resight global

The Automotive Experience Score (AXS)

Whitepaper

Jan Panhoff, Wolfgang Waxenberger, et al., ReSight Global GmbH

Table of contents

Abstract

In the ever-evolving automotive sector, there is a pressing need for a standardized yet flexible framework to consistently evaluate the user experience (UX) of interactions with cars across various models. The Automotive Experience Score (AXS) meets this demand by providing a pioneering approach to benchmarking UX, integrating real user interactions from a broad spectrum of vehicles. This method consolidates the UX assessment process, addressing the industry's fragmentation and the inconsistency of individual original equipment manufacturers (OEMs) studies. Unique in its ability to compare OEM outputs directly, the AXS combines subjective assessments, such as the UX Score and the After Scenario Questionnaire (ASQ), with objective metrics like task completion rates and driver distraction evaluations.

Roots in the highly regulated healthcare industry

The foundational methodology of AXS draws inspiration from the stringent protocols mandated by the U.S. Food and Drug Administration (FDA) for medical device testing. The FDA, an authoritative body responsible for protecting public health by ensuring the safety and efficacy of medical devices (among other responsibilities), sets forth comprehensive guidelines that developers must follow throughout the medical device development cycle. These guidelines emphasize rigorous testing and evaluation to ensure that medical devices meet the highest standards of safety and functionality. The process is designed to identify all significant UX issues, even in studies with limited participant numbers, ensuring that findings are stable, reliable, and reproducible. By mirroring the methodological rigor that the FDA applies to life-critical medical device assessments, the AXS framework achieves reliability and statistical stability in automotive UX evaluation. This approach underscores the commitment to delivering a UX evaluation tool that is both robust and sensitive to the nuances of user interaction with automotive interfaces.

Robust, consistent, international

The robustness of the AXS is further evidenced by comprehensive statistical validation efforts, including correlation testing and factor analysis to neutralize brand perception biases. International testing in key markets—Japan, the USA, and Germany—has validated the AXS framework's consistency across different cultural contexts and highlighted the adaptability of automotive interfaces and the system's responsiveness to regional variations.

As the automotive industry progresses, the AXS is set to become an asset for OEMs, harmonizing UX research initiatives, elevating the importance of user feedback, and enhancing the overall driving experience. This whitepaper outlines the AXS core principles, metrics, and the extensive validation it has undergone, establishing a new standard for automotive UX evaluation and signifying a leap forward in user-centric design and development .

01

Method overview

The development journey of the AXS framework

In the competitive landscape of the automotive industry, OEMs continuously seek methodologies not only to benchmark their products' UX against the competition but also to derive actionable insights to inform future development, establish realistic key performance indicators (KPIs), and adopt best practices. Recognizing this perennial need, the research teams from all ReSight Global companies dedicated two decades to exploring and evaluating automotive UX, harnessing extensive experience to pioneer testing approaches that address the industry's complex demands.

The journey began with a simple premise: provide OEMs with a comprehensive toolkit for assessing and comparing their vehicles' UX in a scientifically robust and practically insightful way . Over the years, the research team has experimented with and refined many testing methodologies, each contributing to understanding how to measure and interpret UX within the automotive domain effectively. This iterative development, application, and validation process has resulted in the ***Automotive Experience Score (AXS)*** as a single quality score for automotive interaction UX and a comprehensive framework.

The AXS framework is the product of rigorous research and practical application designed to meet the nuanced needs of the industry. The framework offers a statistically reliable, top-level market overview essential for strategic product planning and management. However, its ability to delve into the granular details of UX—at the task and domain levels—sets it apart, providing developers and engineers with the in-depth information needed to pinpoint areas for enhancement and innovation.

The AXS framework embodies a comprehensive evaluation tool bridging market insights and actionable feedback, empowering OEMs to excel in an evolving industry and ensuring products exceed discerning consumer expectations.

Research approach and methodology

The AXS allows for a reliable comparison of automotive interface UX between many current car models. Centered around feedback and measurements from real users, the AXS framework incorporates four principal metrics: the After-Scenario Questionnaire (ASQ), UX Score, Task Completion Rate, and Driver Distraction Rating. These metrics collectively encapsulate various facets of UX, blending subjective feedback with objective performance indicators to comprehensively evaluate automotive UX.

Core Principles of the AXS Methodology

The AXS methodology is rooted in authentic user interactions. It focuses on 11 standardized tasks across all tested vehicles to ensure consistency and comparability in assessing each vehicle's UX.

Protocol inspired by medical device testing

Emulating the structured and rigorous summative testing procedures of medical devices, the AXS protocol is characterized by:

Including real users: All metrics are gathered from 40 one-on-one, structured, in-person interviews conducted inside each tested car.

The journey began with a simple premise: provide OEMs with a comprehensive toolkit for assessing and comparing their vehicles' UX in a scientifically robust and practically insightful way.

Uninterrupted tasks: Reflecting real-life usage scenarios, tasks are completed without interruptions from the moderator, ensuring an uninterrupted interaction flow. However, akin to medical device testing, participants are subjected to a constant mental load simulating real-life driving conditions, achieved through concurrent engagement with the Box task and Detection Response Task (BT+DRT). This dual-task approach ensures that assessments accurately reflect the cognitive demands placed on drivers in actual-use scenarios.

Identical task instructions: Uniform task instructions are provided to all participants, ensuring data consistency and reliability across tests.

Clear, uniform instructions: Every participant receives concise, unambiguous instructions, minimizing confusion and standardizing the testing experience.

Two rounds of testing - training and measurement: Rather than relying on a singular 'walk-up-and-use' scenario, the AXS protocol employs a two-phase testing process to capture the nuances of the user experience with the system over time. Initially, participants undergo training to familiarize themselves with the vehicle's interface. This step is crucial for simulating a realistic scenario where a user has had at least a brief period to acclimate to the system, reflecting the expectation that any car UI should be sufficiently intuitive to allow for ease of use after a short familiarization period. Following this, the measurement phase begins, where the performance and experience of users are evaluated. This dual-phase approach is designed to provide a deeper understanding of how users interact with and perceive the system after using it for some time, offering a more comprehensive assessment than possible through a single, immediate measurement.

Structured session guide and digital data collection: The testing process is supported by a detailed session guide, which employs digital data collection tools for accuracy and efficiency in data gathering.

Clear success/fail definitions: Success and failure criteria for each task are explicitly defined, facilitating objective evaluation of task completion and performance.

A novel addition to the AXS method is informing participants of their success in completing each task during the measuring round before administering the UX Score questionnaire. This approach ensures that participants' subjective ratings are informed by a clear understanding of their task performance, leading to more nuanced and accurate reflections of their experience.

Task selection criteria

The team carefully selected tasks to provide insights to help OEMs enhance their vehicles' HMIs. Three fundamental criteria governed the selection process:

1. Complexity and depth of interaction: The primary goal was to select tasks with a certain complexity level. Simple tasks, like those requiring a single button press, barely scratch the surface when distinguishing the UX capabilities across various car models. In contrast, complex tasks force participants to delve deeper into the interface, exposing the intricacy and effectiveness of the information architecture. By demanding more involved interactions, these tasks challenge participants, offering a broader scope for assessing the design and functionality of the interface. This approach ensures the evaluation captures the subtleties of how the interface handles integrated and complex functions, spotlighting designers' significant challenges in creating intuitive and efficient user interactions.

2. Comprehensive representation: To assess the automotive interface, selecting tasks that covered all critical domains and avoided redundancy was essential. Overlapping tasks could lead to biased results, focusing on certain elements over holistic evaluation. By choosing a diverse and distinct set of tasks, the team achieved a balanced overview of the interface's UX, providing comprehensive and focused insights into its strengths and areas for improvement across all key functionalities.

3. Focus on OEM-controlled interfaces: A critical aspect of the team's selection process was the exclusion of tasks that rely heavily on third-party applications or other systems outside the OEM's control. For example, interactions with external apps like Spotify were avoided. This decision reflects the team's focus on evaluating OEM-designed interface aspects, enabling targeted improvements based on controllable factors.

The AXS framework evaluates automotive interfaces by adhering to these principles, focusing on complexity, comprehensiveness, and the OEM's design influence.

Core metrics

At the heart of AXS lie four core metrics: the ASQ, the UX Score, the Task Completion Rate, and the Driver Distraction Rating, the latter of which is derived from the BT and DRT methodologies. Each metric plays a crucial role in capturing distinct aspects of UX, from subjective user feedback to objective performance measures.

After-Scenario Questionnaire (ASQ)

Incorporating the ASQ into the AXS methodology underscores a pivotal step toward quantifying user satisfaction in automotive UX studies. The ASQ, initially developed and psychometrically evaluated by Lewis in 1991 [1] and further elaborated in 1995 [2], has been recognized as a robust tool for assessing user satisfaction following specific scenarios or tasks in computer usability studies. Using this tool in the AXS framework upholds the psychometric rigor set by Lewis and is customized for automotive UX research.

The ASQ's power lies in its three concise questions that capture user reactions and satisfaction levels after an interaction. In the AXS framework, the methodology specifically incorporates only the first two questions from the ASQ, focusing on measuring users' satisfaction with the ease of completing tasks and the time efficiency of these tasks within the automotive interface. This strategic decision is based on the team's empirical research, revealing a wide range of responses to the third question designed to assess system support. These varied responses underscore a pronounced mismatch between the question's intended purpose and how participants, particularly those in the automotive sector, interpret "system support," leading to significant variations in perceptions. This approach helps capture user satisfaction, indicating the effectiveness and intuitiveness of automotive interfaces.

UX Score

Integrating the UX Score into the AXS framework significantly enhances the ability to assess the subjective aspects of user interactions with vehicle interfaces. This approach takes a comprehensive view of measuring UX, relying on user self-reports through a questionnaire. It moves beyond focusing on usability, incorporating emotional and aesthetic reactions to automotive HMIs. The rationale behind adding the UX Score to the AXS framework is rooted in its established structure and importance in capturing the full range of user interactions within the automotive context.

The UX Score's conceptual foundation classifies UX into three primary dimensions: task-oriented qualities (learnability and operability), self-oriented qualities (product fit and inspiration), and aesthetic qualities (product look and feel). This categorization, detailed in the research by Wildner, Kittinger-Rosanelli, and Bosenick (2015)[3], captures the comprehensive nature of user interactions, emphasizing the importance of both functional and emotional aspects of UX. Their international validation of the UX Score, utilizing a detailed 10-item scale, confirms its effectiveness and reliability as a measure of subjective user satisfaction and engagement.

Traditional usability metrics, while essential, do not fully capture the nuances of user experience in the automotive domain, where emotional and aesthetic considerations play a significant role. The UX Score's comprehensive approach to measuring these aspects offers a more holistic understanding of user interactions, enabling automotive manufacturers to better align their products with user expectations and preferences.

The application of the UX Score, as Wildner et al. (2015) demonstrate in their assessment of car infotainment systems, proves its value in pinpointing strengths and areas for improvement across brands and regions. This metric enables comprehensive benchmarking and delivers precise insights into aspects needing enhancement—such as operability, aesthetic appeal, or better alignment with user needs.

Including the UX Score in the AXS framework enables car companies to prioritize user-centric design and engineering. By pinpointing aspects of the HMI for user satisfaction, the UX Score motivates strategic enhancements beyond usability, resulting in a more engaging and enjoyable user experience. This reflects a broader trend in the automotive industry towards designs that value emotional and aesthetic responses as much as functional efficiency.

Task completion rate

The task completion rate is a crucial metric in UX research, indicating the quality of a system. Its reliability and relevance are strengthened by its endorsement in established works and standards, including the "Applying human factors and usability engineering to medical devices" guidance published by the FDA[4], and the International Electrotechnical Commission's (IEC) international standard IEC 62366-1[5]. These documents underscore the task completion rate as a fundamental criterion for evaluating the usability and effectiveness of a system, be it medical devices or automotive interfaces.

The essence of the task completion rate as a measure lies in its straightforward yet powerful premise: it quantifies the percentage of tasks that users correctly complete within a given scenario or system. This metric shows how well the design helps users achieve their goals, making it a key part of system quality. In environments where precision and safety are paramount, such as medical devices or automotive interfaces, the importance of users completing tasks without errors cannot be overstated. Consequently, a high task completion rate directly correlates with a system's usability, indicating that users can efficiently and effectively achieve their objectives.

The validation of task completion rate as a reliable measure of system quality extends across various fields. For instance, in medical devices, guidance on "Applying human factors and usability engineering" recommends using the task completion rate to assess the safety and effectiveness of these devices. This recommendation ensures that the devices meet the highest standards of user-centered design, significantly reducing the risk of user errors that could result in adverse outcomes. Similarly, the IEC 62366-1 standard highlights the necessity of understanding and measuring the effectiveness with which users can interact with a device to perform intended tasks, marking it as a crucial aspect of usability.

In automotive UX, the task completion rate measures how easily drivers or passengers use car interfaces for functions like adjusting climate controls or navigating multimedia systems. A high task completion rate indicates that the vehicle's HMI is designed to allow users to accomplish their tasks with minimal confusion and distraction, thus enhancing both user experience and safety.

The task completion rate is more than just a metric; it reflects the usability and quality of user-centered design. It is widely praised in different industries for its reliability as a UX benchmark. Using the task completion rate in automotive

UX follows best practices and offers a clear, measurable view of how well the interface meets user needs and safety standards.

Driver distraction metric (Box task & Detection Response Task)

BT and DRT stand out as key methodologies for assessing the effects of visual-manual and cognitive distractions within the automotive context.

The Box Task (BT)

The Box Task simulates a car-following scenario, wherein a visual element, referred to as a "box," changes size and position to mimic the dynamics of maintaining a safe headway and lane position. Participants engage with this task through a steering wheel and a gas pedal, adjusting their actions to the box's movements to keep it within predefined boundaries. This setup assesses visual-manual demand by quantifying participants' ability to respond to the changing task parameters, providing insights into how in-vehicle systems might impact fundamental driving tasks.

Detection Response Task (DRT)

The DRT enhances the BT by focusing on the cognitive load and its impacts on driver attention. It presents participants with random, periodic stimuli across visual, tactile, or auditory modalities. Participants must respond to these stimuli while performing primary and secondary driving-related tasks. For example, a red light appears on the screen at random intervals, and the participant must press the brake pedal. Their response times and hit rates act as indicators of cognitive distraction, with longer response times and lower hit rates indicating higher levels of cognitive load.

This approach enables a nuanced assessment of driver distraction and in-vehicle system demand, which is essential for calculating the AXS. It measures the cognitive and visual-manual load on drivers, providing a unique perspective for evaluating the overall user experience of automotive interfaces. This methodology, validated by Morgenstern et al. (2020)[6] and Trommler et al. (2021)[7], is a fundamental component of the AXS.

The BT + DRT differs from traditional evaluation methods, such as actual driving on test tracks or simulators, in several key ways:

Safety and standardization: Unlike real traffic and test track assessments, which can be variable and hazardous, the BT + DRT is conducted in a controlled environment. This ensures participant safety and test consistency, which is critical for standardizing the AXS across different car models and locations.

Eliminates simulator sickness: Simulator-based assessments often result in participants experiencing simulator sickness, which can skew results and complicate the testing process. The BT + DRT method circumvents this issue, offering a distraction assessment tool that avoids the physical discomfort and data reliability challenges associated with simulator sickness.

Comprehensive distraction analysis: The method comprehensively analyzes distractions by measuring cognitive and visual-motor aspects, offering a more rounded understanding than what is seen in simulators or real-world driving alone. This analysis is crucial for developing intuitive interfaces that minimize cognitive load and enhance safety and UX.

Integrating the BT + DRT into the AXS framework brings a scientifically validated, safe, and efficient method for evaluating automotive interface UX. This approach offers detailed insights into how in-vehicle systems affect driver attention and performance, guiding UI design for safety and intuitiveness. Compared to traditional testing, the BT + DRT's advantages underscore its crucial role in enhancing the AXS, ensuring a comprehensive evaluation of automotive UX that aligns with real driver interaction and distraction.

Alternative metrics as candidates for the AXS

Recognizing the significance of a comprehensive approach to assessing automotive UX, the research team considered incorporating well-established metrics such as the System Usability Scale (SUS) and the User Experience Questionnaire (UEQ) into the AXS framework. These tools, known for their effectiveness in measuring usability and UX across various fields, underwent a thorough evaluation to determine their potential to enhance AXS's diagnostic capabilities. However, after a detailed analysis, it became clear that although the SUS and UEQ offer considerable value, they do not meet the specific goals and methodological standards the AXS sets. This chapter explains the reasons behind this decision, emphasizing the commitment to delivering the most precise and relevant insights for the automotive industry.

System Usability Scale

The SUS questionnaire, developed by John Brooke in 1996 [8], serves as a concise and reliable tool for evaluating the usability of various systems and products. It provides a "quick and dirty" method for measuring usability, making it a popular choice among researchers and practitioners looking to assess interface effectiveness efficiently. The SUS uses a ten-item questionnaire with five response options per item, offering a global perspective on subjective usability assessments. Its simplicity and versatility have led to widespread adoption across domains and technologies. Notable research and evaluations by Bangor, Kortum, and Miller (2008) [9], as well as Lewis (2018) [10] and Gao, Kortum, & Oswald (2020) [11], have further highlighted its utility, even expanding its applicability to multiple languages.

While the SUS is useful for measuring usability, it has a narrow focus, mainly assessing system ease of use. However, usability is just one aspect of automotive HMI UX. Modern research requires a broader view, considering user satisfaction, emotional response, and aesthetic appeal. To address this, metrics like the ASQ and UX Score are integrated into the evaluation.

The ASQ gauges user satisfaction after interaction, providing immediate insights into the user experience. Meanwhile, the UX Score covers task-oriented, self-oriented, and aesthetic aspects, offering a detailed view of how users perceive automotive interfaces.

By combining task measurements with the ASQ and UX Score, researchers gain a richer understanding of the user experience. This approach captures the complexity of user engagement, which usability metrics alone cannot fully grasp. For instance, a system might be easy to use but still unsatisfactory if it lacks aesthetic appeal or fails to evoke positive emotions.

Considering participant fatigue, it is crucial to avoid overloading them with excessive questions. The ASQ and UX Score already provide comprehensive insights, making the addition of the SUS unnecessary.

While the SUS is valuable for usability assessment, the research framework prioritizes a holistic approach to automotive UX. By combining task-based and broader UX assessments, researchers can better understand the interconnected nature of user interactions with automotive systems.

User Experience Questionnaire (UEQ)

In the journey to refine automotive UX assessment, the research team faced a pivotal decision: to integrate the UEQ as a primary metric within the AXS framework or to continue leveraging the UX score.

Laugwitz, Schrepp, and Held (2008) [12] extensively researched and validated the UEQ in 2008, with further refinements by Schrepp, Hinderks, and Thomaschewski (2017)[13] [14] in 2017. The UEQ is a comprehensive tool for assessing various dimensions of UX, from attractiveness to efficiency and novelty. It can be applied to diverse interactive products, providing detailed insights into user perceptions and satisfaction.

Despite the UEQ's robust construction and widespread adoption across different domains, several aspects prompted a reassessment of its suitability as the sole UX metric for the AXS framework.

The decision to develop and utilize the UX Score as the primary metric within the AXS framework stemmed from several factors:

1. Complexity and length: The UEQ comprises many items distributed across six scales. In extensive research, the team discovered that Likert scales used in the UEQ can pose challenges for participants, potentially impacting the reliability and precision of their responses. In contrast, the UX Score, streamlined and tailored to automotive UX evaluation, mitigates participant fatigue without compromising the depth of insights.

2. Focused relevance: Specifically designed with automotive interfaces in mind, the UX Score ensures that each component directly pertains to critical aspects of automotive UX, providing actionable feedback relevant to automotive developers and engineers.

3. Statistical reliability: The UX Score has demonstrated high statistical reliability through rigorous testing and application. It offers a dependable measure of automotive UX quality, supporting its role in setting benchmarks and guiding development strategies.

Comparing the UEQ and UX Score emphasizes the need for a tailored approach to automotive UX assessment. The UX Score, with its emphasis on simplicity, relevance, and actionable insights, aligns with the objectives of the AXS framework. While the UEQ remains a powerful tool for a broad range of UX evaluations, the specialized focus of the UX Score on automotive interfaces ensures that the AXS framework delivers targeted, meaningful assessments that drive innovation and enhancement in automotive UX design.

Supplementary metrics and information

The AXS dashboard presents a comprehensive collection of supplementary metrics and detailed insights. This data extends beyond the foundational components of the ASQ, UX Score, task completion rate, and driver distraction rating. The abundance of information provides researchers, engineers, and designers with the necessary tools to understand the factors influencing a car's AXS rating and analyze task successes or failures across domains or activities.

Net Promoter Score (NPS)

Incorporating the NPS into the evaluation framework strategically aligns with industry standards and executive expectations, as they highly value the simplicity of a single metric like the NPS. Originating from the seminal work by Frederick F. Reichheld, "One Number You Need to Grow," in the December 2003 issue of the Harvard Business Review [15], the NPS has become a cornerstone metric across various industries, offering a straightforward measure of customer loyalty and brand perception.

To enhance AXS insights and address concerns about brand perception's impact on UX evaluations, a dual application of the NPS has been implemented. Participants are first asked to rate their likelihood of recommending the car's brand to a friend or colleague before engaging with it. This initial measure provides valuable insight into how pre-existing brand perceptions influence the overall AXS.

After the UX testing phase, researchers ask participants again about their likelihood of recommending the car to a friend or colleague, repeating the NPS question. Prior research demonstrates a strong link between the UX Score and the NPS (see chapter "Validation studies" below for more information on this research), showing that better UX directly increases the willingness to recommend the brand. Researchers include the NPS again because it is a widely recognized key performance indicator (KPI) with broad acceptance and usefulness, even to executives unfamiliar with detailed UX metrics. Furthermore, asking just one additional question adds little burden, keeps participants engaged, and provides deep insights into how UX improvements can boost brand advocacy.

Number of UX issues

Including the metric that quantifies the number of UX issues within the AXS dashboard enriches the toolkit for assessing automotive interface user experience. This metric, though distinct from the AXS, is critical for deepening the understanding of a vehicle's UX by offering detailed insights for targeted enhancements. Combined with the issue frequency and severity measurements, the number of UX issues forms a comprehensive framework for evaluating and benchmarking automotive products' UX.

Enumerating UX issues provides diagnostic clarity, pinpointing specific areas needing interface refinement. This granular visibility into the UX challenges helps stakeholders to identify and prioritize issues for resolution effectively, focusing efforts on enhancing areas critical to user satisfaction.

Utilizing the count of UX issues as a benchmark offers an objective method to compare a vehicle's UX performance against competitors and industry standards. This analysis highlights a brand's position in the UX landscape, identifying opportunities for differentiation and improvement. When used alongside metrics for issue frequency and severity, it offers a nuanced perspective on UX quality, comparing not just the number of issues but their impact and commonality.

Integrating issue frequency and severity metrics with the count of UX issues delivers a multi-dimensional view of the UX. While the total number of issues provides a broad snapshot of the UX landscape, the frequency metric reveals how often users are likely to encounter these problems, and the severity rating indicates the issues' impact. This trio of metrics enables a holistic approach to UX evaluation, guiding strategic improvements by illustrating problems' prevalence, depth of user impact, and location.

Severity of UX issues

In automotive UX research, effectively classifying the severity of UX issues becomes pivotal for prioritizing improvements and ensuring the highest standards of usability and satisfaction. To this end, researchers employ a streamlined three-point scale for UX issue severity classification, designed to provide clear guidance for UX researchers and actionable insights for development teams. Below is an overview of each severity level, accompanied by automotive-specific examples:

Severity 1: Critical (Showstopper)

Critical issues prevent users from completing tasks, representing substantial barriers within the user journey. In the automotive context, these could include unresponsive touchscreen controls for essential functions like climate control or navigation, leading to complete task failure. For instance, if voice commands for adjusting navigation settings fail to recognize user input consistently, researchers classify this as critical. These issues require immediate attention and resolution because they significantly impact the user experience and may pose safety risks.

Severity 2: Moderate (Impediments)

Moderate issues disrupt the interaction flow, causing frustration or confusion but still allowing users to eventually complete their tasks. Examples include poorly organized menu systems that significantly delay access to features like media playback or contact lists. Another example is a delayed response from the HMI when switching between modes (e.g., from navigation to media settings), which increases cognitive load. Addressing these issues is crucial for enhancing usability and reducing user frustration.

Severity 3: Minor (Cosmetic)

Minor issues pertain to aesthetic inconsistencies or minor annoyances that do not directly impede task completion. In an automotive HMI, this could include mismatched icon styles or font sizes across different screens or slight

misalignments in the graphical user interface. While these issues do not affect the core functionality, resolving them contributes to a more polished and cohesive user interface. Cosmetic issues are considered a lower priority than critical and moderate issues but should be corrected to achieve a refined and professional product presentation.

This severity classification system is foundational to the UX evaluation process, ensuring that all identified issues are accurately categorized and prioritized. By providing clear, automotive-specific examples and detailed descriptions of each severity level, the aim is to standardize the assessment process across researchers, facilitating consistent and effective communication of UX findings and recommendations.

Frequency of UX issues

Understanding the frequency of UX issues is crucial for prioritizing interface improvements. The AXS methodology provides a metric of issue frequency, which helps identify the severity of an issue and how often users are likely to encounter it during interactions with the vehicle's HMI.

Moderators note each occurrence of specific issues experienced by participants to capture this data systematically, providing a quantifiable metric of issue frequency. This metric helps automotive UI developers prioritize UX improvements.

Analyzing issue frequency with severity ratings allows teams to strategically address issues that pose the most significant risk to the user experience. Consequently, this methodology ensures that development efforts effectively enhance user satisfaction and safety, optimizing the overall effectiveness of the automotive interface.

UX issue description and video documentation

In refining the approach to capturing and analyzing UX issues within the AXS framework, ReSight Global considers the nuanced complexities of automotive interface evaluation. Drawing upon research and best practices, including insights from Nielsen (1992) [16], Nielsen and Landauer (1993) [17], and Dumas et al. (1995) [18], the process is structured to harness the collective expertise of multiple UX professionals. This approach is guided by the understanding that an evaluation by multiple UX experts can yield a more comprehensive identification of usability issues.

For each vehicle assessed under the AXS, 40 user interviews provide a rich dataset of real user interactions. To complement this data, at least three UX experts from ReSight Global reviewed all identified UX issues. This multi-expert review ensures a nuanced interpretation of each issue, grounded in observed user behaviors and expert analysis. This strategy aligns with findings that leveraging a greater number of experts can enhance the effectiveness of usability evaluations instead of relying on a deeper review by fewer specialists.

The research team provides detailed information on identified UX issues, clearly distinguishing between the observation—what was directly noted during user testing—and the interpretation— experts' analysis of why an issue occurred, based on expertise and user feedback. This dual perspective ensures an empirical account of the UX challenges and an expert assessment of underlying causes. Moreover, the AXS dashboard includes video evidence of each problem to bring these issues to life, allowing stakeholders to see the issues "in action." This feature illuminates the practical implications of each identified problem and aids in prioritizing fixes and enhancements.

Time on task

The "mean time on task for successful task completion" metric in the AXS dashboard is strategically valuable for competitive analysis. It provides an average duration for participants to complete specific tasks, allowing a direct comparison of interface efficiency with competitors.

For developers, this metric is instrumental in benchmarking their vehicle's user interface (UI) against the industry standard and competitors. By comparing mean time on task, developers can pinpoint precisely how their UI performs in efficiency and user-friendliness. A shorter mean time suggests a more intuitive interface, indicating a competitive UX advantage. A longer time reveals opportunities for design improvements to enhance usability and potentially surpass competitors.

This insight enables development teams to allocate resources for UI improvements effectively. It illuminates areas where enhancements can significantly reduce task completion time, impacting user satisfaction and overall perception of the vehicle. Tracking changes in mean time on task over iterations allows for a clear assessment of progress in UI optimization efforts relative to competitors.

Essentially, the “mean time on task for successful task completion” metric in the AXS dashboard empowers developers with the data necessary for targeted improvements. It facilitates a better understanding of their vehicle’s UX strengths and weaknesses and provides a clear benchmark for measuring success in achieving a more efficient and competitive automotive interface.

Percentage of time on task spent in incorrect navigation paths

The AXS dashboard features the percentage of time participants spend in incorrect navigation during task completion, a valuable metric for information architects and UI designers. This metric helps identify navigational challenges, guiding refinements to reduce inefficiencies and confusion in the UI. Streamlining navigation improves efficiency and user experience by reducing cognitive load and distractions, promoting safer interactions. Tracking changes over time benchmarks UI improvements, revealing the impact of design interventions on user behavior.

Car Interface Metrics

To deepen the understanding of automotive UX, Resight Global introduces “Car Interface Metrics” within the AXS dashboards and toolkit. This suite of metrics goes beyond traditional UX evaluation to include detailed specifications of the car’s user interface elements, offering OEMs a new dimension of insight. By integrating these metrics, the research teams aim to illuminate the intricate relationships between user performance, perception, and the specific characteristics of interface elements.

Car Interface Metrics provide a structured compilation of data points that characterize the user interface of vehicles, including but not limited to:

Screen specifications: Sizes and resolutions of the main cluster display, Central Information Display (CID), and passenger displays. These metrics offer a foundational understanding of the visual interface’s capacity for information presentation.

Voice recognition system: Details include the number of domains covered and the interaction style. This information helps assess the system’s versatility and intuitiveness in understanding and executing voice commands.

Steering wheel-mounted controls: This metric evaluates the number and style of interaction elements on the steering wheel and the domains they control. It also evaluates how these controls integrate with the vehicle’s overall UX.

Head-up display (HUD): The availability and range of functions presented. The HUD’s inclusion and capabilities contribute to a safer and more immersive driving experience by minimizing driver distraction.

Haptic feedback systems: The presence and types of haptic feedback provided by the interface are crucial for enhancing user engagement and reducing visual attention requirements.

Customization options: The extent to which users can personalize interface settings to suit their preferences, impacting overall satisfaction and ease of use.

Connectivity features: Details on integrating external devices and services, such as smartphone connectivity and internet-based services, which are increasingly important for users.

Gesture control capabilities: The inclusion and range of gesture-based controls offer users a modern and potentially more intuitive interaction method.

Task-specific hardware interaction elements: This metric identifies hardware controls, like satellite levers or physical buttons, dedicated to specific tasks, highlighting alternatives to touchscreen interaction. It assesses the diversity and accessibility of interaction modes, impacting the ease of use and the overall driving experience.

The compilation of Car Interface Metrics alongside measured UX dimensions enables OEMs to make informed decisions based on the correlation between interface specifications and UX outcomes. Understanding how specific interface characteristics relate to user performance and perception in tasks helps OEMs pinpoint areas for improvement and innovation. This holistic approach facilitates targeted enhancements in car interface design and helps set benchmarks for future development, ensuring that vehicles meet evolving user expectations.

02

Validation studies

Stability / Accuracy

The development and application of the AXS necessitated rigorous validation across diverse cultural contexts. The research team conducted a study, testing the same car models in Japan, the United States, and Germany to ensure consistent results. Participants completed tasks, and researchers evaluated their experiences using core AXS metrics. Normalizing scaled feedback from each country was essential to mitigate cultural biases in survey responses.

The analysis revealed remarkable stability in the AXS metrics across the tested regions, with no significant differences in task completion rates, ASQ responses, and UX Scores—except for two notable instances that underscored the methodological accuracy of the AXS:

Task completion rate anomaly in the US: Researchers observed a significantly higher task completion rate for a specific task in the US. It was discovered that this anomaly was due to an over-the-air update by the manufacturer, which resolved a critical UX issue previously identified in tests conducted in Germany and Japan. This finding validated the sensitivity of the AXS to detect interface modifications and highlighted its potential in guiding iterative design improvements.

Performance discrepancy in Japan: The second anomaly involved a significant divergence in task completion rates, the second ASQ question ratings, and certain UX Score aspects for the Japanese car, compared to its US and German counterparts. Further investigation, in collaboration with the car's manufacturer, revealed that the Japanese car model was equipped with a slower processor. This hardware limitation resulted in a less responsive system, directly impacting participants' ability to complete tasks efficiently and affecting their subjective ratings. This instance demonstrated the AXS's unintended yet valuable capability to pinpoint hardware-related UX bottlenecks.

These findings confirm the AXS framework's statistical stability across cultures and its sensitivity to software updates and hardware differences. Originally for measuring

automotive interface UX, the AXS also detects minute yet significant performance differences, showcasing its robustness as a global standard. Through its application, ReSight Global has established a methodologically sound and culturally adaptive framework that accurately reflects automotive UX complexities.

Impact of brand perception on AXS

To understand the interplay between brand perception and UX within the automotive sector, researchers conducted a comprehensive analysis leveraging the NPS and various subjective UX metrics. This analysis aimed to discern how pre-existing brand perceptions influence the evaluation of automotive user interfaces, as reflected through the AXS framework.

To achieve this, researchers measured the NPS for the brand of the tested cars before participants engaged with the vehicles. This pre-test NPS aimed to capture participants' initial perceptions of the brand as a baseline for understanding how these perceptions might color their subsequent UX evaluations. After interacting with the cars, the team collected participants' responses to the ASQ and the overall UX Score. They also gathered a post-test NPS, reflecting participants' likelihood to recommend the specific car after testing.

The researchers employed the Pearson correlation coefficient to quantify the relationship between brand perception and UX evaluations. This statistical measure evaluates the linear correlation between two variables, producing a value between -1 and 1. A coefficient close to 1 implies a strong positive correlation, indicating that as one variable increases, the other does likewise. Conversely, a coefficient close to -1 suggests a strong negative correlation, where one variable increases as the other decreases. A coefficient around 0 indicates little to no linear correlation between the variables.

The data conclusively indicates that enhancing the UX directly and potently affects users' likelihood to recommend a car, far outweighing the initial biases brought by brand perception.

The Pearson correlation analysis of n > 350 participants revealed enlightening insights:

- The pre-test NPS showed a very low correlation with the first and second ASQ questions (0.079 and 0.062, respectively) and a modest correlation with the UX Score (0.301). These results suggest that initial brand perceptions have minimal influence on the specific assessments of the user experience.

- In contrast, the post-test NPS for the specific car tested exhibited a robust correlation with the UX Score (0.713), indicating that the direct experience with the car significantly impacts the likelihood of participants recommending it.

These findings highlight a critical insight: the quality of the user experience, as quantitatively evaluated through the UX Score, plays a pivotal role in shaping advocacy and brand loyalty, as measured by the post-test NPS. The stark difference in correlation coefficients between the pre-test and post-test NPS underscores the importance of intrinsic user experience over pre-existing brand perceptions. This analysis reinforces that delivering an exceptional UX within the automotive industry is paramount for fostering positive word-of-mouth and enhancing brand equity.

Using Pearson correlation analysis, this study underscores the significance of focusing on the quality of user interactions and interface design in automotive development. The data conclusively indicates that enhancing the UX directly and potently affects users' likelihood to recommend a car, far outweighing the initial biases brought by brand perception. These insights serve as a valuable guide for OEMs in prioritizing UX enhancements to drive brand advocacy and loyalty.

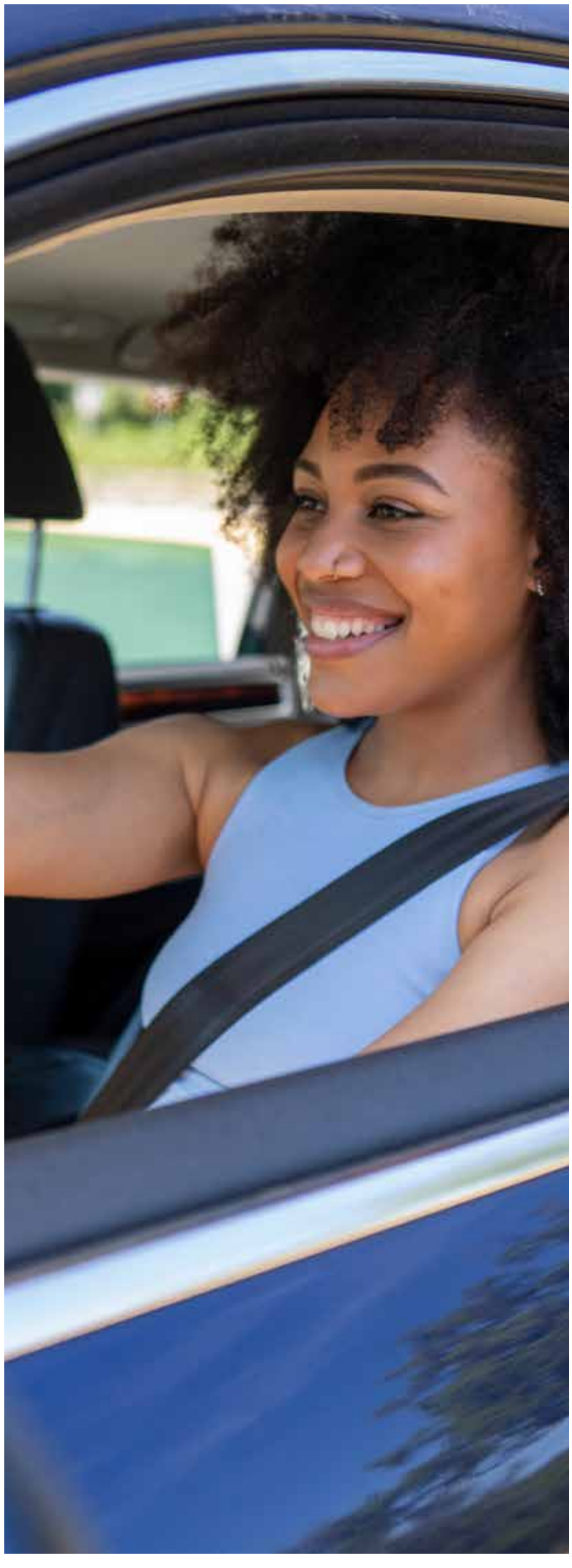
03

Conclusion

The AXS framework is a testament to ReSight Global's dedication to revolutionizing automotive design and development through comprehensive user experience evaluation. It integrates a suite of meticulously designed metrics, such as the UX Score, ASQ, Task Completion Rate, and innovative methods like the BT and DRT, offering an unparalleled perspective on automotive UX.

Drawing inspiration from the rigorous standards of medical device testing, the methodology behind the AXS ensures stable, reliable, and reproducible results across a diverse and global user base. This meticulous attention to detail and methodological rigor establish the AXS as a pivotal asset for navigating the complex demands of modern automotive UX evaluation. The framework's successful application and validation in multiple cultural contexts underline its robustness and adaptability, cementing its relevance and usefulness worldwide.

For OEMs, the AXS becomes both a tool and a strategic ally. It empowers organizations at various levels, from product planning and management to development and engineering teams, providing insights from high-level market overviews to granular task and domain-level analyses. This versatility enables OEMs to make informed decisions that enhance user satisfaction, drive innovation, and, ultimately, shape the future of automotive experiences.



References

1. *Lewis, J. R. (1991). Psychometric evaluation of an after-scenario questionnaire for computer usability studies: The ASQ. SIGCHI Bulletin, 23(1), 78-81.*

2. *Lewis, J. R. (1995). IBM Computer Usability Satisfaction Questionnaires: Psychometric Evaluation and Instructions for Use. International Journal of Human-Computer Interaction, 7(1), 57-78.*

3. *Wildner, R., Kittinger-Rosanelli, C., & Bosenick, T. (2015). How Good is Your User Experience? Measuring and Designing Interactions. GfK MIR, 7(2). Available at: https://www.researchgate.net/publication/284086367_How_Good_is_Your_User_Experience_Measuring_and_Designing_Interactions*

4. *FDA: Applying human factors and usability engineering to medical devices. Available at: https://www.fda.gov/media/80481/download*

5. *IEC 62366-1: Medical devices - Part 1: Application of usability engineering to medical devices. Available at: https://webstore.iec.ch/preview/info_iec62366-1%7Bed1.1%7Db.pdf*

6. *Morgenstern, T., Wögerbauer, E.M., Naujoks, F., Krems, J.F., & Keinath, A. (2020). Measuring driver distraction - Evaluation of the box task method to assess in-vehicle system demand. Applied Ergonomics, 88, 103181. doi: 10.1016/j.apergo.2020.103181.*

7. *Trommler, D., Morgenstern, T., Wögerbauer, E.M., Naujoks, F., Krems, J.F., & Keinath, A. (2021). The box task - a method for assessing in-vehicle system demand. MethodsX, 8, 101261. doi: 10.1016/j.mex.2021.101261.*

8. *Brooke, J. (1996). SUS: A 'quick and dirty' usability scale. In Jordan, P.; Thomas, B.; Weerdmeester, B. & McClelland, I. L. (ed.) Usability Evaluation in Industry. S. 189-194, Taylor & Francis.*

9. *Bangor, A., Kortum, P. T. & Miller, J. T. (2008). An Empirical Evaluation of the System Usability Scale. International Journal of Human-Computer Interaction, 24(6), 574 - 594.*

10. *Lewis, J. R. (2018). The System Usability Scale: Past, Present, and Future. International Journal of Human–Computer Interaction, 34(7), 577-590. DOI: 10.1080/10447318.2018.1455307.*

11. *Gao, M., Kortum, P. & Oswald, F. L. (2020). Multi-Language Toolkit for the System Usability Scale. International Journal of Human–Computer Interaction, 36(20), 1883-1901. DOI: 10.1080/10447318.2020.1801173.*

12. *Laugwitz, B., Schrepp, M., & Held, T. (2008). Construction and evaluation of a user experience questionnaire. In A. Holzinger (Ed.), USAB 2008, LNCS 5298, pp. 63-76.*

13. *Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Construction of a benchmark for the User Experience Questionnaire (UEQ). International Journal of Interactive Multimedia and Artificial Intelligence, 4(4), 40-44.*

14. *Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Design and Evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). IJIMAI, 4(6), 103–108.*

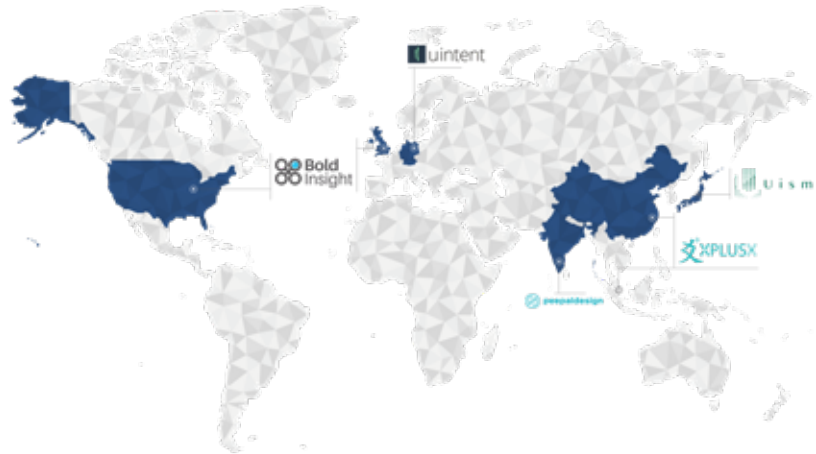
15. *Reichheld, Frederick F. (December 2003). "One Number You Need to Grow". Harvard Business Review, 81(12): 46–54, 124. PMID 14712543.*

16. *Nielsen, Jakob (1992). "Finding Usability Problems Through Heuristic Evaluation." In CHI '92 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. New York: ACM Press.*

17. *Nielsen, Jakob, and Thomas Landauer (1993). "A Mathematical Model of the Finding of Usability Problems." In Proceedings of the ACM/IFIP INTERCHI '93 Conference. New York: ACM Press.*

18. *Dumas, Joseph S., John C. Sorce, and Robert A. Virzi (1995). "Expert reviews: How many experts is enough?" Human Factors and Ergonomics Society Annual Meeting Proceedings, vol. 39, no. 1, pp. 228-232. Available at: https://www.researchgate.net/publication/273604589_Expert_Reviews_How_Many_Experts_is_Enough*

Let's talk!



axs@resight.global

www.resight.global

Instagram . @resightglobal

LinkedIn . @resight-global

Our research team

CHINA (Shanghai) **.** XplusX

GERMANY (Hamburg, Munich) **.** uintent

INDIA (Bangalore, Pune) **.** Peepal Design

JAPAN (Tokyo) **.** Uism

SINGAPORE (Singapore) **.** XplusX

UK (London) **.** Bold Insight

US (Chicago) **.** Bold Insight

resight global

Why us?

ReSight Global is a family of **180+ user experience and human factors researchers** with offices in seven countries.

Our experience in the auto industry spans **over 20 years**, leading the industry in innovative application of ground-breaking frameworks and trusted methodologies to shape technology that delights and engages drivers and passengers. **We help manufacturers craft experiences that drive brand loyalty.**

Over 10 **leading OEMs worldwide** have trusted us with their research.

Welcome to the future of car UX

When innovation is moving at a breakneck pace, you want to know you're spending precious time and money developing features and benefits that are meaningful to users. The ***Automotive Experience Score (AXS)*** offers a comprehensive solution to support confident design decisions and elevate your vehicle's user experience.

Learn more: **resight.global/AXS**