

Create a single response test suite.

Evaluating individual responses is very useful when you want to test your agents on how they answer specific questions, on the functions they call, and on the precise wording they use in their responses.

Single-response reviews examine your agent on each individual question, rather than the entire conversation. For example, a single-response review for a customer service agent asks, "What are the company's business hours?", records the agent's answer to that question, and then begins with a new question, "How can I find my order history?".

Single-response evaluations are useful when you want to test your agents on how they respond to specific questions, the functions they call, and the precise wording they use in their responses. You can also run conversational evaluations, allowing you to assess agent behavior over a longer interaction.

The reviews utilize test suites. A single-action review test suite includes a group of up to 100 test cases. When running an agent review, you select a test suite, and Copilot Studio will run all the test cases in that suite against your agent.

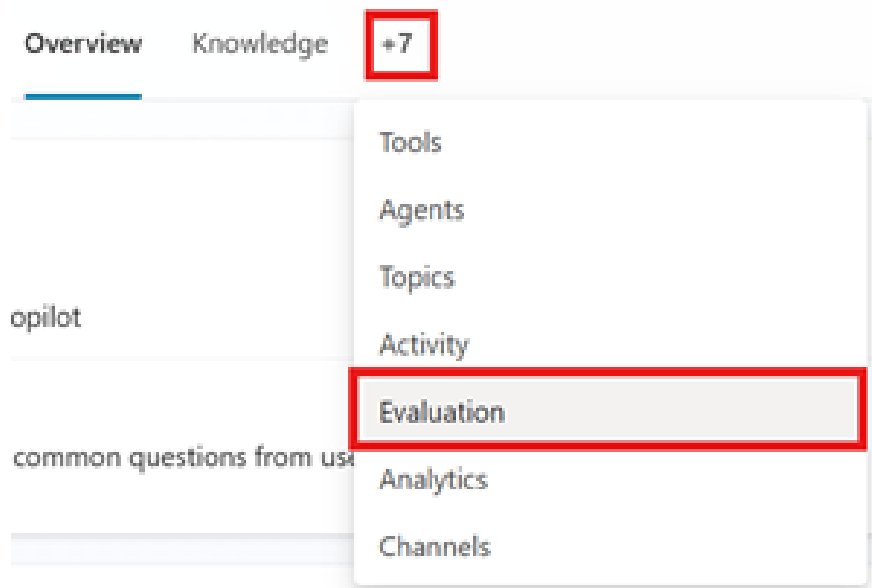
You can create test cases in a test suite manually, import them using a spreadsheet, or use AI to generate messages based on the agent's design and resources. Then, you can choose how you want to measure the quality of the agent's response for each test case in the test suite.

For more information on how agent evaluation works, see [Agent Evaluation Overview](#) .

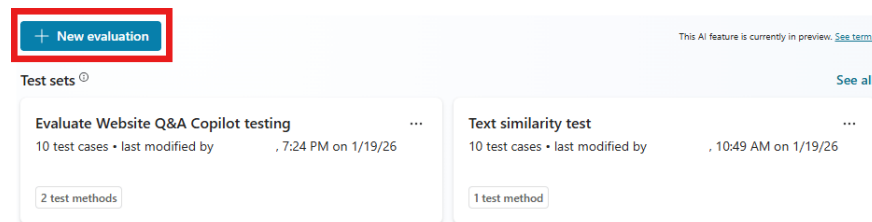
Important note : Test results will be available in Copilot Studio for 89 days. To save your test results for a longer period, export them to a CSV file.

Create a new test toolkit.

1. Access the agent's **Evaluation** page .



2. Select **New evaluation > Single response** .

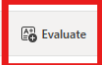


3. Choose the method you want to use to create your test suite. A test suite can have up to 100 test cases.

1. **The Quick Question Set** allows Copilot Studio to automatically generate test cases based on the agent's description, instructions, and capabilities. This option creates 10 questions for running small, quick assessments or for starting to build a larger test suite.
2. **A full question set** allows Copilot Studio to create test cases using knowledge sources or topics and select the number of questions to generate.
3. **Use your test chat conversation** to automatically populate the test suite with the questions you provided in the test chat. This method uses questions from the most recent test chat. You can also start the assessment from the test chat by using the rating icon.



1. Import test cases from a file by dragging your file into the designated area, selecting **Browse** to upload the file, or choosing one of the other upload options.
2. Alternatively, you can write some questions yourself to manually create a test suite. Follow the steps to edit the test suite to add and modify test cases.
3. Utilize production data based on topics from the agent's analysis.

Name		Total questions	Answered questions	Response quality	Reactions
Troubleshooting	 Evaluate	499	39%	51%	👍 157 👎 184
Billing & Payments		474	18%	49%	👍 16 👎 224

4. In the **Name** field , enter a name for your test suite.

5. Change or add the testing methods you want to use:

1. Adding a new method:

1. Select **Add test method** .
2. Select all the methods you want to test, then select **OK** .
3. Some methods require passing scores. Passing scores determine which points lead to passing or failing. Set a score, then select **OK** .
4. Some methods require adding expected responses or keywords to each of your test cases.

2. Choose an existing test method to edit or delete.

Testing method	Measurement	Type of test kit	Scoring	Configuration
General quality	How good is the test case response based on specific characteristics?	Single response or conversation	Score on a 100% scale	Are not
Connect oven	The degree of agreement between the meaning of the answer in the test case and the expected answer.	Single response	Score on a 100% scale	Passing score, expected answer
Capability use	Does the test case utilize all or any of the resources expected?	Single response	Pass/Fail	Predictability
Keyword match	Does the test case use all or any of the expected keywords or phrases?	Single response or conversation	Pass/Fail	Expected keywords or phrases
Text similarity	The degree of agreement between the content of the test answer and the expected answer.	Single response	Score on a 100% scale	Passing score, expected answer
Exact match	Do the results of the test case exactly match the expected results?	Single response	Pass/Fail	Expected answer
Custom	Does the response from the test case meet your defined criteria or expectations?	Single response or conversation	Pass/Fail (meets the defined label criteria)	Name, assessment guidelines, label

6. Refine the test case details. All test methods, except for general quality, require expected feedback or keywords.

7. Select **User profile** , then choose or add the account you want to use for this test suite, or proceed without authentication. The evaluation process uses this account to connect to knowledge sources and tools during testing. If you select an account other than the one with the authentication connection for evaluation, the agents using the connector or tools will fail.

Note : Automated testing uses the authentication of the selected test account. If your agent has knowledge sources or connections that require specific authentication, select the appropriate account for your testing process. When Copilot Studio creates test cases, it uses the authentication information of the connected account

to access the agent's knowledge sources and tools. The created test cases may include sensitive data that the connected account can access. Any manufacturer with access to the agent can also view the test toolkits associated with that agent.

8. Select **Save** to update the test suite without running test cases, or **Evaluate** to run the test suite immediately.

Limitations on creating test cases

The test case creation will fail if one or more questions violate the agent's content moderation settings. Possible reasons include:

1. The agent's instructions or themes cause the model to generate content that the system flags.
2. The connected knowledge sources include sensitive or restricted content.
3. The agent's content moderation settings are too strict.

To resolve the issue, try different actions, such as adjusting the knowledge source, updating the guidelines, or modifying moderation settings.

A test suite can contain up to 100 test cases.

Create test kits from knowledge or topics.

You can test your agent by creating questions using information and conversational resources that your agent already possesses. This testing method is good for testing how your agent utilizes existing knowledge or topics, but not good for testing for information gaps.

You can create test cases using the following knowledge sources:

1. Document
2. Microsoft Word
3. Microsoft Excel
4. PDF file
5. SharePoint content

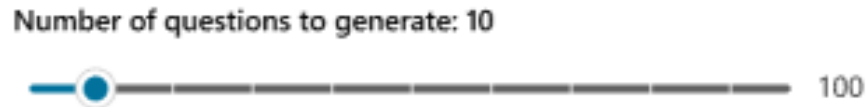
You can use files with a maximum size of 5 MB to create test questions.

To create a test kit:

1. In the **New evaluation** section , select **Full question set** .
2. Choose **Knowledge** or **Topics** .
 1. **Knowledge** works best for agents using orchestration generation. This method generates questions using one of the agent's knowledge sources.
 2. **Topics** work best for agents using classic dispatching. This method generates questions by using the agent's topics.
3. For **Knowledge** , select the knowledge source you want to use to create the question.

Name	Type	Last edited
<input checked="" type="radio"/> Copilot-STudio-build-agent-intro.docx	Word Document	2 months ago
<input type="radio"/> Choose the right Copilot Studio experience.docx	Word Document	2 months ago

4. For **Knowledge** and **Topics** , select and drag the slider to choose the number of questions to create.



5. Select **Generate** .

6. In the **Name** field , enter a name for your test suite.

7. Change or add the testing methods you want to use:

1. Adding a new method:
 1. Select **Add test method** .
 2. Select all the methods you want to test, then select **OK** . You can add more methods.
 3. For some methods, set a pass score, then select **OK** . The pass score will determine which method results in a pass or a fail.
 4. Some methods require adding expected answers or keywords to each of your test cases.
2. Choose an existing test method to edit or delete.

8. Modify the details of the test cases. All test cases using the methods, except for general quality, require an expected response.

9. Select **Save** to update the test suite without running test cases, or select **Evaluate** to run the test suite immediately.

Create a test toolkit file for import.

Instead of directly building your test cases in Copilot Studio, you can create a spreadsheet file with all your test cases and import them to create the test suite. You can draft each test question, define the testing methods you want to use, and specify the expected answers for each question. Once you've finished creating the file, save it as a .csv or .txt file and import it into Copilot Studio.

Important note :

1. The file can contain up to 100 questions.
2. Each question can be up to 1,000 characters long, including spaces.
3. The file must be in comma-separated value (CSV) or text format.

To create an import file:

1. Open a spreadsheet application (e.g., Microsoft Excel). You can download the CSV template in the **Data source** section after selecting **New evaluation** .
2. Add the following headings, in this order, to the first row:
 1. Question
 2. Expected answer
 3. Testing methods
3. Enter your test questions in the **Question** column . Each question can be up to 1,000 characters long, including spaces.
4. Enter one of the following testing methods for each question in the **Testing method** column :
 1. Overall quality
 2. Compare the meanings
 3. Similarity
 4. Precision fit
 5. Keyword Matching
5. Enter the expected responses for each question in the **Expected response** column . Expected responses are optional when importing the test suite. However, you need expected responses to run the match, similarity, and significance test cases.
6. Save the file as a .csv or .txt file.
7. Import the file following these steps in order.

Create a theme-based testing toolkit.

Create a test suite with questions derived from conversations with real users. This method utilizes themes found in the agent's analytics section.

Themes are groups of questions taken from a user question bank, triggering automated responses. When creating a test suite using themes, you will generate test cases from user-submitted questions related to that theme.

Use these test tools to conduct assessments focused on a specific area or topic within the agent's scope of work. For example, if you have customer service agents, you could track the quality of answers to payment and billing questions separately from other use cases such as troubleshooting.

Note : Before creating a test suite from themes, you need to have access to the themes in the analytics section.

1. On the agent's **Analytics** page , go to the **Themes** list .
2. Hover your mouse over a theme, then select **Evaluate** .

Themes PREVIEW See all (19)

Analyze your users' questions by auto-classified themes, and see how well your agent responds. [Learn more](#) [Preview terms](#)

Name		Total questions	Answered questions	Response quality	Reactions
<u>Troubleshooting</u>	Evaluate	499	39%	51%	157 184
Billing & Payments		474	18%	49%	16 224
Feature Requests		367	78%	44%	98 64
Mobile App Support		334	82%	45%	76 144
Order Status		310	59%	74%	108 103

AI-generated content may be incorrect.

You can also select **See all** to view more themes, then select **Evaluate** .

3. Select **Create and open** .

4. Modify the details of the test toolkits and test cases. All test cases using the method, except for general quality testing, require expected feedback.

5. Select **Save** to update the test suite without running test cases, or select **Evaluate** to run the test suite immediately.

You finished reading the article "**Create a single response test suite.**" edited by the [TipsMake](#) team. We hope this article has provided you with many useful tech tips and tricks. You can search for similar articles on tips and guides. Thank you for reading and for following us regularly.