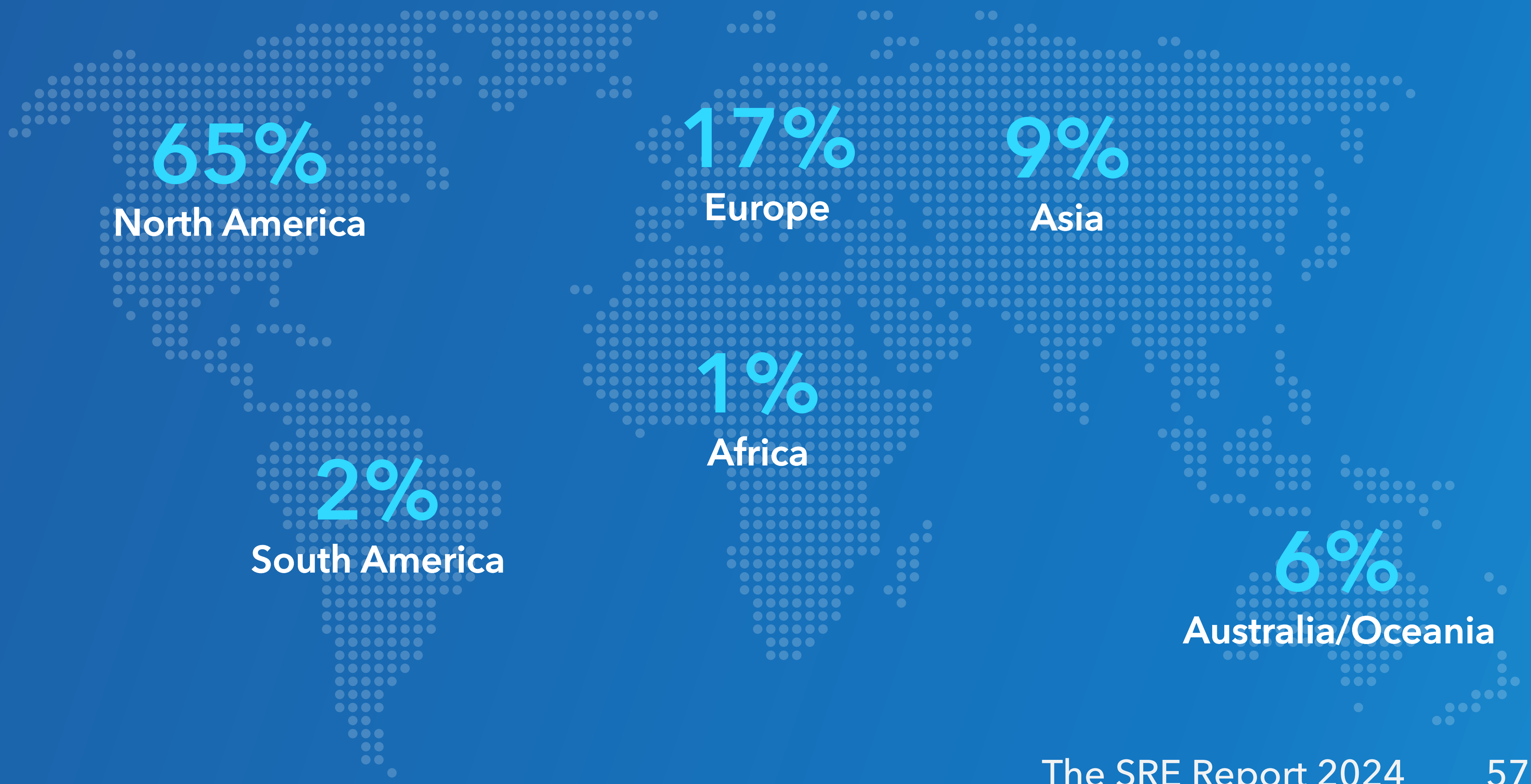
In what industry/vertical does your company operate (with multiple select)?

Technology or Technology Provider	40%
Financial Services	13%
Healthcare or Chemicals	12%
Government or Non-profit	7%
Professional Services or Consulting	7%
Consumer Packaged Goods or Retail	7%
Media or Entertainment	7%
Manufacturing	6%
Telecom	6%
Conglomerate: Across many industries/verticals	4%
Energy	3%
Travel or Accommodation	2%
Transportation	2%
Other	9%

How many employees does your company have?

2%	Up to ten
10%	Between 11-100
10%	Between 101-250
10%	Between 251-500
15%	Between 501-1,000
25%	Between 1,001-10,000
18%	Between 10,000-100,000
11%	More than 100,000

Where are you (personally) located?



Acknowledgements

Catchpoint wishes to thank the SRE community for providing the survey data which was used to create The SRE Report. We also want to thank the entire report production team for their hard work, dedication, and grit. Without the community and production team, this report would not have been possible.

Survey Authors and Report Analysts Leo Vasiliou and Kurt Andersen

Report Writers Leo Vasiliou and Kurt Andersen

Report Editors Anna Jones and Leo Vasiliou

Report Graphic Designer Beril Unal

Report Project Coordinator Jennay Wangen

Report Passioneer Steve McGhee

Report Contributors (in order of appearance) Sarah Butt, Alex Elman, Kurt Andersen, Dr. Laura Maguire, Niall Murphy, Alex Hidalgo, Leon Johnson, and Susan Shimamura

Executive Sponsors Drit Suljoti, Gerardo Dada, Howard Beader, and Mehdi Daoudi

catchpoint

About Catchpoint

Catchpoint is the Internet Resilience Company™. The top online retailers, Global2000, CDNs, cloud service providers, and xSPs in the world rely on Catchpoint to increase their resilience by catching any issues in the Internet Stack before they impact their business. Catchpoint's Internet Performance Monitoring (IPM) suite offers web synthetics, internet synthetics, RUM, performance optimization, high fidelity data and flexible visualizations with advanced analytics. It leverages thousands of global vantage points (including inside wireless networks, BGP, backbone, last mile, endpoint, enterprise, ISPs and more) to provide unparalleled observability into anything that impacts your customers, workforce, networks, website performance, applications, and APIs.

Learn more at: https://www.catchpoint.com/

Follow us on LinkedIn: in