APPLAUSE

2024 State of Digital Quality

An in-depth exploration of the frameworks, strategies and challenges of building exceptional user experiences in today's digital age.

Contents

Executive summary

As the world leader in testing and digital quality, Applause has a distinct perspective. Members of our team are intimately involved in testing and test strategy for global technology leaders and innovators. The State of Digital Quality Report draws on some of our learnings to help guide organizations in their efforts to improve overall digital quality while speeding releases and boosting efficiency.

For the third year, Applause analyzed trends in software development, testing and quality assurance to benchmark the state of digital quality. In addition to mining our crowdtesting platform data, we also surveyed the community and our enterprise customers to better understand how development organizations fit into the digital quality continuum.

Last year, we introduced frameworks to help organizations assess their progress toward achieving comprehensive digital quality, including the core capabilities, practices and processes they should adopt. The frameworks cover functional testing, accessibility and inclusive design, payments, localization – and now, user experience (UX) and artificial intelligence (AI).

These frameworks provide a means for organizations and teams to benchmark their current capabilities and map out a plan for improvement. While certain teams have achieved excellence in some areas, only a select few have established cultures where quality is thoroughly embedded into the organization's DNA across every aspect of their digital experiences. It's a tough goal to achieve and even more difficult to maintain — every release comes with new code to test, documentation to write, test cases to update and execute, and decisions to make. It's a constantly moving target. And, it's a goal the teams at Applause – and most end users – believe is worth pursuing.

Highlights from this year's report

- **Measuring success is still a challenge.** Organizations rely on a variety of measurements and KPIs to assess digital quality across multiple dimensions, with customer-centric measures the most common indicators of success. Despite reliance on multiple KPIs, many teams still have an incomplete picture of how well they're performing.

- **Process and documentation are often overlooked to meet tight deadlines.** Fewer than one-third of respondents reported that their organization has comprehensive documentation for test cases and test plans. Though most teams recognize the fundamental value of clearly defined test methodologies, accurate documentation, and testing and feedback throughout the software development lifecycle (SDLC), these strategic imperatives are often sacrificed in favor of speed.

- **Despite being a priority, digital accessibility resources are lacking.** Our 2024 accessibility and inclusive design survey revealed that, though 42% of respondents rated accessibility as a top priority for their organization, 29% do not involve people with disabilities (PWD) directly in the design or testing process. Furthermore, 44% of respondents said they have limited or no in-house resources for accessibility testing — a fact that seems at odds with an issue considered a top priority.

- **High-quality user experiences make or break Gen AI.** While generative AI (Gen AI) holds the potential for hyper-personalization in the future, at present, the best way for organizations to make customers feel recognized and appreciated is to focus on UX, delivering experiences that account for preferences across different markets and demographics. Our AI survey found that 27% of Gen AI users have swapped one service for another due to UX and performance issues. The ruthless pursuit of opportunities to delight and relentless commitment to rooting out flaws, friction and bias are the only ways to stand out in markets where everyone is fighting for wallet share.

Methodology

We analyzed results from a representative sample of closed test cycles executed in 2023. A test cycle is how Applause defines each unique set of tests: a client sends us testing parameters — builds, scope, coverage, etc. — and we create a test cycle that includes the specific test cases and scenarios to be tested. In addition, we conducted several surveys of the uTest community and Applause customers throughout the year to identify trends in development, testing and consumer preferences across different types of digital experiences.

The data includes various industries, testing categories and regions spanning more than 748,000 test runs, tens of thousands of devices, and thousands of device/OS/browser combinations. Testing included websites, IoT devices, mobile web and mobile apps in real-world scenarios.

Mobile	
Mobile makes	**126**
Mobile models	**2,102**
Mobile OSes	**17**
Mobile OS versions	**420**
Mobile web browsers	**37**
Mobile carriers	**527**

Other devices	
Set-top/streaming devices	**187**
Gaming consoles	**35**
Smart TVs	**2,270**
TV providers	**87**
ISPs	**89**
Smart home devices	**61**

Desktop	
Desktop web browsers	**35**
Desktop OSes	**9**
Desktop OS versions	**279**

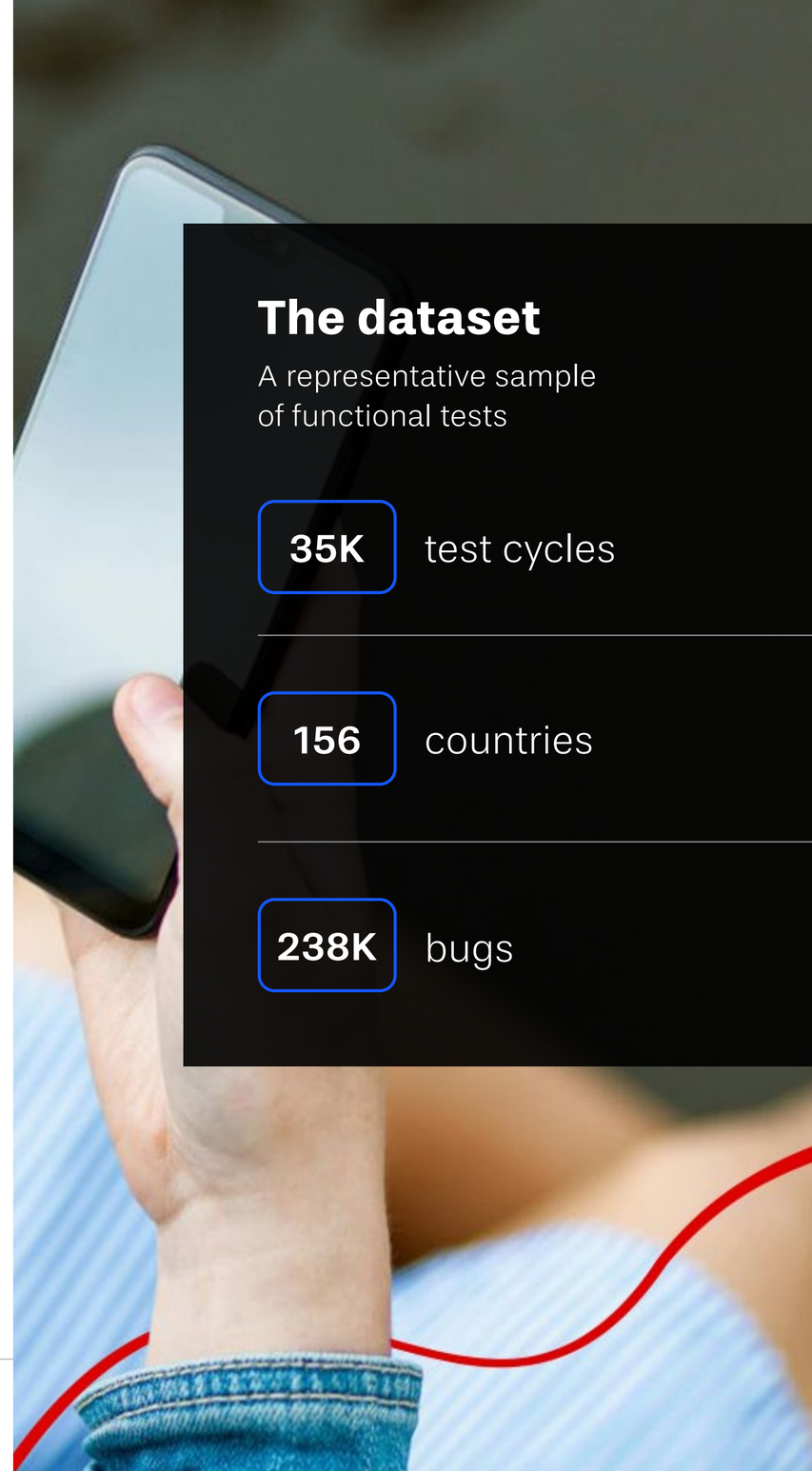
Payment methods	
Credit/debit cards	**2,208**
E-wallets	**50**
Mobile wallets	**87**
Alternate payments	**80**

Functional testing

Last year, Applause testers executed nearly three-quarters of a million test cases. Though about 80% of them passed, the remaining 20% of failed test cases reflected defects across all levels of severity and scope, from reputation-destroying disasters to minor usability challenges. This rate, fairly average across industries, is typically accepted as a good rate. It leaves room to improve on each release and test cycle due to changes in the feature base and code being introduced. No one expects perfection — yet teams continue to seek out and address issues that seriously detract from the user experience.

Often, organizations are so focused on releasing quickly that they don't take time to set themselves up for long-term success. Keeping code clean, writing good test cases, documenting test run results and leveraging data to focus efforts allows development and QA teams to become both more effective and efficient. These tasks also lay the groundwork so they can scale with minimal growing pains.

Ideally, digital quality focuses less on finding defects and more on creating systems and processes at all stages of development that prevent them from occurring in the first place. All frameworks assume that many/most of the elements from the corresponding stage of the functional testing framework are in place.



Functional testing framework

Digital quality emergence:

Lack of consistent systems, processes and documentation — while individuals may have their own methods and documentation, the organization has no consistent methodology or approach to quality.

Examples of testing activities and processes:

- Dogfooding
- Conducting reactive testing after development
- Running tests without documenting test cases or test run results; or documenting poorly – defects may not be reproducible based on bug reports
- Running test cases inconsistently
- Skipping test case documentation or writing test cases in ways that make them difficult to follow or reproduce
- Practicing some exploratory testing
- Using disparate environments for development, testing and production

Digital quality essentials:

Early stages of defining and documenting processes and procedures; establishing some consistency and structure around test efforts. Teams may have their own unique processes, but efforts may still be siloed.

Examples of testing activities and processes:

- Documenting test cases for feature-based tests
- Ensuring test cases are written clearly
- Defining a device coverage matrix
- Maintaining a known issues/bug backlog list
- Testing releases pre-production
- Conducting unit, smoke and regression testing for major app components or workflows
- Performing exploratory testing for new features/app changes
- Recording test run results
- Automating frequently executed/rarely changed tests

Digital quality expansion:

Clear processes and a broad range of testing types in place. Some reporting is in place. Focus is on coverage, scalability and efficiency across the organization.

Examples of testing activities and processes:

- Maintaining a defined device coverage matrix based on data about website/app usage
- Conducting regression testing for all workflows
- Testing user acceptance and UX for new features/app changes
- Leveraging test automation for repetitive tests; reviewing and updating automation scripts regularly
- Documenting test cases/suites for all features
- Measuring quality KPIs with data and reporting

Digital quality excellence:

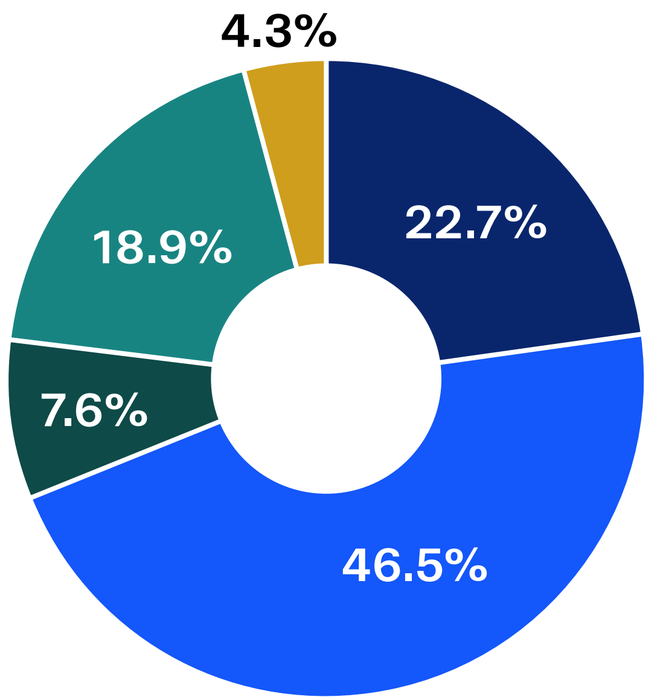
Quality is embedded in the company's DNA and built into all products and experiences from end to end.

Examples of testing activities and processes:

- Testing throughout the SDLC, in-sprint and in staging/pre-production
- Incorporating the voice of the customer into product design and development
- Delivering exceptional UX across all touchpoints
- Maintaining a strong test case management process
- Automating all repetitive tests that humans can't do better
- Reviewing and refining testing processes regularly
- Proactively balancing testing across manual functional, exploratory and automated testing; documenting when to use each test type
- Exploring new testing processes to maintain high levels of quality, efficiency and coverage
- Driving innovation throughout the SDLC
- Using reports to analyze trends and identify areas for improvement

To determine where most organizations fit into the framework, Applause surveyed software development and quality professionals. Here are some of the findings.

Roles n=6,286



22.7% Engineering

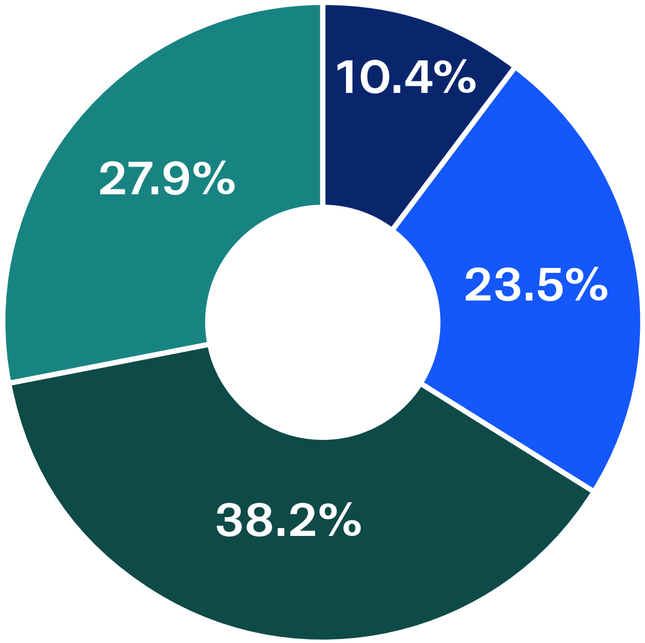
46.5% QA

7.6% DevOps

18.9% Product

4.3% Ops

How would you rank your organization's structure and consistency when it comes to digital quality? n=3,934



10.4% Emergence

23.5% Essentials

38.2% Expansion

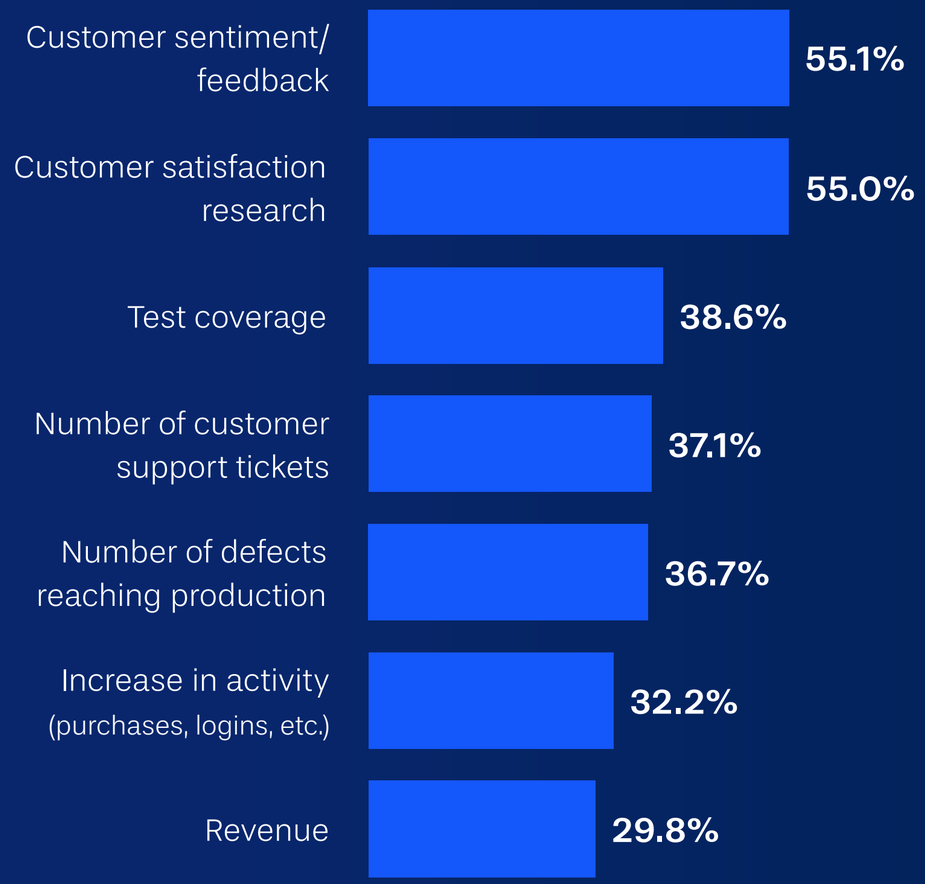
27.9% Excellence

While 27.9% of respondents stated that their organization has comprehensive end-to-end testing processes and documentation that are used consistently across the organization for all products, 10.4% are still in the earliest stages of establishing a digital quality program. The largest pool of respondents — 38.2% — have some clear testing processes, documentation and reporting in place that they're working to scale across the organization, while the remainder have some building blocks in place, but are still solidifying other essential elements of a quality program.

Digital quality is more than just ticking a box or conducting certain activities. So how do different groups define quality and monitor their progress? Organizations are using a combination of metrics to assess digital quality: 73.9% report that their organization uses three or more metrics, 46.9% report their organization uses five or more metrics, and 12.9% use 10 or more.

Most common quality indicators

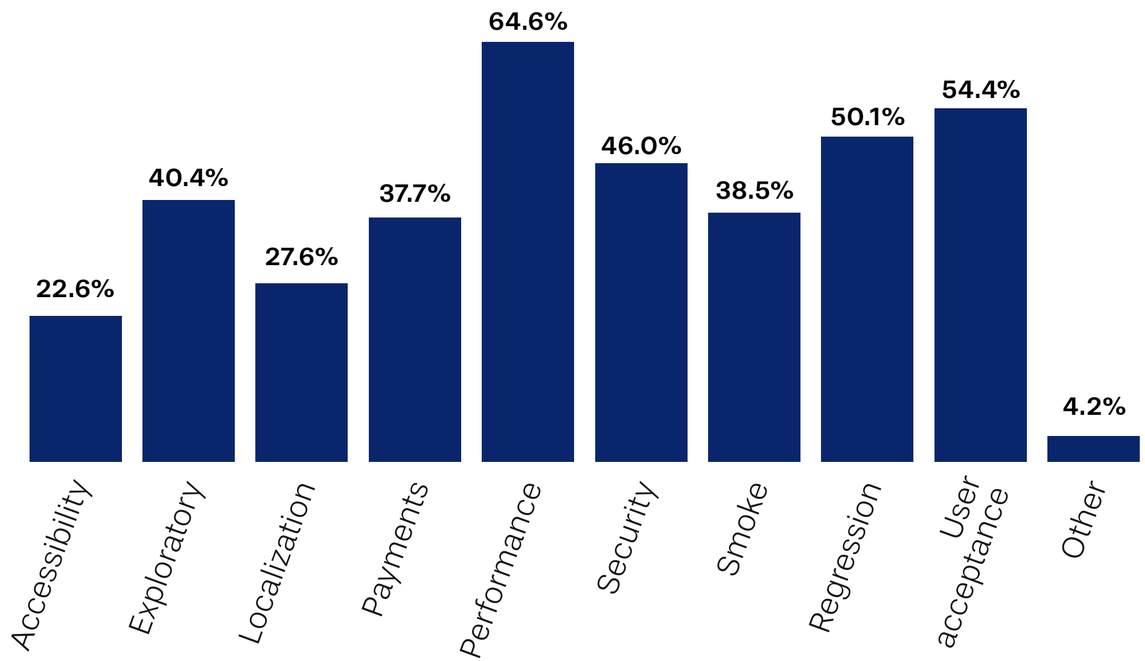
n=4,824



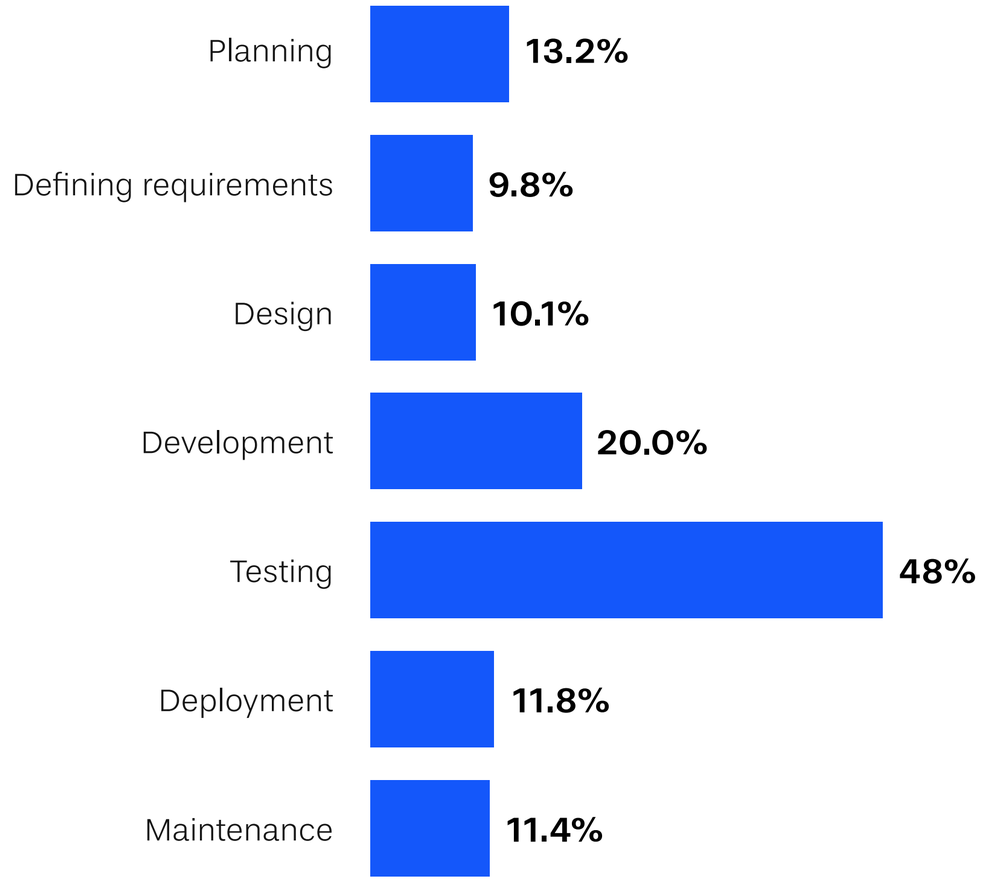
Most organizations are conducting a variety of different types of tests as well. Although most organizations focus predominantly on functional testing, we see a number of companies increasing their investment in non-functional testing to become more competitive.

Types of tests organizations are conducting

n=4,659



Where do you test and collect feedback in the SDLC? n=3,709



While testing throughout the SDLC is a best practice, many organizations aren't doing so. In this year's survey, 41.7% reported that their organization only conducts tests during the testing phase of the SDLC; another 7.1% only test during deployment.

When asked what types of testing and quality control procedures teams conduct during development, respondents reported using a variety of different approaches. The most common activities:

61.5%
Unit testing

58.9%
Integration testing

52.5%
Checking for bugs in staging

38.3%
Peer review

28.0%
Static code review

Exploratory testing

Nearly 60% of the teams that conduct exploratory testing incorporate it into every release, and a similar number said they explore whenever they've changed a feature or workflow.

Having testers spend a little time exploring aspects of the application that have changed offers an easy way to uncover flaws before customers encounter them during their typical use patterns. Focused exploratory testing, especially in the hands of creative testers, can quickly uncover issues and ensure better releases.

How often do you conduct exploratory testing? n=1,708

59.0%
Every release

58.8%
Whenever we change a feature or workflow

54.1%
Whenever we launch a new feature

How do you define exploratory testing?

> Simultaneous learning, test design, and execution. It focuses on discovery and relies on the guidance of the individual tester to uncover defects that are not easily covered in the scope of other tests."

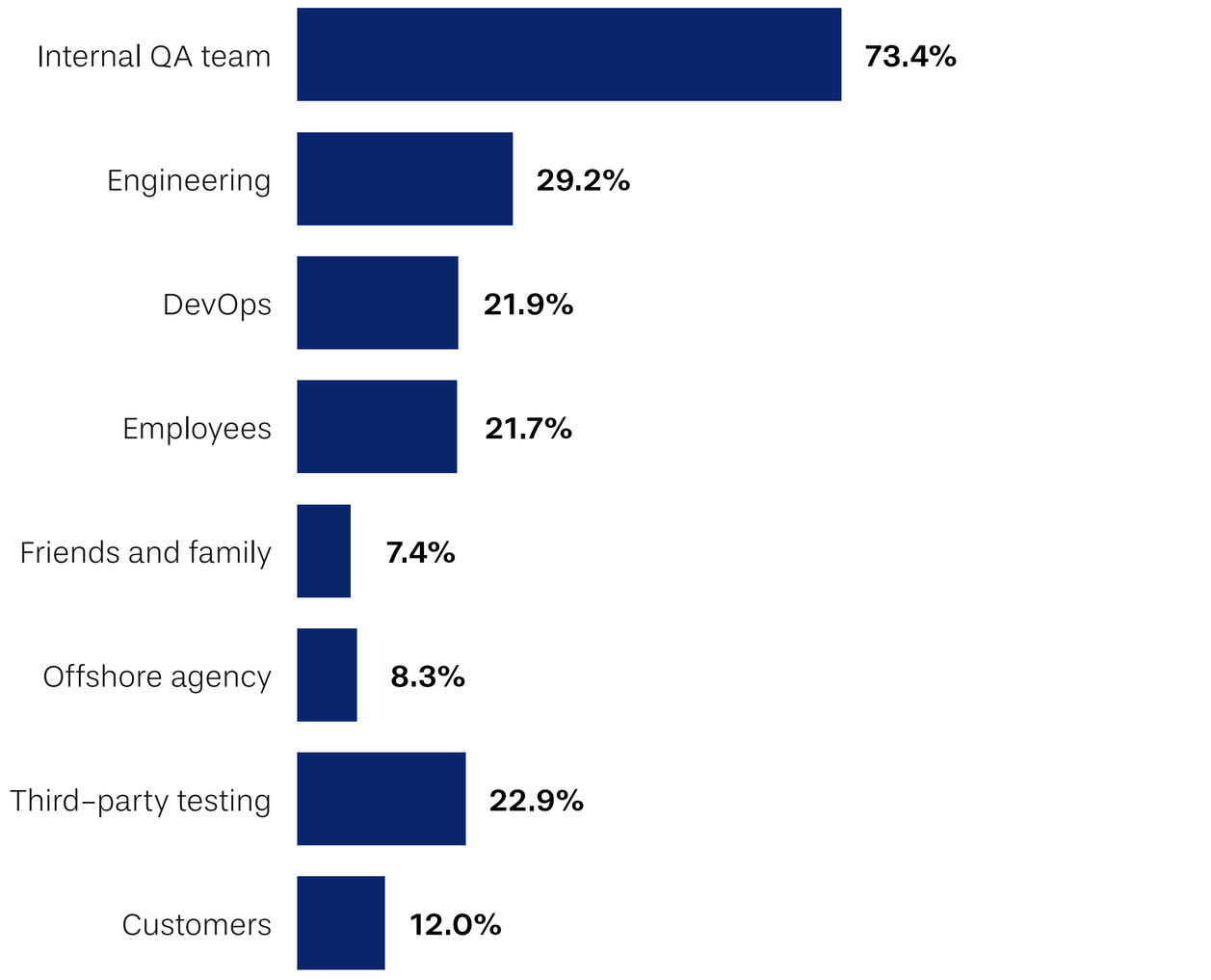
> Exploratory testing is an informal testing approach where testers dynamically explore the software, design and execute test cases on the fly, and adapt their testing based on immediate feedback. It emphasizes learning about the system, uncovering defects, and refining test scenarios as the testing process unfolds."

> Flexible tests where the skills and experience of the testers are of utmost importance. These tests allow us to find defects that were not initially considered."

Most organizations reported that a QA team conducts testing (73.5%), with 30.6% of respondents stating that QA members are the only ones involved in testing. Typically, multiple groups provide testing at some point, with 29.2% stating engineering staff performs tests and 21.9% including DevOps.

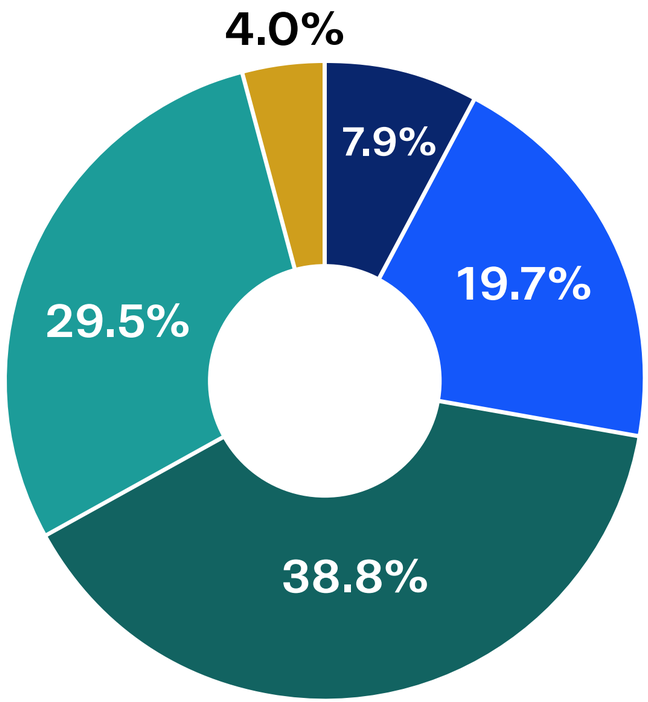
No matter who performs testing, ensure that staff members have the opportunity to test different portions of the app or conduct varied types of tests to maintain a fresh perspective and reduce the possibility of blind spots or tester fatigue.

Who carries out testing for your organization? n=3,941



How well does your organization document test cases and test plans?

n=3,939



7.9% Little to no documentation

19.7% Some documentation, but not comprehensive

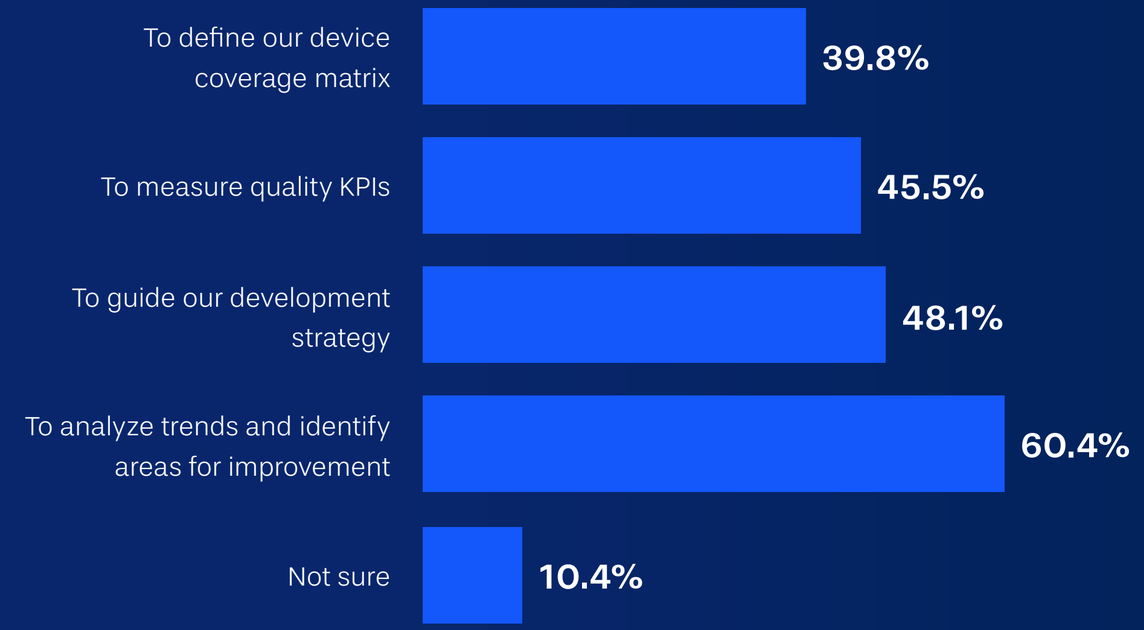
38.8% Clear documentation for most features

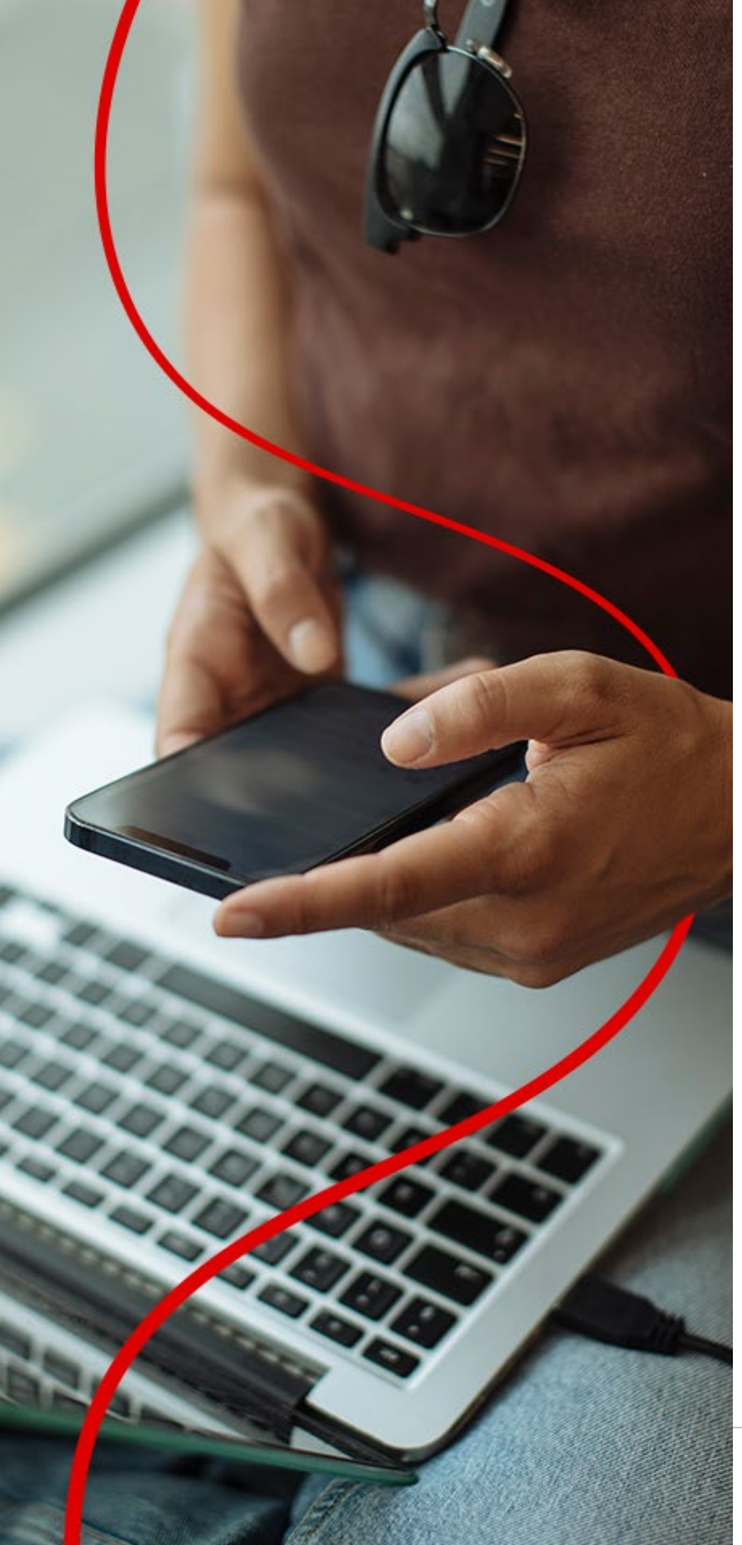
29.5% Comprehensive documentation

4.0% I'm not sure

How is your team using test reporting and metrics?

n=3,934





Functional testing recommendations

Embrace change as part of the process of continuous improvement. While consistency is part of quality, so is adaptability. Regularly revisit each process to review why it is in place, whether it is still necessary, and whether there's a better way to meet the same goal. Adjust how the organization performs certain tasks as needed to make better use of available resources, eliminate inefficiency and reduce gaps in coverage.

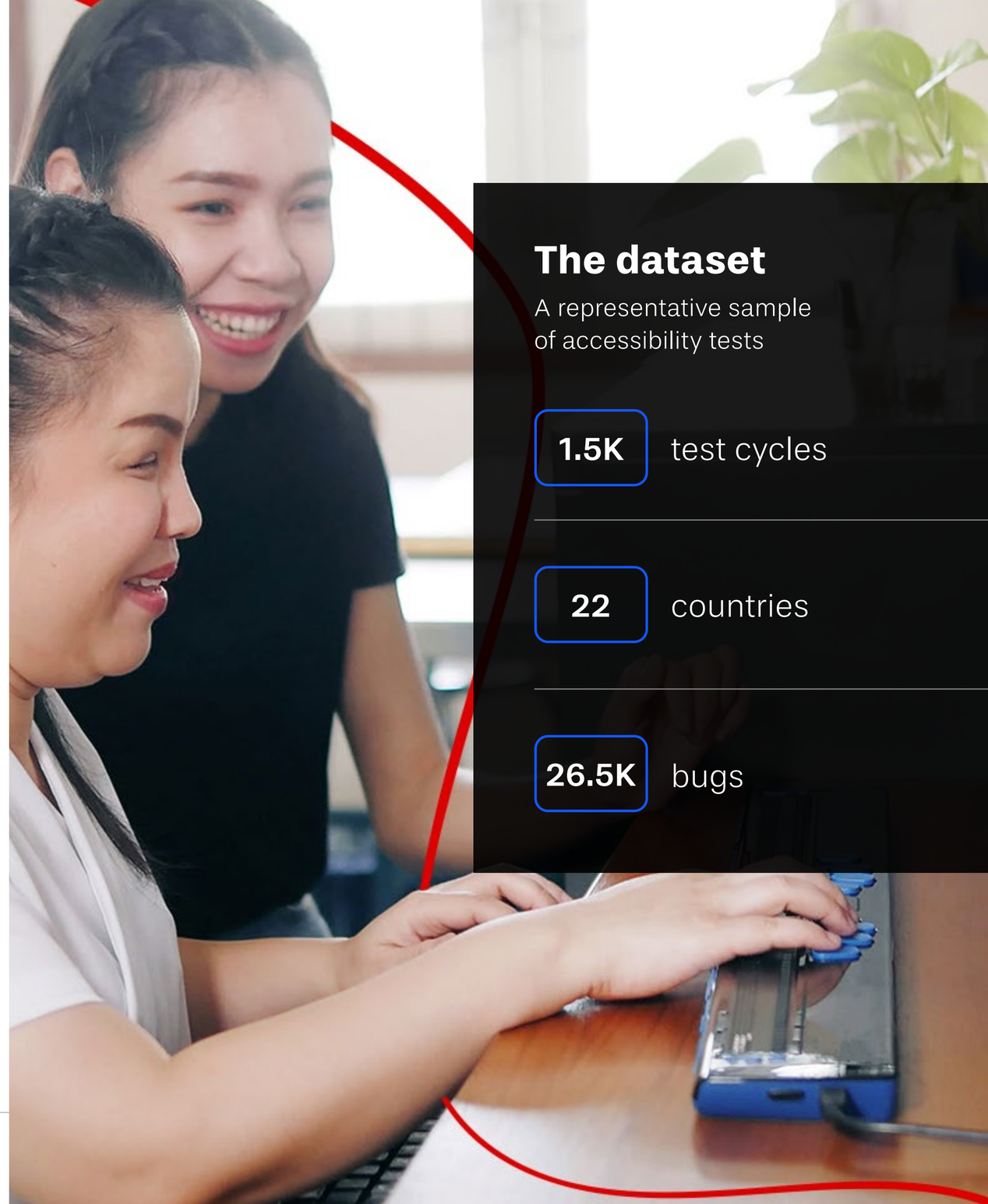
Conduct testing and collect feedback throughout design, development and deployment. The earlier in the process you start, the easier it is to adjust, adapt and evolve, which saves both time and money. Those critical resources offer a competitive advantage as businesses work to release better and faster while maximizing profits.

Focus on building repeatable, scalable systems and processes as the foundation for great digital quality. When everyone involved in development and testing is clear on roles, responsibilities, requirements and testing procedures, it's easier to identify — and fill — gaps. Let data be your guide. Use existing reports to determine which areas require most attention so you can improve coverage, consistency and code across the organization.

Accessibility and inclusive design

While accessibility typically starts with conformance and compliance, once organizations embrace empathy-based design and development, the user experience improves dramatically. Our customers who have adapted to focus on inclusive design report that this shift left creates a cycle of ongoing innovation that benefits all users, not just PWD.

With the EU Accessibility Act (EAA) coming into effect in June 2025, global organizations have heightened their focus on accessible and inclusive designs, raising the bar for all application developers and technology teams.



The dataset

A representative sample of accessibility tests

1.5K test cycles

22 countries

26.5K bugs

Accessibility framework

Digital quality emergence:

Limited understanding of applicable laws and regulatory requirements. Lack of formal systems, processes and documentation, no consistent methodology or approach to accessibility or inclusivity.

Examples of testing activities and processes:

- Identifying some accessibility issues; resolving or remediating high-priority A11y issues that are identified
- Assessing conformance to WCAG and/or locally applicable regulations once or infrequently
- Emphasizing risk mitigation; making changes in response to customer complaints or threats of legal action
- Relying solely on automated tools

Digital quality essentials:

Early stages of defining and documenting processes and procedures; establishing some consistency and structure around accessibility.

Examples of testing activities, resources and processes:

- Designating an owner of the accessibility program and identifying organizational champions
- Performing periodic assessments and automated checks to ensure conformance to WCAG and compliance with locally applicable regulations
- Offering training on accessibility best practices and inclusive design
- Providing developers and/or product owners access to people with disabilities (PWD) for initial product feedback
- Conducting design reviews and in-sprint testing to identify problems earlier in the SDLC
- Engaging PWD to perform usability tests and provide feedback

Digital quality expansion:

The organization has a clear process for ensuring that accessibility is in place and uses various testing types. Some reporting is in place. Focus on coverage, scalability and efficiency across the organization.

Examples of testing activities, resources and processes:

- Documenting best practices and checking them during development
- Maintaining a knowledge base and offering training on best practices
- Holding empathy-based design and development workshops with PWD
- Incorporating input from PWD into the design & development process; benchmarking to understand the validity of customer criticisms around accessibility and usability
- Prioritizing accessibility and inclusivity in the organization's design/UI kit
- Providing attestation (VPATS)

Digital quality excellence:

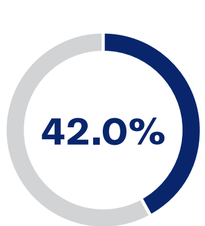
The organization prioritizes inclusivity at all levels; testing and feedback from PWD occur throughout the SDLC.

Examples of testing activities, inclusive design resources and processes:

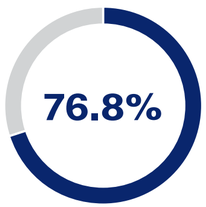
- Creating a Center of Excellence with guidance on design and development of best practices
- Requiring ongoing, mandatory training on accessibility and inclusive design best practices
- Following inclusive hiring practices to develop a diverse workforce
- Going beyond attestation to demonstrate thought leadership in inclusive design and product development
- Incorporating accessibility into the company's contracting and procurement process
- Using analytics to drive priorities and continuous improvement

Where organizations fall on the spectrum

In our survey focused on accessibility and inclusive design, 42% of more than 3,000 respondents indicated that accessibility is a top priority for their organization and 76.8% indicated there is a group or person responsible for ensuring products are accessible.

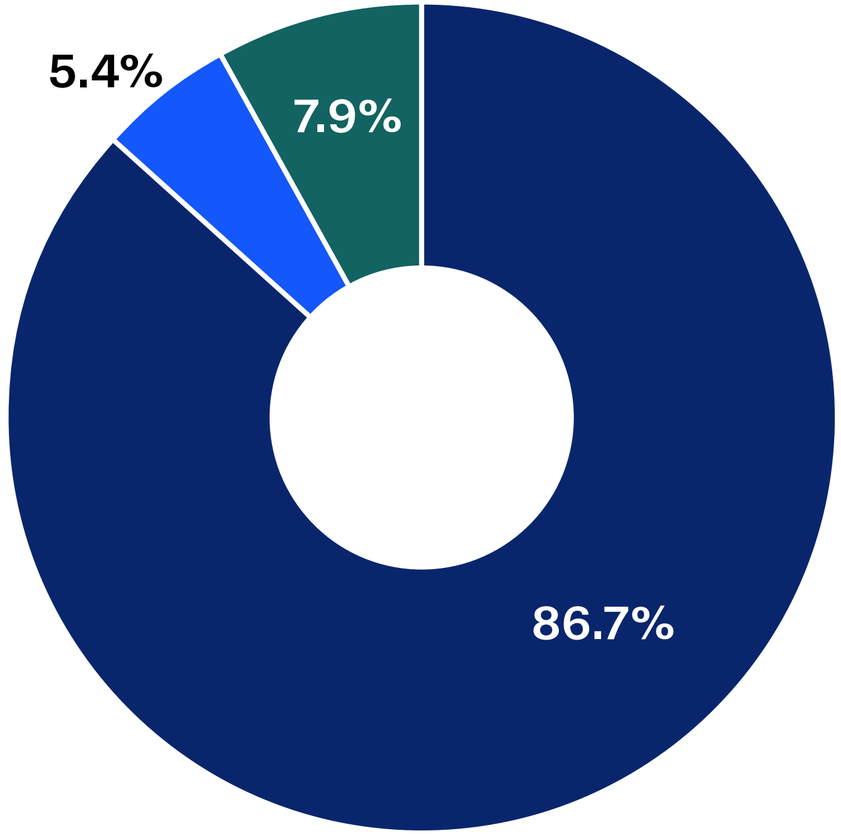


42.0% indicated that accessibility is a top priority for their organization



76.8% indicated there is a group or person responsible for ensuring products are accessible

Does the team or expert responsible for accessibility in your organization employ inclusive design principles? n=3,939



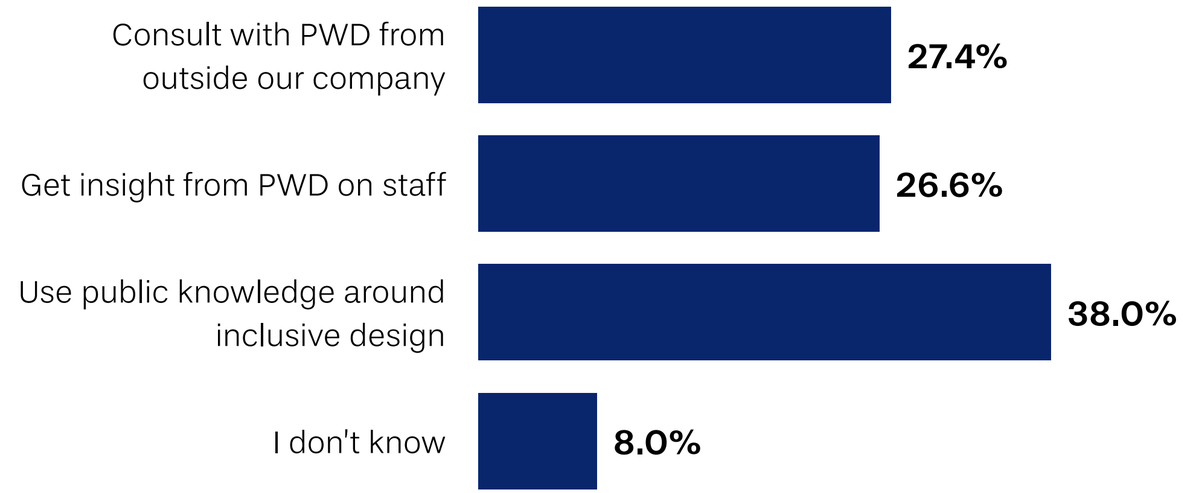
86.7%	Yes
5.4%	No
7.9%	I don't know

While most respondents reported their organization captures input from PWD at some point in the design and development process, 28.7% stated that their organizations rely solely on publicly available knowledge around inclusive design and have no direct interaction with PWD.

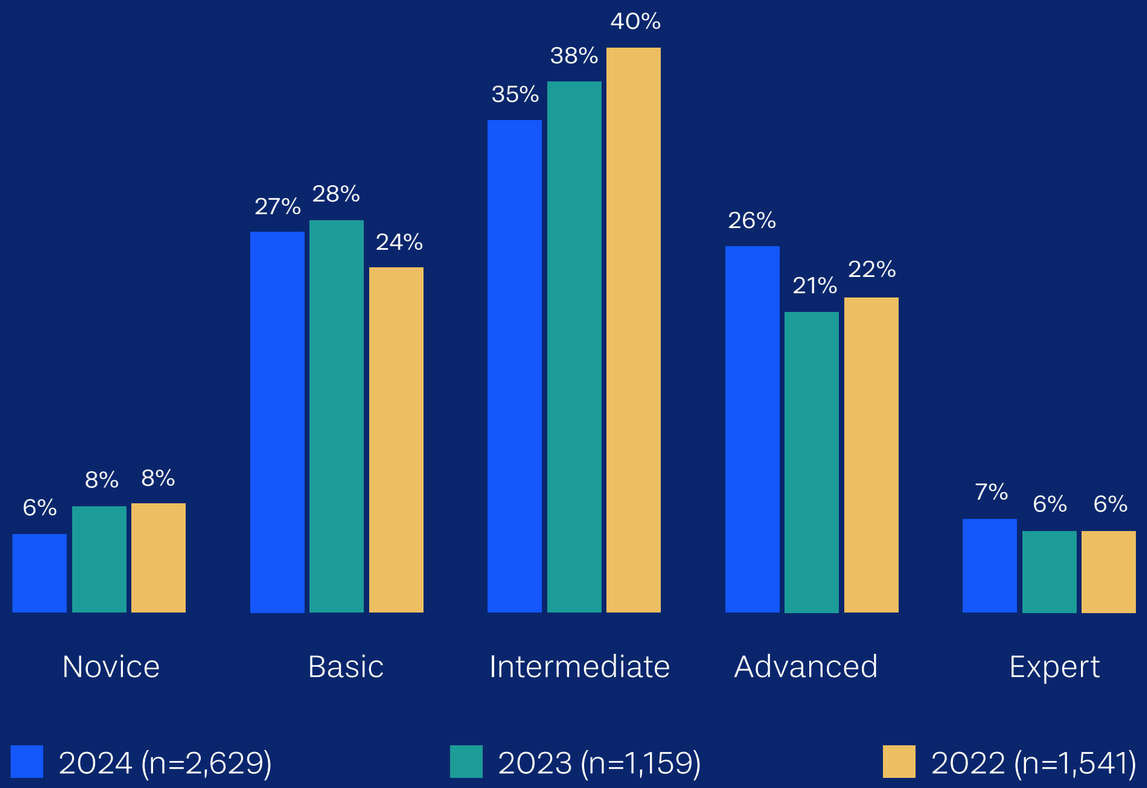
Understanding of accessibility is on the rise, if only slightly. Over the last two years, the percentage of respondents reporting advanced and expert knowledge of digital accessibility has risen while the number rating themselves as novices has decreased. The number reporting that they or someone they know have benefited from digital accessibility has also increased slightly in that time.

How does your organization incorporate inclusive design?

n=2,763 responses from 1,887 respondents



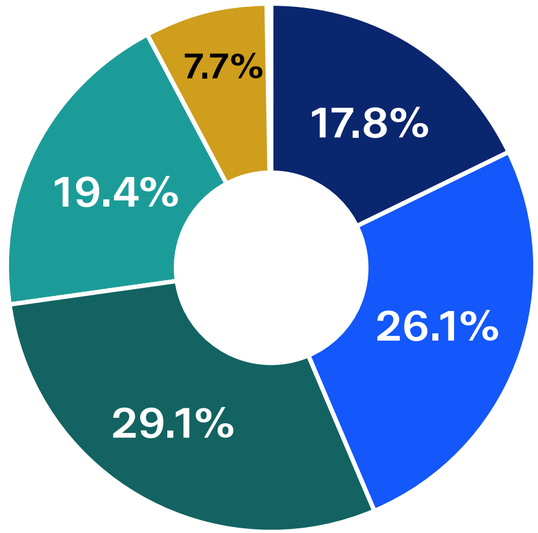
How would you rate your current understanding of digital accessibility?



Despite increasing understanding of digital accessibility and a desire to do better, many organizations still lack adequate resources.

How well-equipped is your company to test for accessibility without external help?

n=2,608



17.8% No in-house expertise/ resources

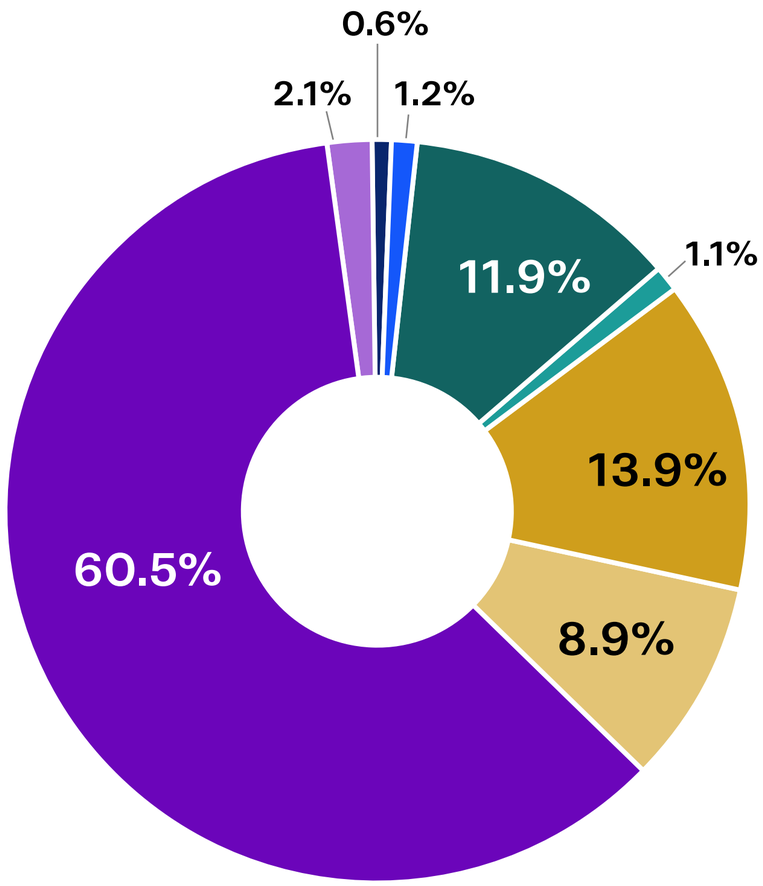
26.1% Limited in-house expertise/ resources

29.1% Some in-house expertise/ resources, but could use more

19.4% Adequate in-house expertise/ resources

7.7% Plenty of in-house expertise/ resources

Accessibility issues Applause reported n=26,479



0.6% Animation & video

1.2% Color

11.9% Color contrast

1.1% HTML validator

13.9% Keyboard navigation

8.9% Other A11y

60.5% Screen readers

2.1% Zoom

For the third year in a row, screen reader errors are the most common accessibility issue, by a large margin. There seems to be a knowledge gap in how to consistently create code that works well with screen readers.

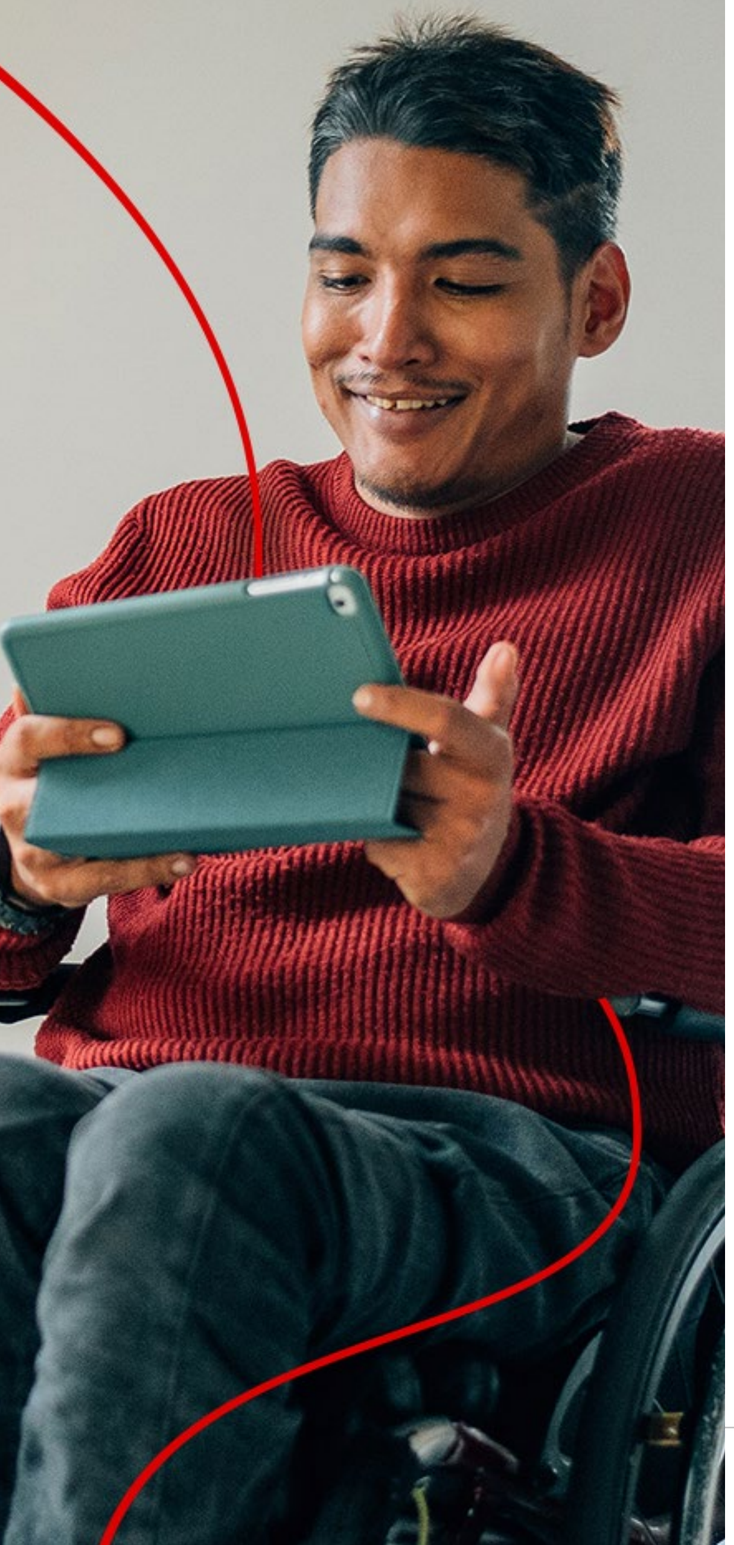
There's another gap in preparing to comply with the EAA. When surveyed less than 15 months before the deadline, more than a quarter of respondents said they had not begun to prepare, and another 32% indicated that they were behind schedule.

Is your company preparing for the EAA ahead of the June 2025 compliance deadline? n=1,531

39.7%
Yes – we are on schedule

32.0%
Somewhat – we are a bit behind schedule

28.2%
No – we have not started



Accessibility and inclusive design recommendations

Start with the foundations that shape your websites and apps: your design system/ UI kit. Focusing on the elements that are most commonly used across your apps and websites delivers an immediate boost to usability and inclusivity. If the design system lacks built–in accessibility, that cascades across all the forms, templates and pages that incorporate those elements. Starting with the design system and UI kit helps scale accessibility from the earliest stages of design and development.

Include PWD and their insights from the earliest stages of design and development to ensure the best user experience possible. Make sure that your staff understands how your apps and sites work with different adaptive technologies for users with different disabilities, and how those may combine to create a variety of different disability profiles. For example, how does the experience for blind users change when they also have mobility limitations or hearing loss? Build partnerships and empathy with PWD to create better user experiences for everyone.

Make sure your organization has adequate expertise and resources. While most organizations have a team or person responsible for digital accessibility, inclusivity and conformance must scale across all products and websites. Invest in ongoing training and testing for accessibility and inclusivity to reduce risk and ensure the organization is prepared to conform to WCAG standards and comply with applicable laws, like the EAA.

UX testing

It's not enough for websites and apps to function: they must do so in ways that users expect and enjoy. Attracting and retaining customers is hard enough without having to recapture those who have been driven off by poor user experiences. Despite the ever-increasing amount of data organizations have at their fingertips, they often fail to address customer concerns or perform root cause analysis on common points of abandonment. With so many users willing to abandon an app at the first sign of friction, it's imperative for teams to have a strategy for delivering the best possible UX for the largest possible pool of users.

As organizations look to speed deployment, automate more and capitalize on AI's potential, they must remember that human feedback still plays a crucial role throughout the development process. From generative and exploratory research at the earliest stages of design to ongoing evaluative studies that drive product iteration, organizations must capture insight on the real user experience and identify ways to improve.

UX testing framework

Digital quality emergence:

Lack of formal systems, processes and documentation, no consistent methodology or approach to UX testing.

Examples of testing activities and processes:

- Dogfooding; testing with friends and family
- Guerilla usability testing
- Reading customer reviews
- Collecting user feedback (e.g., intercept surveys, etc.), without analyzing or synthesizing the data

Digital quality essentials:

Early stages of defining and documenting processes and procedures; establishing some consistency and structure around UX testing.

Examples of testing activities and processes:

- Regularly reviewing data from intercept surveys or similar sources
- Conducting annual heuristic (expert) evaluations
- Identifying 1–2 key user groups based off of initial trends noticed in data
- Conducting user research without a clearly established plan to act upon the findings
- Conducting 1–2 light usability evaluations each year

Digital quality expansion:

The organization has a clear process in place to assess the most common customer journeys, including digital, physical and hybrid interactions. Some reporting is in place. Focus on coverage, scalability and efficiency across the organization.

Examples of testing activities and processes:

- Gathering data in a variety of ways to create user personas and archetypes
- Regularly conducting a variety of UX studies at various points in the development lifecycle
- Conducting iterative usability testing to ensure any changes lead to stronger user experience
- Establishing strong relationships with key stakeholders to take action on UX study findings

Digital quality excellence:

The organization thoroughly tests all customer journeys. Detailed reporting is in place. Teams relentlessly seek out and eliminate friction and understand the variations in preferences across markets.

Examples of testing activities and processes:

- Regularly using personas throughout the design process
- Conducting generative research before creating products or designs
- Conducting iterative testing across all platforms, devices and designs
- Employing both qualitative and quantitative methodologies to uncover what is happening, why is it happening, and how the design should be improved
- Building tight partnerships across the product, design, development and research teams

Case Study

Scoring better with users

Like many of its contemporaries, the Austrian social gaming company Funstage used to rely on qualitative user surveys to gather UX feedback. Beyond giving a vague idea of user sentiment, this did little to answer business-critical questions, like why users always dropped out of a game at a specific point.

Today, through detailed documentation in the form of user videos and written analysis, Applause helps Funstage see exactly where users encounter issues. And, instead of sorting through thousands of survey responses, Funstage counts on Applause's UX expert to filter and prioritize user feedback and provide actionable recommendations.

Localization and payments are increasingly becoming part of the UX equation, yet fewer than 40% of our survey respondents reported that they typically test these elements. Without attention to these crucial aspects of the customer journey, brands leave revenue on the table.

What types of tests does your organization conduct? n=894

27.8%
Customer journey testing

27.6%
Localization testing

37.7%
Payments

54.4%
User acceptance testing

65.0%
User experience

UX recommendations

Collaborate on the best ways to act on insights and user feedback. Bring designers, developers and UX researchers together regularly to review UX test results and user feedback as a team, then brainstorm and implement improvements. Taking an interdisciplinary approach to evaluating and acting on feedback allows for teams to better understand priorities and to work together to come up with ways to enhance UX.

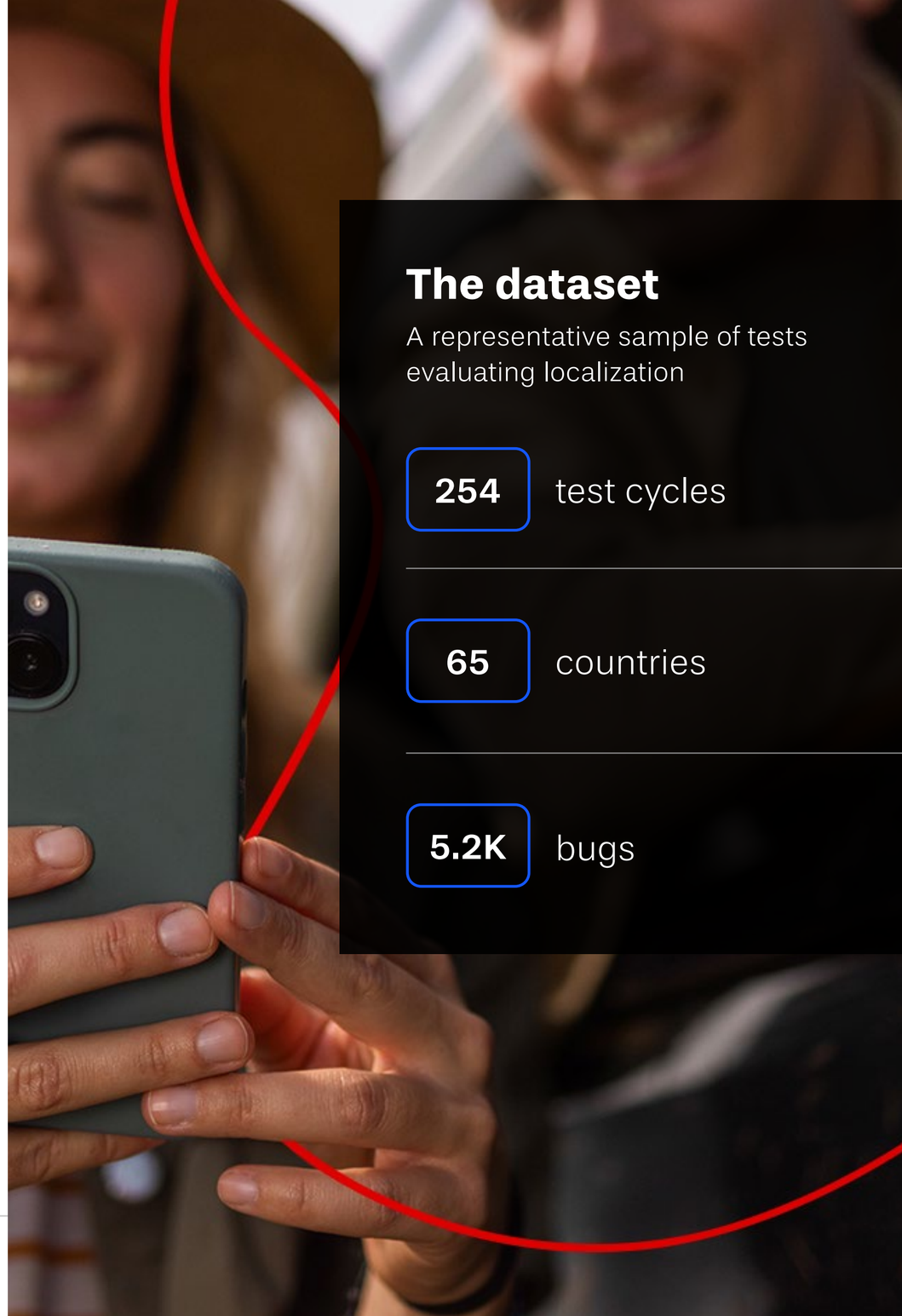
Create detailed user personas and use them for scenario-based testing. Develop detailed user personas representing different types of users or customers, based on data collected in generative and evaluative research studies — make sure the personas accurately reflect your users and their needs. Use these personas to guide UX testing and ensure your digital experiences meet the needs of various groups. Test specific tasks and workflows within the software from the perspectives of those personas to understand their unique wants and needs.

Implement continuous feedback mechanisms. Integrate in-app tools that allow users to provide real-time feedback directly within the software. Conduct regular user studies to gather ongoing feedback about the user experience and identify areas for improvement. Incorporate both quantitative and qualitative data, and measure changes over time.

Localization testing

Accurate localization can help attract and keep customers. While many users are willing to try apps in languages other than their native one, they expect accurate translations and culturally relevant offerings. In a survey of more than 4,800 consumers, 24% of respondents reported that they would never use an app in a foreign language or only do so if absolutely necessary.

Another 45% stated they do not mind using foreign-language apps as long as the content provides value. While 74% of respondents said they had used a foreign-language app within the last year, users will abandon apps that don't live up to expectations: 52% reported that they have abandoned an app or website due to poor translations or inaccurate localization.



The dataset

A representative sample of tests evaluating localization

254 test cycles

65 countries

5.2K bugs

Localization testing framework

Digital quality emergence:

Lack of formal systems, processes and documentation; no consistent methodology or approach to translation or localization.

Examples of testing and processes:

- Translating some high-priority content without validating that translations are contextually correct

Digital quality essentials:

Early stages of defining and documenting processes and procedures; establishing some consistency and structure around localization.

Examples of testing activities and processes:

- Ensuring all content that should be localized is accurately translated
- Verifying that all currencies, dates, symbols and measurements are converted to the appropriate units and formats
- Validating that forms work correctly
- Localizing content for the business's top markets

Digital quality expansion:

The organization has a clear process for ensuring accurate localization and uses various testing types to validate. Some reporting is in place. Focus on coverage, scalability and efficiency.

Examples of testing activities and processes:

- Using native speakers in market to validate translations and idioms
- Ensuring that visual elements fit translations
- Validating that imagery and colors are culturally appropriate
- Assessing workflows to ensure they align with expected processes
- Putting a process in place to resolve conflicts and update assets
- Conducting pre-production validation for strings
- Documenting preferred translations for key words and phrases in a glossary
- Using a minimum of double-blind support to validate translations, not including the content creator
- Localizing applications for all markets where the business operates

Digital quality excellence:

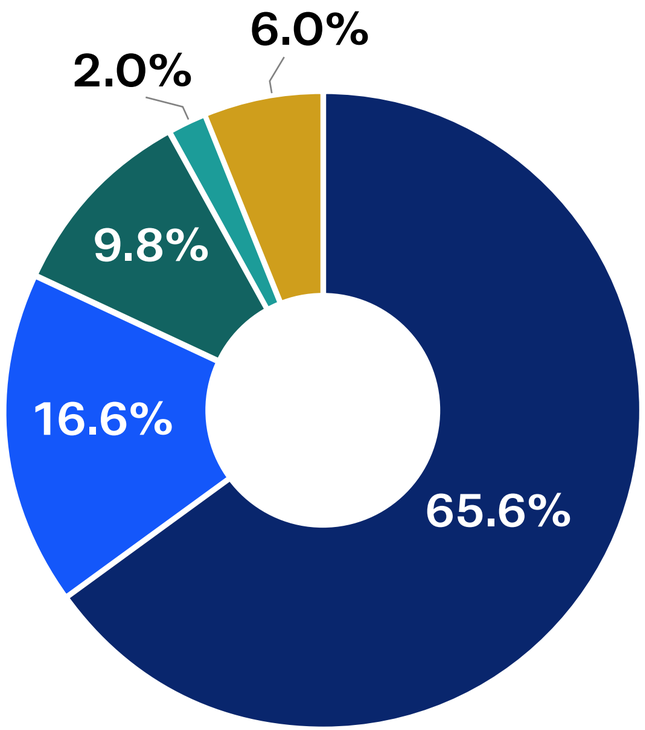
Awareness of cultural differences and commitment to respecting the norms in different markets serves as a competitive differentiator.

Examples of testing activities and processes:

- Leveraging cultural values and norms to create relevant product features and offerings
- Differentiating between distinct dialects, such as Portuguese vs. Brazilian Portuguese
- Providing a channel for customers to report any issues
- Having a process to manage issues
- Assessing how customer journeys vary across markets
- Factoring in accessibility concerns, such as whether screen readers work for right-to-left languages
- Demonstrating geopolitical awareness in sensitive areas; correcting errors that can damage earnings or reputation
- Committing to localization in native languages; factoring in regional dialects and differences to develop a solution that works for all stakeholders

In a survey of more than 1,900 software development, product and QA professionals, 64% indicated that their apps and websites are available in multiple languages.

In how many different languages is your application available? n=899



65.6% 2–5

16.6% 6–9

9.8% 10–24

2.0% 25–49

6.0% 50+

More than half of those who make apps available in multiple languages do all the development in-house.

Do you develop localized sites/apps in-house? n=900

54.1%
Yes, we develop all of our localized sites/apps in-house

22.2%
We develop some in-house and outsource some, depending on the language

15.6%
No, we outsource development for all localized sites/apps

10.0%
I don't know

How do you validate that applications are properly localized? n=896

41.3%
We use automated internationalization (i18n) testing tools

59.0%
We have developers and QA staff who are native speakers for each language

24.0%
We hire a translation agency

27.6%
We beta test with in-market customers

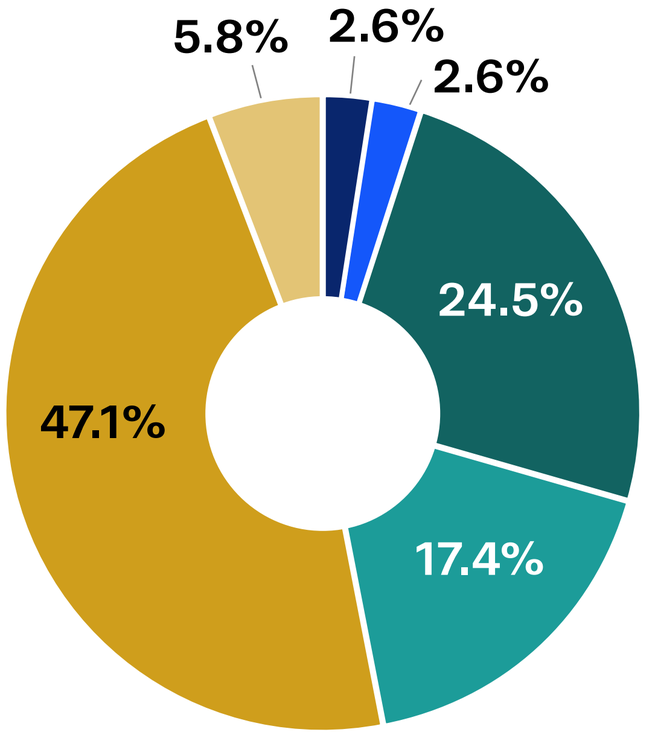
27.0%
We use a third-party testing partner

4.1%
Other

To maintain accuracy and consistency, 70% report that their organization maintains a glossary of preferred translations for key words and phrases used in their apps. In addition, 66% of 1,394 respondents reported that they adjust their apps' customer journeys and checkout flows to align with expectations in different countries and cultures.

Localization issues Applause reported

n=5,226



2.6% Corrupted characters

2.6% Currency and number format

24.5% Missing translation

17.4% Other/general

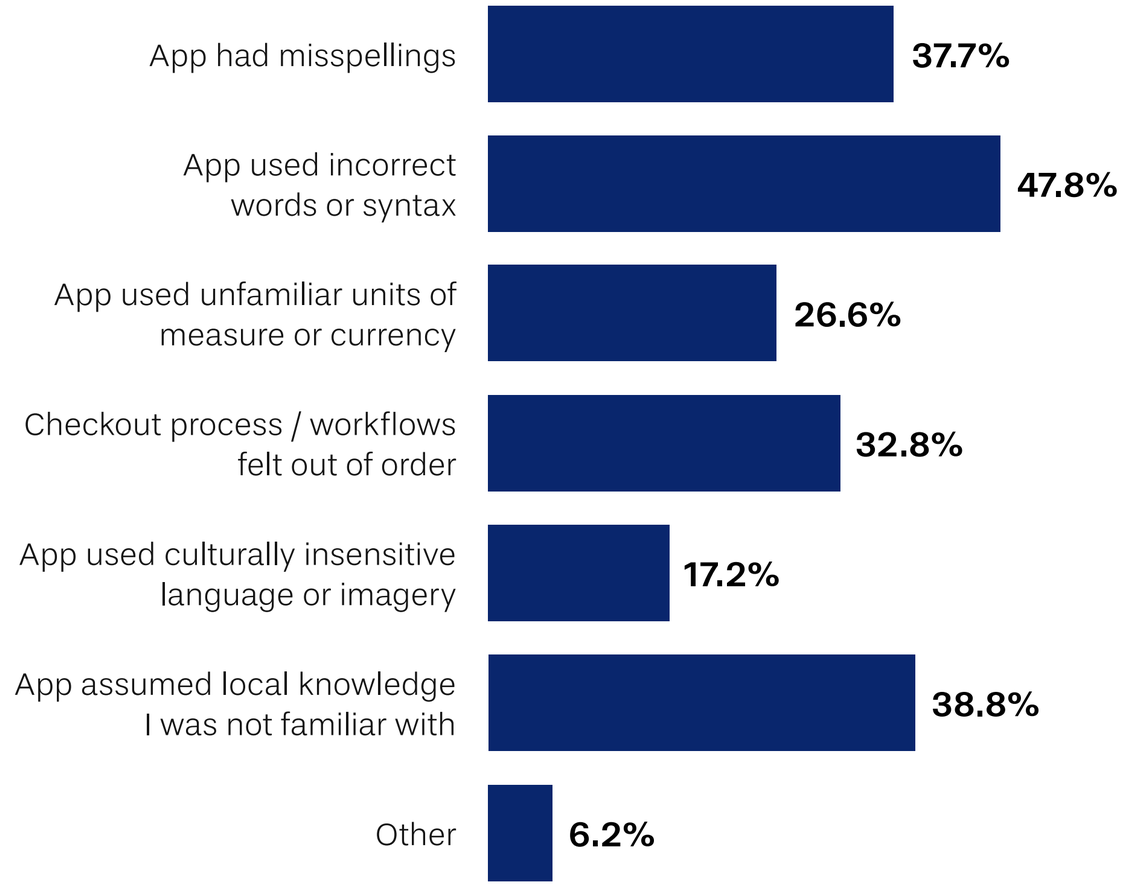
47.1% Poor translation

5.8% Truncation and overlap

Consumers who had abandoned apps or websites due to poor translation and localization described the issues that led them to give up on the experience.

What types of issues led you to abandon the app or site?

n=2,402



Case Study

Winning love at first sight

When a company wanted to roll out its dating app in new countries, its small internal QA team needed to understand user expectations and ensure culturally appropriate experiences in the new markets.

Applause community members from each of the four target countries provided feedback on the onboarding process and offered unique culturally-based insights and recommendations, such as not asking about religious preference in Turkey.

As a result, the company was able to avoid errors that would have been immediate turn-offs for its new customers and win them over to the app.

Localization testing recommendations

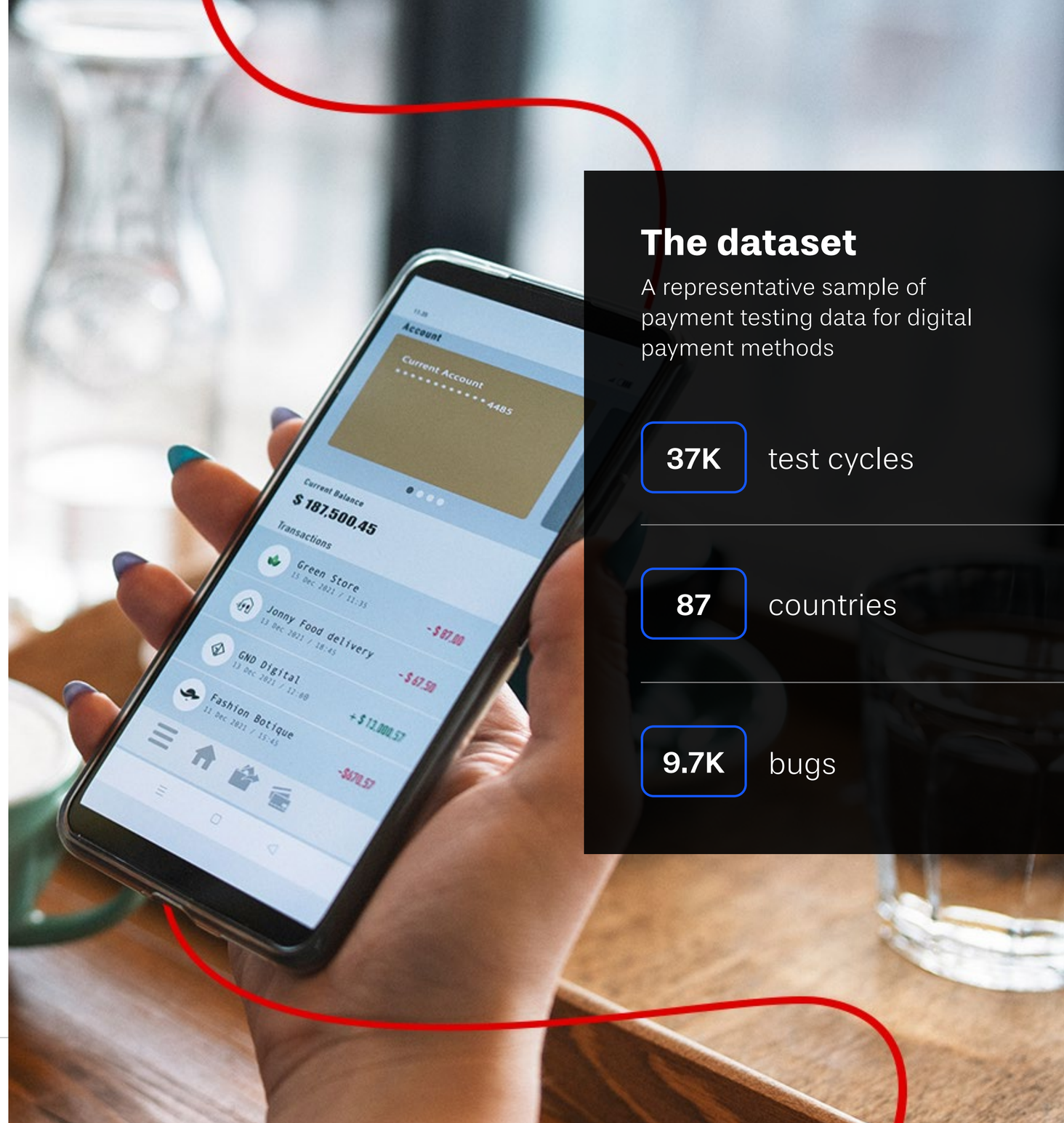
Take the time to translate, then validate spelling and syntax. The most common localization defects Applause testers uncovered were poor translations, followed by missing ones. Attention to these details is crucial to establish trust with customers in different markets and to create seamless user experiences.

Learn what the locals expect and adjust your customer journeys accordingly. With a third of consumers reporting they abandoned an app because a sign-up process or checkout flow felt out of order, it's important to align with the expectations in different markets. There's no substitute for feedback from local, in-market participants that match your customer profiles.

Payment testing

While it seems like payment testing would build a case for itself – if you can't accept payments, you can't make money – it's often overlooked. And, even organizations that validate functionality for payments often only partially test various payment instruments or other types of transactions.

Finding and fixing a single critical payment blocker before it reaches production can have a significant impact on the bottom line. Taking the time to thoroughly test payment workflows and instruments pays off.



The dataset

A representative sample of payment testing data for digital payment methods

37K test cycles

87 countries

9.7K bugs

Payment testing framework

Digital quality emergence:

Lack of formal systems, processes and documentation, no consistent methodology or approach to payment testing.

Examples of testing activities and processes:

- Testing with dummy cards and accounts
- Dogfooding; testing with friends and family

Digital quality essentials:

Early stages of defining and documenting processes and procedures; establishing some consistency and structure around payment testing.

Examples of testing activities and processes:

- Validating that transactions work using live payment instruments: testing both purchases and returns
- Defining a payment instrument coverage matrix, including at least the 5–10 most popular payment instruments in market
- Testing when changing payment gateways; validating code for new gateways or payment processors
- Verifying code for launches in new countries

Digital quality expansion:

The organization has a clear process in place to ensure the full range of payment activities work properly and uses various testing types to assess functionality and UX. Some reporting is in place. Focus on coverage, scalability and efficiency across the organization.

Examples of testing activities and processes:

- Creating a payment instrument coverage matrix based on data about website/app usage, local user preferences
- Validating new payment methods prior to launch
- Collecting and acting on UX feedback
- Conducting payment instrument regression testing for all features, including product returns and customer service interactions
- Testing payments using specific devices, banks and BIN combinations
- Sharing card statements upon request

Digital quality excellence:

The organization continuously and comprehensively tests end-to-end payment workflows and UX, including a wide variety of payment instruments, and understands the variations in preferred payment instruments and workflows across markets.

Examples of testing activities and processes:

- Conducting ongoing testing for loyalty programs, promotional offers, and options like BNPL
- Testing negative/unavailable payment instruments (soft or hard decline, blocked, expired)
- Performing wallet cycling testing with multiple payments in different balance states
- Live testing one-of-a-kind payments which require holder presence (PIX, Blink, Bancontact, iDEAL)
- Verifying purchase states live directly with the bank for card issuers, card networks and fintechs
- Verifying purchase results live directly with the bank for merchants

In how many countries does your organization go to market? n=1,933

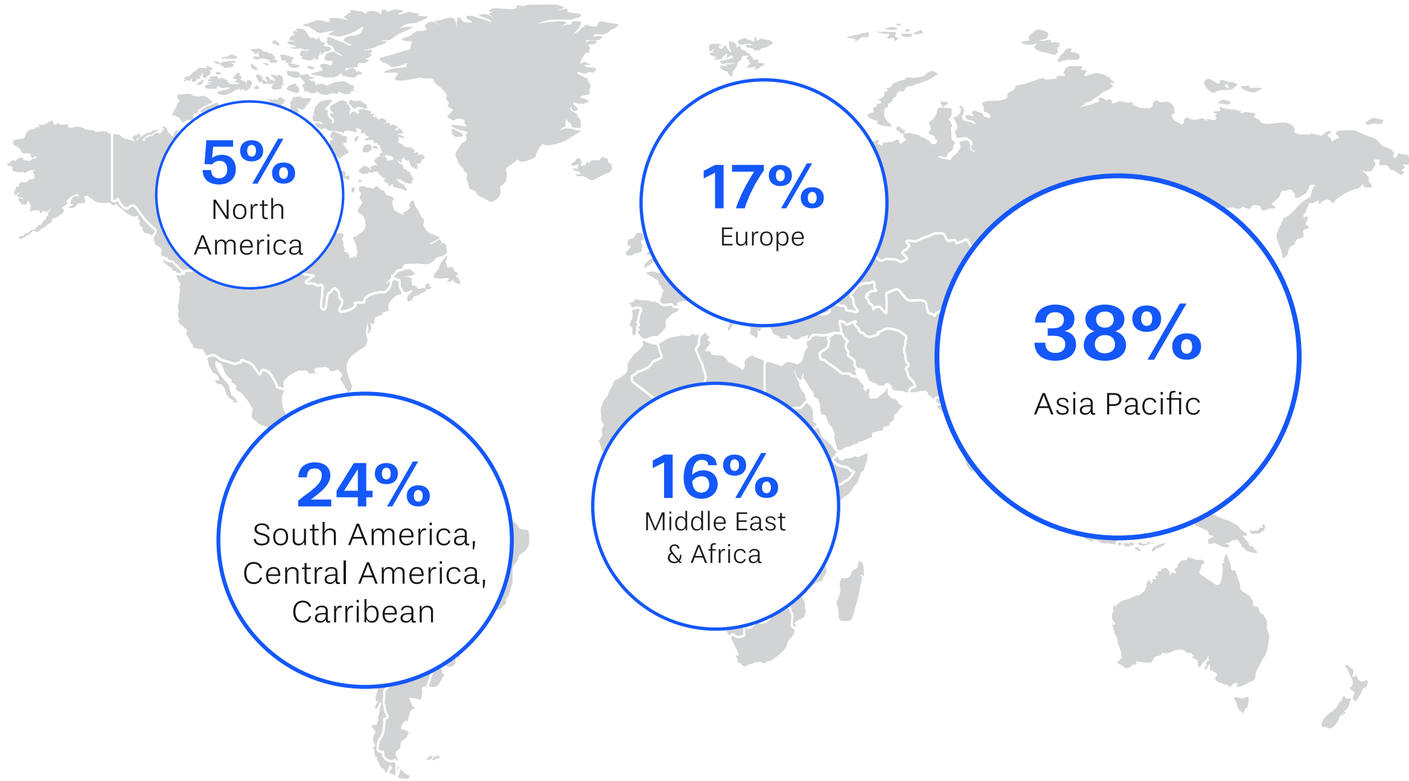
66.8%
1–9 countries

18.4%
10–49 countries

14.8%
50+ countries

For organizations to succeed, they need to be able to capture payments. Digital payments are becoming the norm. No longer reserved for websites or apps, digital payments are increasingly making their ways into storefronts, restaurants and other physical environments. As companies roll out new payment options, whether it's QR codes on restaurant checks or launching an app in a new region, testing is essential.

In which regions does your organization go to market? n=1,937

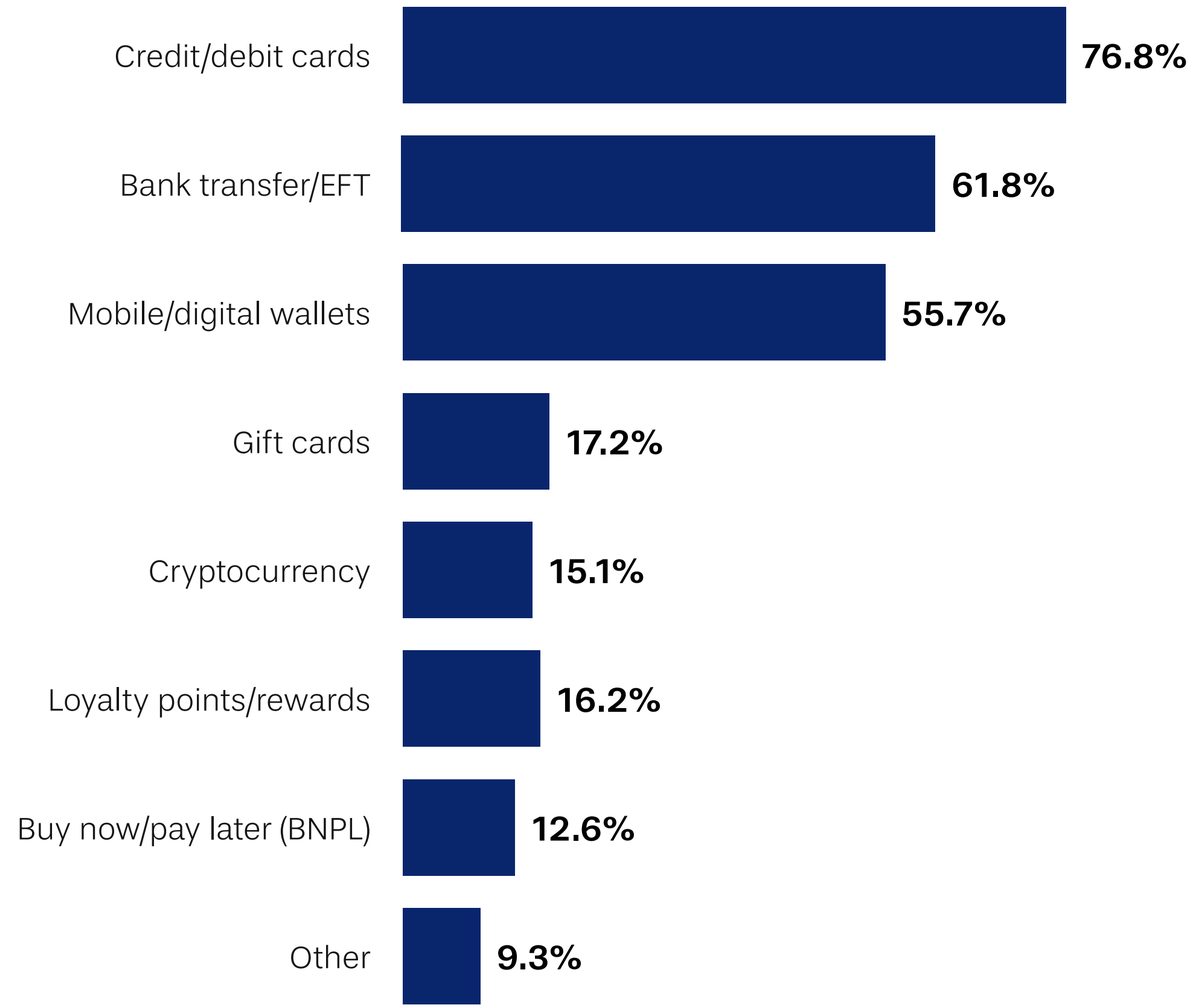


Case Study

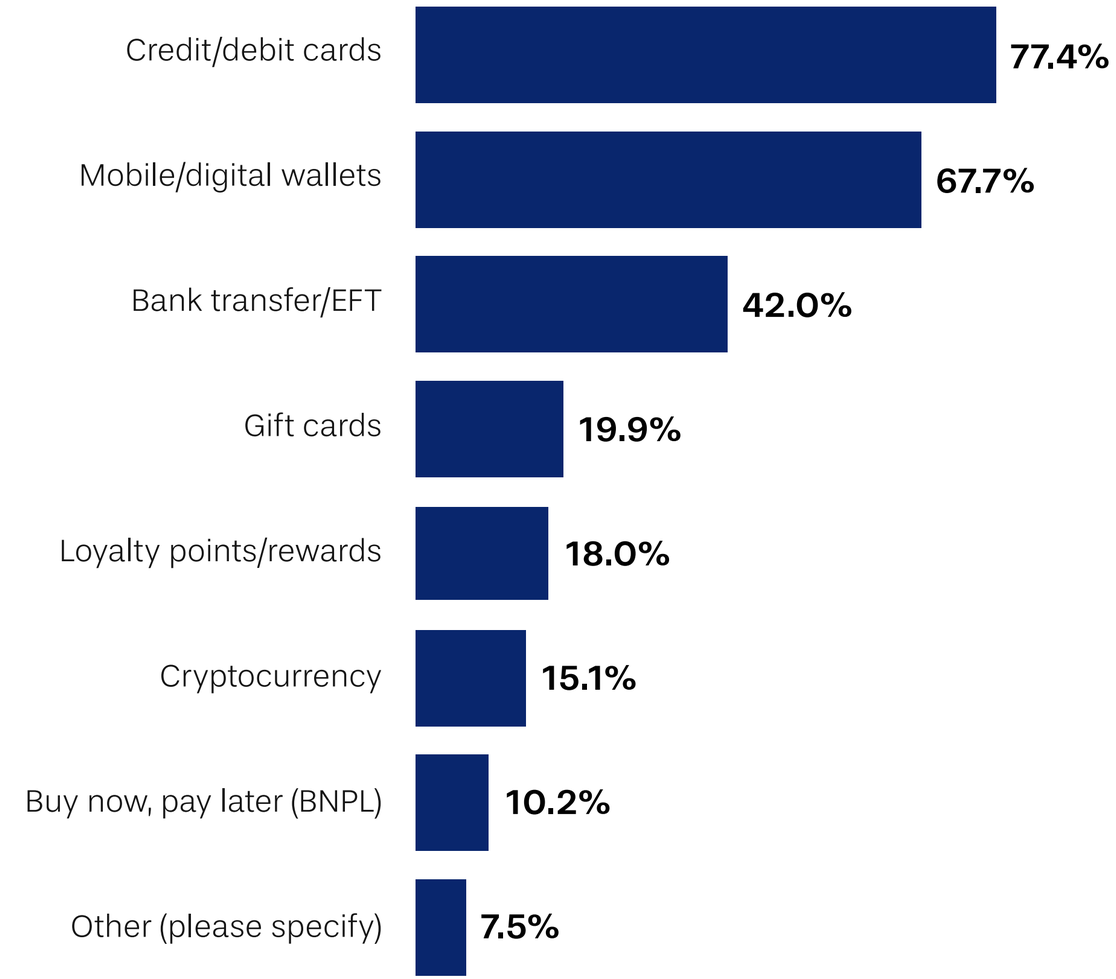
Protecting sales with payment testing

One global athletics brand worked with Applause to test payments across 58 countries. Testers identified 12 payment bugs with an estimated value of $1.1M per critical blocker. The company avoided $13.3M in lost revenue by addressing just those 12 defects.

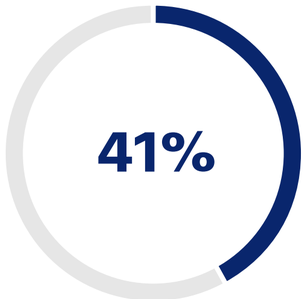
Which payment options are available through your app? n= 1,386



Which of these payment methods do you use online? Choose all that apply. n=4,694

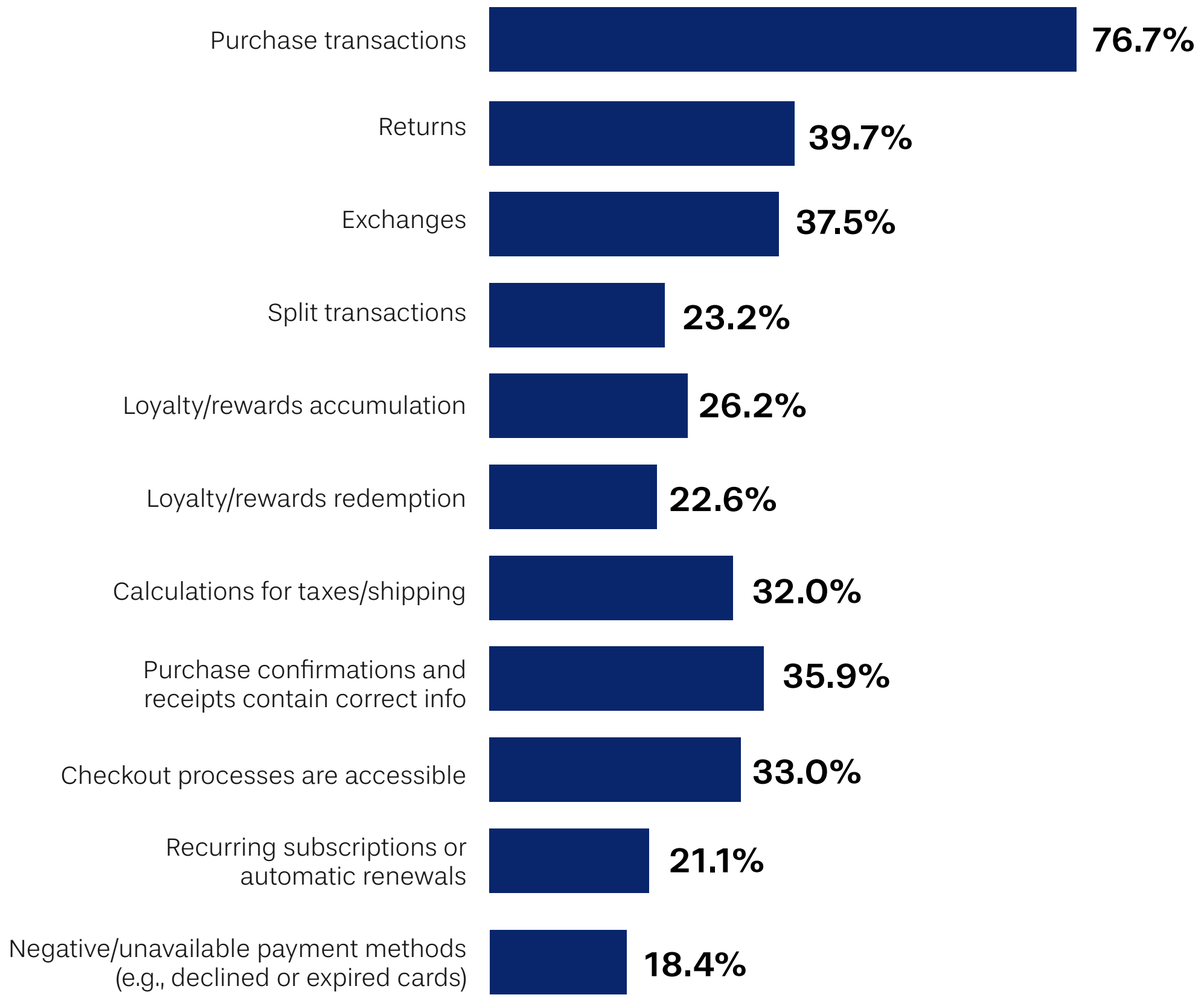


In Applause's consumer survey, 41% of respondents said they were extremely likely to abandon a purchase if a website or app did not accept their preferred payment method. Another 35% said they were somewhat likely to abandon; only 4% said they were extremely unlikely to do so. Yet, only 62% of the software development and QA professionals surveyed reported that customer preference influences the payment methods their organizations accept.

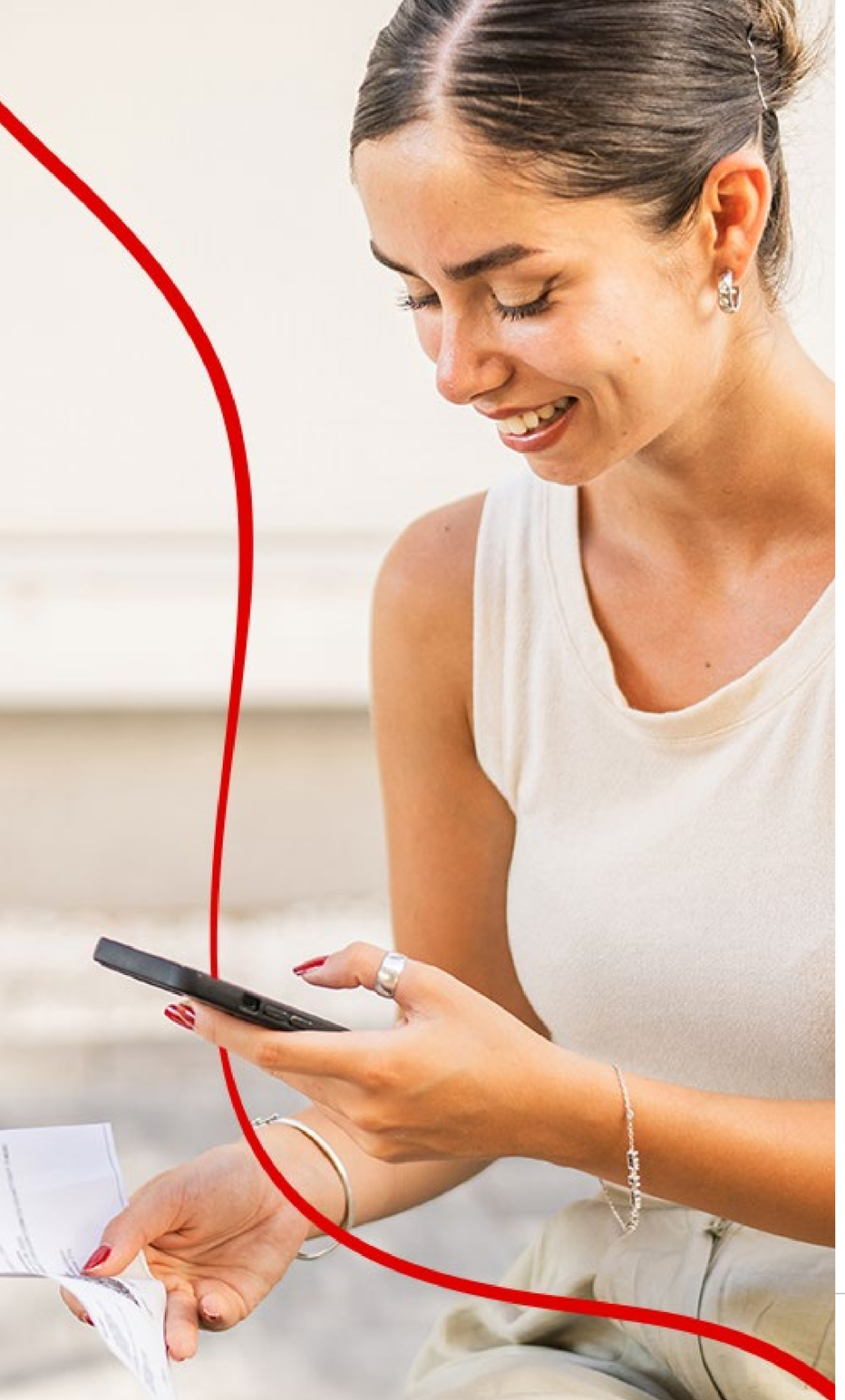


41% of respondents said they were extremely likely to abandon a purchase if a website or app did not accept their preferred payment method

Which elements of the payment process does your organization test or evaluate? Choose all that apply. n=1,338



Failure to test processes like rewards accumulation or redemption can be costly — in both financial terms and customer satisfaction. The same holds true for validating shipping and tax calculations: too much and you'll anger customers, too little and you cut into profits and potentially accrue penalties. Even something as simple as validating the way a transaction shows up on a statement can have an impact — a customer who doesn't recognize a charge due to an unfamiliar vendor name may dispute it with their bank or credit card provider, or cancel a subscription moving forward.



Payment testing recommendations

Test all aspects of the payment process, not just purchases. End-to-end payment testing across all parts of the transaction is the only way to detect issues that can eat away at margins. Returns, exchanges, recurring subscriptions and renewals, autotransfers, loyalty programs: test all the different ways funds and credit change hands. It's the only way to ensure every transaction succeeds, every time.

Test a broad range of payment instruments across different devices to uncover functional flaws that block payments. Whether it's the inability to accept a specific payment method or a visual error that fails to display a card entry form on certain devices, preventing these errors from reaching customers is crucial.

Collect customer feedback on their preferred payment methods in different markets. With more than three-quarters of consumers likely to abandon a purchase if a website or app doesn't accept their preferred payment method, it's imperative to provide the options customers want, especially as providers change and new payment methods enter the scene. Preferred payment methods can vary widely from one country to another, so make sure you understand local preferences.

AI training and testing

The meteoric rise of AI – and generative AI in particular – poses new opportunities and complexities for consumers and organizations alike. As the number of users and use cases increases, so does the need for thorough testing and risk mitigation. While consumer adoption has soared and businesses are looking to capitalize on AI's potential to increase efficiency, many people remain skeptical about the technology, raising concerns about inaccuracy, bias, safety, security and other risk factors.

To reduce risk and develop AI apps that users trust – and that comply with ever-evolving legislation – organizations must adopt a rigorous testing strategy and demonstrate transparency around data collection and model decisions.

AI testing framework

Digital quality emergence:

Lack of formal systems, processes and documentation, no consistent methodology or approach to testing AI.

Examples of testing activities and processes:

- Dogfooding; testing with friends and family
- Performing basic data quality checks to ensure training data is clean and accurate
- Conducting simple tests to identify any glaring bias or fairness issues in the model

Digital quality essentials:

Early stages of defining and documenting processes and procedures; establishing some consistency and structure around testing AI.

Examples of testing activities and processes:

- Assessing bias and fairness across multiple dimensions, including different demographic and contextual factors
- Defining and monitoring critical KPIs for the application to measure model performance
- Maintaining documentation and version control: keeping detailed records of data sources, model parameters, and configuration settings to ensure reproducibility and traceability
- Executing basic error analysis and troubleshooting

Digital quality expansion:

The organization has a clear process in place to ensure the full range of AI interactions work properly and uses various testing types to assess functionality and UX. Some reporting is in place. Focus on coverage, scalability and efficiency across the organization.

Examples of testing activities and processes:

- Systematically collecting and acting on UX feedback to improve the user experience
- Conducting regular data validation to ensure ongoing data integrity and quality
- Testing for model explainability and transparency to ensure the model's decisions can be understood and justified
- Using data from KPIs to improve the application's performance over time, iteratively enhancing accuracy and efficiency based on KPIs

Digital quality excellence:

The organization thoroughly tests end-to-end workflows and UX for AI experiences, including a wide variety of inputs and prompts, and understands the variations in preferences across markets.

Examples of testing activities and processes:

- Conducting comprehensive end-to-end testing from data ingestion to final output to ensure consistency and accuracy
- Performing advanced, ongoing red teaming exercises to proactively identify and mitigate vulnerabilities
- Ensuring rigorous, comprehensive data validation to maintain high standards of quality and reliability, including regular audits and compliance checks
- Maintaining extensive testing documentation and sharing knowledge and learnings internally
- Continuing to retrain models over time, updating them using live data
- Executing real-time monitoring to detect and correct anomalies as they arise

In 2024, Applause conducted a survey on generative AI use: 6,361 consumers, software developers and QA professionals participated in the survey, revealing that while user satisfaction is increasing, opportunities for improvement still exist.

More than 2,900 respondents indicated they use at least one Gen AI chatbot daily. In addition, 37.5% of respondents reported that they use different chatbots for different tasks, while 26.5% stated that they have swapped one service for another, typically due to performance issues.

For most users, chatbots replace existing search engines and research tools — 91% of respondents have used chatbots to conduct research, and 33% of those respondents use Gen AI for research daily. For 81% of survey participants, chatbots have replaced search engines for queries; 32% of those use chatbots for search daily. Other popular use cases include language translation, creative writing and drafting emails, proposals or similar business communications.

Though the number of users who have encountered responses that are biased, offensive or contain hallucinations has grown slightly over last year, more survey respondents reported that Gen AI understood their prompts than last year.

The most common Gen AI UX errors n=4,174

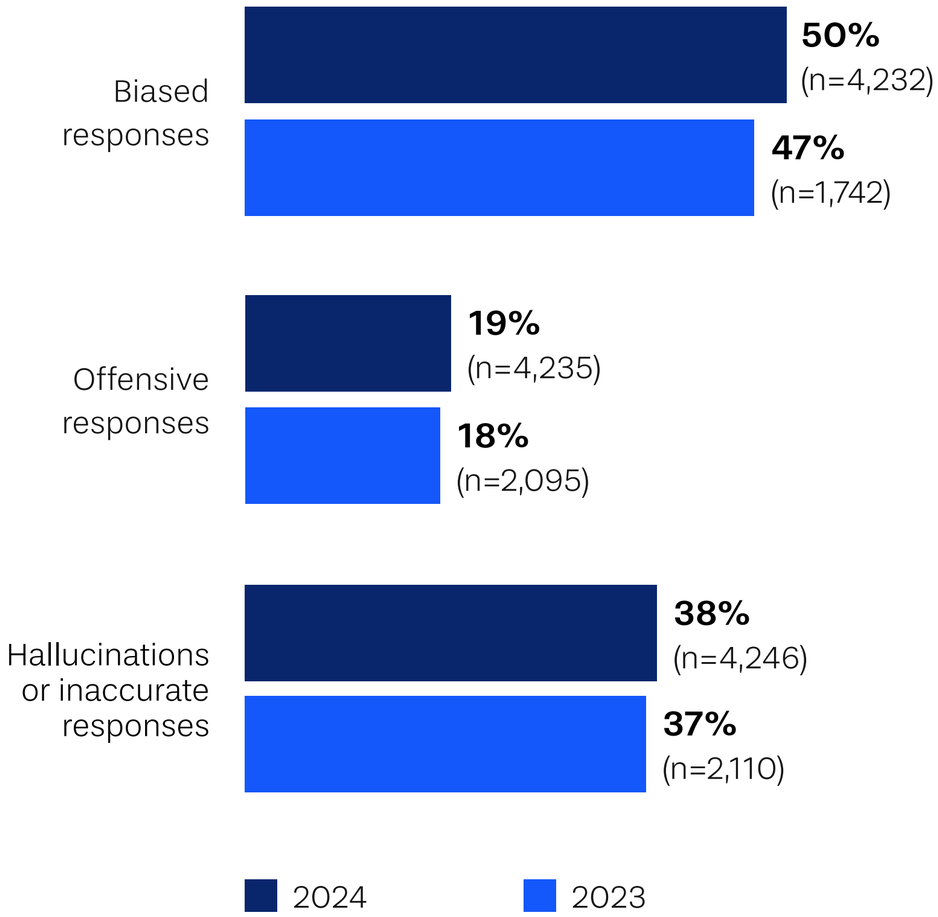
17.5%
Gave a general answer that did not provide enough detail

10.7%
Gave a convincing but slightly incorrect answer

16.8%
Misunderstood my prompt

10.3%
Generated obviously wrong answers

Have you experienced any of the following when using generative AI?



Gen AI is changing software development and testing

Gen AI has myriad use cases in the software development and testing process. We asked survey respondents who indicated they are using chatbots to write or debug code, build test cases, or for test reporting to expand on how they are using chatbots for testing.

Top use cases for Gen AI in testing n=1,548

60.9%
Test case generation

53.7%
Text generation for test data

52.3%
Test reporting

Case Study

Harnessing human feedback with prompt response grading

A leading financial services firm sought to enhance its AI chatbot, built on an open-source large language model and trained on proprietary data. It worked with Applause to fine-tune the chatbot's responses and align them with user needs through reinforcement learning from human feedback (RLHF).

A diverse testing team evaluated thousands of chatbot responses weekly, grading accuracy and flagging content such as bias, toxicity and inaccuracies. This enabled the company to identify and fix specific weaknesses in the model, significantly reducing safety concerns and improving user satisfaction.

AI testing recommendations

Use red teaming to identify vulnerabilities and mitigate risk. Assemble testing teams that include diverse perspectives and deep domain expertise to reduce the risk of bias and inaccuracy. Effective red teaming can also improve UX and scale testing by providing both depth and breadth.

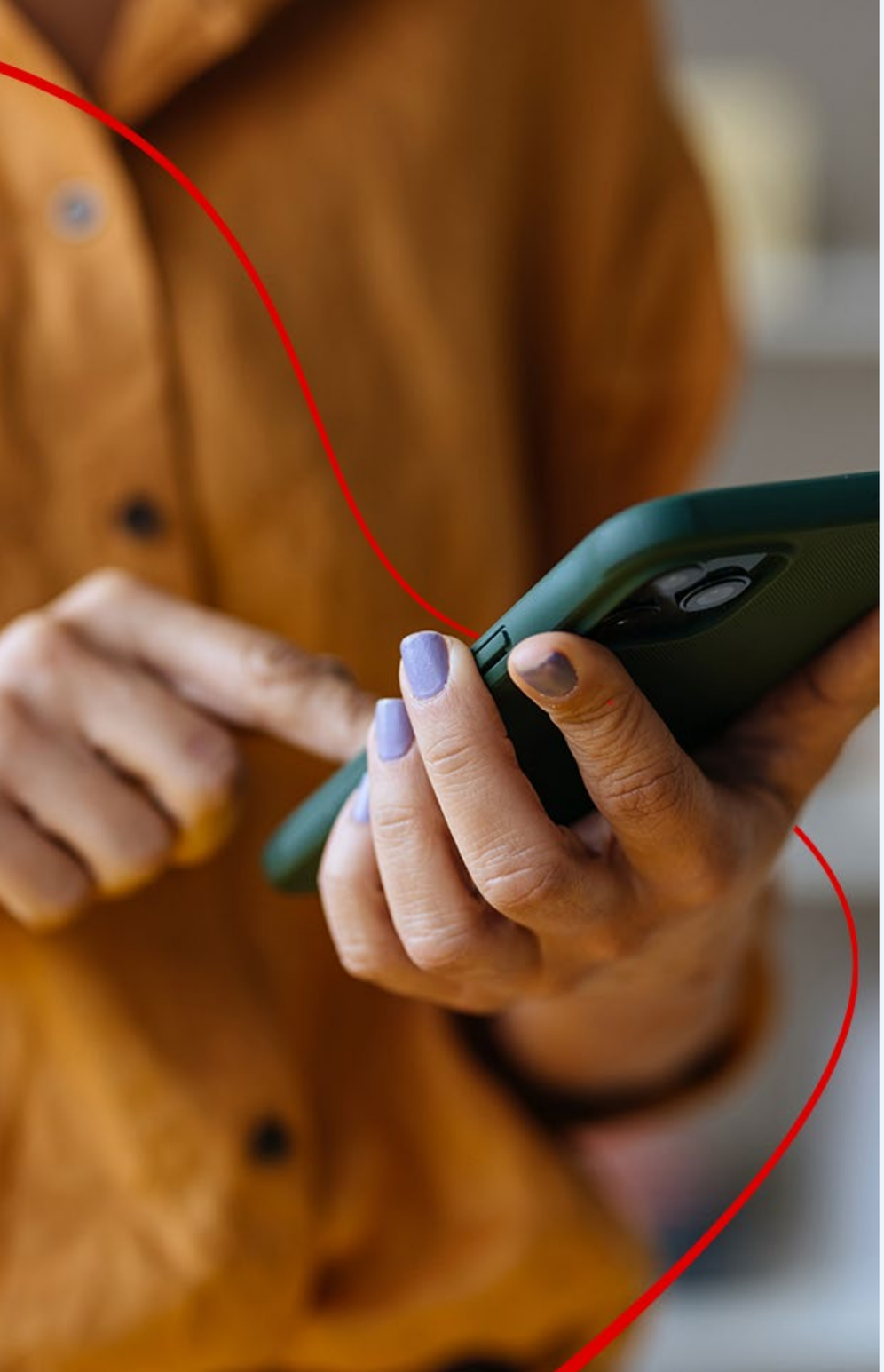
Integrate human feedback throughout the AI model development lifecycle. Nuanced human feedback is essential while training models and in evaluating model behavior. Human oversight is also required to ensure AI responses meet user expectations and adhere to legal, regulatory and ethical requirements. A programmatic approach to capturing and responding to ongoing and iterative human feedback allows organizations to fine-tune models to improve quality and mitigate risks.

Collect diverse training data that covers real-world use scenarios. To reduce the risk of bias, inaccuracy or poor user experience, it's crucial to start with high-quality data that's fit for purpose. Off-the-shelf datasets from a data warehouse can offer a starting point, but can't provide the specific details and level of nuance that AI requires.

What's next in digital quality

Whether your organization has well-established quality systems or you're just starting out, we hope you'll find insights in this report that help you improve and innovate. Use this report to identify your organization's current capabilities and gaps, and leverage your internal data to pinpoint which areas of focus will yield the greatest boosts to quality.

As AI adoption takes off, users develop higher expectations for apps across industries, and since technologies constantly evolve, it's critical to invest in the activities and infrastructure that allow your organization to maintain quality at speed and scale.



What does digital quality mean to you?

Where does your organization fall on the various frameworks? Which trends are you monitoring most closely? What are your secret weapons for creating apps users can't live without? We'd love to learn more about how you define, measure and pursue exceptional digital quality.

View additional materials at **StateOfDigitalQuality.com**.

Contact us at **feedback@stateofdigitalquality.com**.

Acknowledgments

Thank you to the business leaders, community members and subject matter experts who shared their time, knowledge and insight to shape this report.

Report Authors
Jenn Waltner, David Carty, Paul Hoffman, Annabel van Daalen

Creative Design:
Joe Stella, Karley Searles, Megan Gawlik, Samsudeen Sallah

About Applause

Applause is the world leader in digital quality – built by innovators, powered by people and dedicated to the comprehensive digital testing and feedback needs of our global enterprise customers. Our fully managed solutions harness a powerful combination of community-based testing and advanced technology to ensure organizations can move quickly to release apps, devices and experiences that are consistently functional, intuitive and inclusive in any market. Our experts steward you through the entire testing process, from strategy to execution, seamlessly supplementing your existing resources. We deliver actionable, real-time insights early in the development process that ensure projects adhere to roadmaps and launch successfully, driving customer retention and revenue. We're proud to be an essential partner to the most innovative names in the digital economy, as we work to ensure technology works for everyone, everywhere. Learn more: **applause.com**

APPLAUSE