AI Experience Impact Score (AXIS) for evaluating AI-based customer support interactions

by Kenji Hayward
Head of Support, Front

A new way of measuring AI performance to drive better customer experiences

As AI becomes integral to customer support, measuring its impact on the customer experience will be crucial for retaining customers.

Businesses must ensure their customers continue having a high-quality experience in AI-led interactions. The three most common points of friction in an AI-led customer interaction are resolution accuracy, interaction effort, and handoff smoothness.

The **AI Experience Impact Score,** or **AXIS**, measures each component to help organizations pinpoint which area of the AI-driven customer experience can be improved. This will help customer service and support teams define their focus areas, prioritize resourcing, and track service quality trends.

This white paper details the AXIS framework, its calculation, and the implementation strategies for both integrated and stand-alone AI systems. If you have any feedback or questions, feel free to reach out.

Kenji Hayward
Head of Support at Front
LinkedIn | front.com

AXIS FRAMEWORK FOCUSES ON:

01

Resolution Accuracy (RA)

Measures how precise AI is in resolving customer inquiries or correctly routing them without human intervention.

02

Interaction Effort (IE)

Evaluates how easily customers interact with AI, focusing on the clarity and simplicity of the communication.

03

Handoff Smoothness (HS)

Assesses how effective the transition is from AI to human agents, emphasizing the continuity of context and minimization of customer effort.

What's a good AXIS score?

1 - 2.9	3.0 - 3.9	4-5
Poor	Fair	Excellent

AXIS scoring rubric

Resolution Accuracy (RA)
Measures AI's ability to correctly identify and address the customer's issue.

Score		Assessment
5	EXCEPTIONAL	AI resolves the issue with total accuracy, providing the correct solution on the first attempt without any errors, unnecessary steps, or additional clarifications.
4	HIGH	AI provides the correct solution with minor clarifications needed, or it resolves the issue accurately but requires minimal extra input from the customer.
3	MODERATE	AI partially resolves the issue, but the customer needs to provide additional details or clarification, or the solution provided isn't entirely accurate, requiring customer input to correct it.
2	LOW	The AI fails to resolve the issue initially, and the customer must rephrase the issue or go through multiple interactions to get a partially correct response.
1	POOR	AI doesn't understand or address the issue, repeatedly giving incorrect answers or irrelevant suggestions, leading to frustration or a required handoff to a human agent for resolution.

Interaction Effort (IE)
Evaluates the ease and intuitiveness of the AI interaction,
focusing on how simple it is for the customer to reach a resolution.

Score		Assessment
5	EFFORTLESS	AI provides a quick and clear solution with minimal back-and-forth; the customer reaches the resolution with only a few interactions and without having to repeat information.
4	EASY	The customer achieves a resolution with minimal effort, but one or two follow-up questions or clarifications are necessary for the AI to fully understand or complete the request.
3	MODERATE EFFORT	The customer must provide several responses or clarify their issue more than once to reach a resolution, making the process somewhat cumbersome but not overly frustrating.
2	HIGH EFFORT	The customer needs to repeat themselves or go through many questions or steps before achieving a resolution, making the process feel drawn-out or inefficient.
1	VERY HIGH EFFORT	The interaction requires excessive effort from the customer, with numerous repetitive or irrelevant prompts that make it hard to achieve a resolution, leading to a frustrating and overly complex experience.

Handoff Smoothness (HS)
Assesses the quality and efficiency of the transition
from AI to a human agent when needed.

Note: If the handoff isn't required, treat HS as a "neutral" value of 5, assuming the AI has handled the interaction completely on its own.

Score		Assessment
5	SEAMLESS	The handoff to a human agent is flawless. The agent has full access to the customer's previous interactions with the AI, requiring no repetition of details, and the transition feels natural and timely.
4	SMOOTH	The transition is effective, with minimal disruption. The agent has most of the customer's details, and any repetition required from the customer is minimal and does not detract from the experience.
3	MODERATE	The transition works but is somewhat inefficient. The agent has partial information but needs the customer to repeat some key details or context, leading to a slight disruption in flow.
2	ROUGH	The handoff is problematic. The agent has little or no context, and the customer must repeat most of the information, creating noticeable friction and an inefficient experience.
1	POOR	The handoff is very poor or fails altogether. The customer is either transferred incorrectly, needs to rephrase their issue entirely, or is left without clear support, leading to a frustrating experience.

Front

Calculate AXIS of an AI-led customer interaction

The industry-standard version of the AXIS formula is calculated with equal emphasis across Resolution Accuracy (RA), Interaction Effort (IE), and Handoff Smoothness (HS).

The unweighted AXIS score can be calculated as follows:

$$\text{AXIS} = \frac{(RA + IE + HS)}{3}$$

Calculate your overall AXIS score by taking the average of the AXIS scores of the total conversations.

Industry-standard AXIS score example

Suppose in a particular customer interaction with AI:

- **Resolution Accuracy (RA)** is scored as 4 out of 5.
- **Interaction Effort (IE)** is scored as 2 out of 5.
- **Handoff Smoothness (HS)** is scored as 5 out of 5.

The AXIS score of 3.67 out of 5 suggests a generally effective AI interaction, with excellent handoff smoothness and resolution accuracy, but indicates room for improvement in reducing customer effort.

$$\text{AXIS} = \frac{(RA + IE + HS)}{3}$$

$$3.67 \approx \frac{(4 + 2 + 5)}{3}$$

Customize AXIS to your customer experience goals

Support leaders can also customize AXIS by weighting the components differently to emphasize certain aspects more than others.

Here are some examples of when a support leader might care about one component over another:

Testing new AI tools

When implementing or testing new AI customer service agent tools, companies can give higher weight to the **Resolution Accuracy (RA)** component.

Reducing customer effort

For organizations aiming to streamline customer support interactions and minimize customer effort, they can place greater emphasis on the **Interaction Effort (IE)** component.

Optimizing internal routing

Organizations prioritizing quick and efficient routing to the right internal team can give more weight to the **Handoff Smoothness (HS)** component.

The weighted AXIS score can be calculated as follows:

$$AXIS = \frac{[(w1 \times RA) + (w2 \times IE) + (w3 \times HS)]}{(w1 + w2 + w3)}$$

Where w1, w2, and w3 are the weights assigned to RA, IE, and HS respectively, depending on their importance in the specific context of the analysis.

This gives a balanced assessment that can be tailored to prioritize certain aspects of the AI's interaction quality based on the goals and priorities of the support team.

Front

Customized AXIS score example

Let's apply the AXIS formula using an example with specific scores and weights for each component:

To calculate a weighted AXIS score, you would assign a weight (between 1 and 5) to each component based on its relative importance in your context.

For example:
- If you require AI to quickly identify the customer issue, you might give **RA** a higher weight (w1 = 4).
- If your workflow typically requires a few customer confirmation steps, you could give **IE** a lower weight (w2 = 3).
- If seamless handoffs are essential, **HS** might receive a higher weight (w3 = 5).

This resulting AXIS score of 4.17 out of 5 indicates a generally high-quality AI interaction in terms of resolution accuracy, customer effort, and handoff smoothness, with particular strength in handoff smoothness.

The weighted AXIS score would be calculated as:

$$\text{AXIS} = \frac{[(w1 \times RA) + (w2 \times IE) + (w3 \times HS)]}{(w1 + w2 + w3)}$$

$$\text{AXIS} = \frac{[(4 \times 4) + (3 \times 3) + (5 \times 5)]}{(4 + 3 + 5)}$$

$$= \frac{(16 + 9 + 25)}{12}$$

$$= \frac{50}{12} \approx 4.17$$

Front

AXIS generative AI prompt

Note: Copy and paste these prompts into a generative AI tool like ChatGPT or Claude alongside a screenshot or export of your AI-led customer support interaction to calculate your AXIS score. Be sure to exclude any sensitive customer information and follow your company's data privacy policy before using generative AI. Always review the output to confirm the interaction was accurately evaluated.

INDUSTRY-STANDARD VERSION

Hello! Evaluate the uploaded interaction using the AXIS framework by scoring the following components on a scale from 1 to 5, where 1 is poor and 5 is excellent:

Resolution Accuracy (RA) - Measures the AI's ability to correctly identify and address the customer's issue:
- 5 – Exceptional: The AI resolves the issue with total accuracy on the first attempt without errors or unnecessary steps.
- 4 – High: Minor clarifications needed, but the AI accurately resolves the issue with minimal extra input.
- 3 – Moderate: Partial resolution, requires more customer details or the solution isn't entirely accurate.
- 2 – Low: Multiple interactions needed to get a partially correct response.
- 1 – Poor: Incorrect answers or irrelevant suggestions, leading to frustration or handoff.

Interaction Effort (IE) - Assesses the ease and intuitiveness of the AI interaction:
- 5 – Effortless: Quick and clear solution with minimal back-and-forth.
- 4 – Easy: Minimal effort with one or two follow-up questions necessary.
- 3 – Moderate Effort: Several responses or clarifications needed to reach a resolution.
- 2 – High Effort: Many questions or steps before achieving a resolution.
- 1 – Very High Effort: Numerous repetitive prompts, hard to achieve resolution.

Handoff Smoothness (HS) - Evaluates the transition from AI to human agent:
- 5 – Seamless: Flawless handoff; no repetition of details, natural and timely transition.
- 4 – Smooth: Minimal disruption; slight repetition but does not detract from the experience.
- 3 – Moderate: Inefficient transition; partial information, some repetition required.
- 2 – Rough: Problematic handoff; little or no context, noticeable friction.
- 1 – Poor: Very poor or failed handoff; incorrect transfer or lack of support.

Calculate the AXIS score by adding the scores for RA, IE, and HS and then dividing by 3. This will provide an overall measure of the AI interaction quality. Please share your results with us. Thank you!

AXIS generative AI prompt

Note: Enter your weight classification before copying and pasting these instructions into the chat.

CUSTOM WEIGHTED VERSION

Hello! Evaluate the uploaded interaction using the AXIS framework by scoring the following components on a scale from 1 to 5, where 1 is poor and 5 is excellent:

Resolution Accuracy (RA) - Measures the AI's ability to correctly identify and address the customer's issue:
- 5 – Exceptional: The AI resolves the issue with total accuracy on the first attempt without errors or unnecessary steps.
- 4 – High: Minor clarifications needed, but the AI accurately resolves the issue with minimal extra input.
- 3 – Moderate: Partial resolution, requires more customer details or the solution isn't entirely accurate.
- 2 – Low: Multiple interactions needed to get a partially correct response.
- 1 – Poor: Incorrect answers or irrelevant suggestions, leading to frustration or handoff.

Interaction Effort (IE) - Assesses the ease and intuitiveness of the AI interaction:
- 5 – Effortless: Quick and clear solution with minimal back-and-forth.
- 4 – Easy: Minimal effort with one or two follow-up questions necessary.
- 3 – Moderate Effort: Several responses or clarifications needed to reach a resolution.
- 2 – High Effort: Many questions or steps before achieving a resolution.
- 1 – Very High Effort: Numerous repetitive prompts, hard to achieve resolution.

Handoff Smoothness (HS) - Evaluates the transition from AI to human agent:
- 5 – Seamless: Flawless handoff; no repetition of details, natural and timely transition.
- 4 – Smooth: Minimal disruption; slight repetition but does not detract from the experience.
- 3 – Moderate: Inefficient transition; partial information, some repetition required.
- 2 – Rough: Problematic handoff; little or no context, noticeable friction.
- 1 – Poor: Very poor or failed handoff; incorrect transfer or lack of support.

Use these criteria to determine the scores for RA, IE, and HS. Then, apply the formula with weights if applicable, to calculate the final AXIS score: $AXIS = [(w1 \times RA) + (w2 \times IE) + (w3 \times HS)] / (w1 + w2 + w3)$

Where:
- w1, w2, and w3 are the weights for RA, IE, and HS respectively and they are defined as follows: **{{enter custom weights}}.**

If weights are not specified, use equal weighting for simplicity.

This will provide an overall measure of the AI interaction quality. Please share your results with us. Thank you!

Next steps for customer support leaders after knowing their AXIS score

○ **Interpret and analyze your score:**

- Review your AXIS score to identify strengths and weaknesses
- Reflect on how your score aligns with your organizational goals

○ **Set specific, measurable, achievable, relevant, and time-bound (SMART) goals:**

- Use SMART goals to address low-scoring areas
- Align your goals with team objectives and broader company priorities

○ **Seek feedback:**

- Utilize 360° input from customers and staff to validate your findings
- Engage with support team for guidance in improving targeted areas

○ **Foster cross-functional collaboration:**

- Promote collaboration to address inter-departmental challenges
- Create new or improved strategies, tools, or approaches to effectively address and bridge gaps

○ **Monitor progress and iterate:**

- Regularly review and refine your strategies based on updated AXIS scores on a monthly, quarterly, and annual basis

APR ● 4.0

MAY ● 3.67

FEB ● 3.0

MAR ● 2.67

JAN ● 2.33

⟶ Average AXIS across total conversations
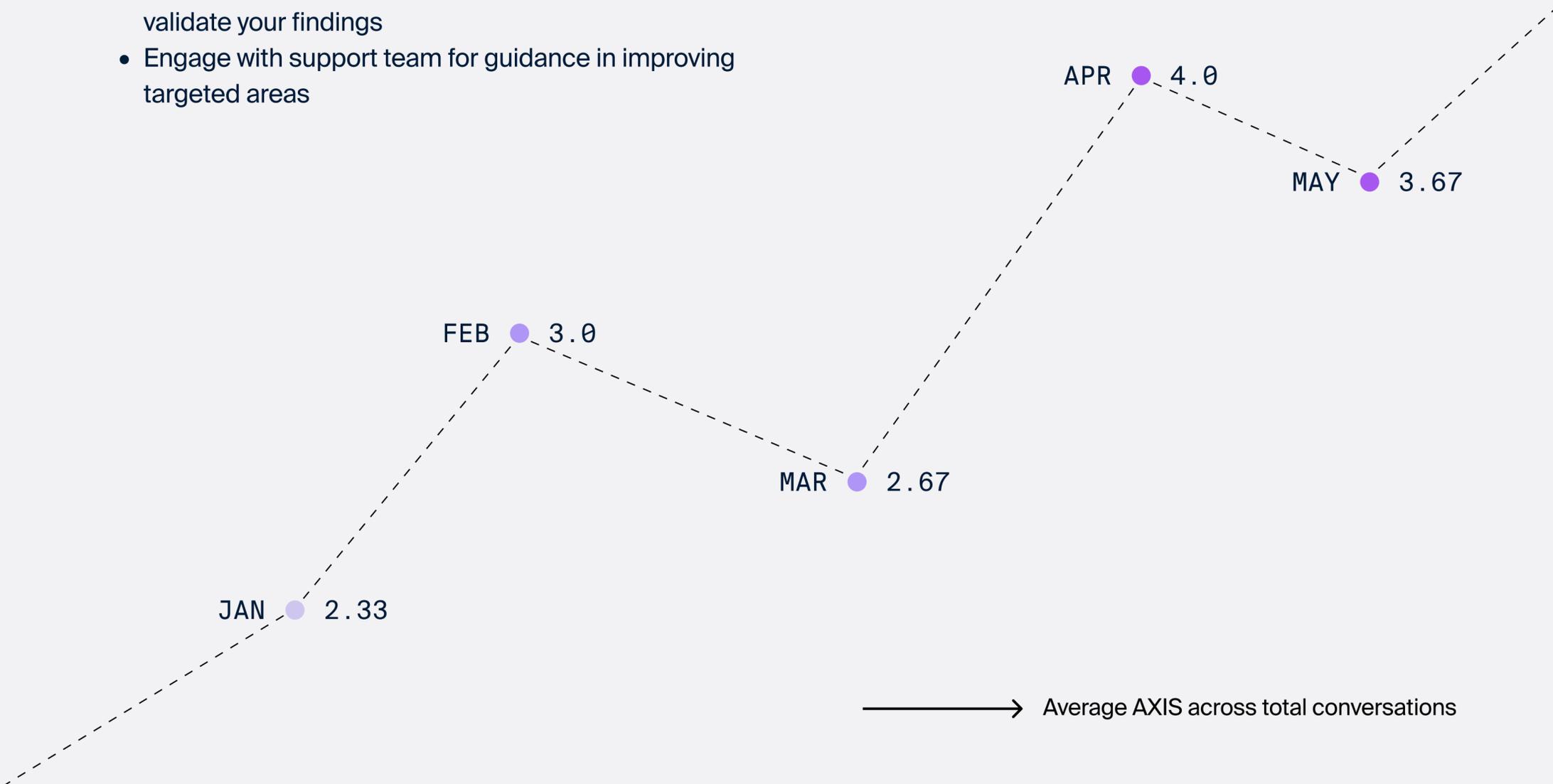
Front

Improving the breaking points of AI-led interactions at Front

We're currently rolling out AXIS to our own customer conversations here at Front. While it's still a manual process since we're in the early discovery phase, we're excited to see what the data will reveal. By adding actionable, quantitative insights, we'll know the *why* behind our deflections to make our customer experience the best it can be.

Future research will focus on refining AXIS weights and expanding the framework to accommodate different AI technologies and interaction models, adapting to diverse operational needs across industries.

If you try AXIS with your team and have any feedback or suggestions, feel free to share them over at Front Community.

Front is the all-in-one customer service platform built for collaboration.

We help over 9,000 businesses give 5-star service by bringing core support channels in one intuitive, AI-powered workspace where teams can build collaborative workflows across the entire customer lifecycle.

As your customer base and its needs grow, Front's AI features and customer intelligence serve up the efficiency and insights needed to keep your entire organization customer-first, every day. Learn more at front.com.