ChatGPT Automation Execution Results

Complexity	Coverage
- Input: 26 Questions - Context: 14 Questions - Total: 40 Questions	- Category Partition Coverage - Input Coverage - Context Coverage - Output Coverage
Bugs	**QA Metrics**
8 Bugs detected	

QA Metrics	Weight
Domain Knowledge	35%
Keyword-Based Similarity	35%
Language-Based Similarity	30%

- Highest: 98.13%
- Lowest: 24.14%
- Average: 84.51%

Conventional Testing	Automated Testing
Complexity	**Complexity**
- Equivalence Partition: 16 Questions - Decision Table: 12 Questions - Boundary Value: 6 Questions - Scenario: 3 Questions - Total: 37 Questions	- Input: 26 Questions - Context: 14 Questions - Total: 40 Questions
Coverage	**Coverage**
- Equivalence Partition Coverage - Decision Table Coverage - Boundary Value Coverage - Scenario Coverage	- Category Partition Coverage - Input Coverage - Context Coverage - Output Coverage
QA Metrics	**QA Metrics**
Bugs	**Bugs**
7 bugs detected	8 bugs detected

QA Metrics (Conventional Testing):

QA Metrics	Weight
Domain Knowledge	80%
Language-Based Similarity	20%

QA Metrics (Automated Testing):

QA Metrics	Weight
Domain Knowledge	35%
Keyword-Based Similarity	35%
Language-Based Similarity	30%

Gemini AI Automation Execution Results

Complexity	Coverage
- Input: 26 Questions - Context: 14 Questions - Total: 40 Questions	- Category Partition Coverage - Input Coverage - Context Coverage - Output Coverage
Bugs	**QA Metrics**
9 bugs detected	(see table below)

QA Metrics	Weight
Domain Knowledge	35%
Keyword-Based Similarity	35%
Language-Based Similarity	30%

- Highest: 89.76%
- Lowest: 16.85%
- Average: 70.31%

Conventional Testing	Automated Testing
Complexity	**Complexity**
- Equivalence Partition: 16 Questions - Decision Table: 12 Questions - Boundary Value: 6 Questions - Scenario: 3 Questions - Total: 37 Questions	- Input: 26 Questions - Context: 14 Questions - Total: 40 Questions
Coverage	**Coverage**
- Decision Table Coverage - Equivalence Partition Coverage - Boundary Value Coverage - Scenario Coverage	- Category Partition Coverage - Input Coverage - Context Coverage - Output Coverage
QA Metrics	**QA Metrics**

QA Metrics	Weight
Domain Knowledge	80%
Language-Based Similarity	20%

QA Metrics	Weight
Domain Knowledge	35%
Keyword-Based Similarity	35%
Language-Based Similarity	30%

Conventional Testing	Automated Testing
Bugs	**Bugs**
11 bugs detected	9 bugs detected

Perplexity AI Automation Execution Results

Complexity	Coverage
- Input: 26 Questions - Context: 14 Questions - Total: 40 Questions	- Category Partition Coverage - Input Coverage - Context Coverage - Output Coverage
Bugs	**QA Metrics**
16 Bugs detected	(see below)

QA Metrics table:

QA Metrics	Weight
Domain Knowledge	35%
Keyword-Based Similarity	35%
Language-Based Similarity	30%

- Highest: 67.5%
- Lowest: 40%
- Average: 63.44%

Conventional Testing	Automated Testing
Complexity	**Complexity**
- Equivalence Partition: 16 Questions - Decision Table: 12 Questions - Boundary Value: 6 Questions - Scenario: 3 Questions - Total: 37 Questions	- Input: 26 Questions - Context: 14 Questions - Total: 40 Questions
Coverage	**Coverage**
- Decision Table Coverage - Equivalence Partition Coverage - Boundary Value Coverage - Scenario Coverage	- Category Partition Coverage - Input Coverage - Context Coverage - Output Coverage
QA Metrics	**QA Metrics**
<table><tr><th>QA Metrics</th><th>Weight</th></tr><tr><td>Domain Knowledge</td><td>80%</td></tr><tr><td>Language-Based Similarity</td><td>20%</td></tr></table>	<table><tr><th>QA Metrics</th><th>Weight</th></tr><tr><td>Domain Knowledge</td><td>35%</td></tr><tr><td>Keyword-Based Similarity</td><td>35%</td></tr><tr><td>Language-Based Similarity</td><td>30%</td></tr></table>
Bugs	**Bugs**
3 Bugs detected	16 Bugs detected

Kindroid Automation Execution Results

Complexity	Coverage
- Input: 26 Questions - Context: 14 Questions - Total: 40 Questions	- Category Partition Coverage - Input Coverage - Context Coverage - Output Coverage
Bugs	**QA Metrics**
13 Bugs detected	(see table below)

QA Metrics	Weight
Domain Knowledge	35%
Keyword-Based Similarity	35%
Language-Based Similarity	30%

- Highest: 79%
- Lowest: 2%
- Average: 37%

Conventional Testing	Automated Testing
Complexity	**Complexity**
- Equivalence Partition: 16 Questions - Decision Table: 12 Questions - Boundary Value: 6 Questions - Scenario: 3 Questions - Total: 37 Questions	- Input: 26 Questions - Context: 14 Questions - Total: 40 Questions
Coverage	**Coverage**
- Decision Table Coverage - Equivalence Partition Coverage - Boundary Value Coverage - Scenario Coverage	- Category Partition Coverage - Input Coverage - Context Coverage - Output Coverage
QA Metrics	**QA Metrics**

Conventional Testing QA Metrics:

QA Metrics	Weight
Domain Knowledge	80%
Language-Based Similarity	20%

Automated Testing QA Metrics:

QA Metrics	Weight
Domain Knowledge	35%
Keyword-Based Similarity	35%
Language-Based Similarity	30%

Conventional Testing	Automated Testing
Bugs	**Bugs**
14 bugs detected	13 Bugs detected

Character.ai Automation Execution Results

Complexity	Coverage
- Input: 26 Questions - Context: 14 Questions - Total: 40 Questions	- Category Partition Coverage - Input Coverage - Context Coverage - Output Coverage
Bugs	**QA Metrics**
32 Bugs detected	see table below

QA Metrics	Weight
Domain Knowledge	35%
Keyword-Based Similarity	35%
Language-Based Similarity	30%

- Highest: 72.12%
- Lowest: 11.95%
- Average: 64.41%

Conventional Testing	Automated Testing
Complexity	**Complexity**
- Equivalence Partition: 16 Questions - Decision Table: 12 Questions - Boundary Value: 6 Questions - Scenario: 3 Questions - Total: 37 Questions	- Input: 26 Questions - Context: 14 Questions - Total: 40 Questions
Coverage	**Coverage**
- Decision Table Coverage - Equivalence Partition Coverage - Boundary Value Coverage - Scenario Coverage	- Category Partition Coverage - Input Coverage - Context Coverage - Output Coverage
QA Metrics	**QA Metrics**
<table><tr><th>QA Metrics</th><th>Weight</th></tr><tr><td>Domain Knowledge</td><td>80%</td></tr><tr><td>Language-Based Similarity</td><td>20%</td></tr></table>	<table><tr><th>QA Metrics</th><th>Weight</th></tr><tr><td>Domain Knowledge</td><td>35%</td></tr><tr><td>Keyword-Based Similarity</td><td>35%</td></tr><tr><td>Language-Based Similarity</td><td>30%</td></tr></table>
Bugs	**Bugs**
14 Bugs detected	32 Bugs Detected

Total Chatbot Testing Results

Conventional Testing	Automated Testing
## Complexity	## Complexity
- Equivalence Partition: 16 Questions - Decision Table: 12 Questions - Boundary Value: 6 Questions - Scenario: 3 Questions - Total: 37 Questions	- Input: 26 Questions - Context: 14 Questions - Total: 40 Questions
## Coverage	## Coverage
- Decision Table Coverage - Equivalence Partition Coverage - Boundary Value Coverage - Scenario Coverage	- Category Partition Coverage - Input Coverage - Context Coverage - Output Coverage
## QA Metrics	## QA Metrics

Conventional Testing – QA Metrics

QA Metrics	Weight
Domain Knowledge	80%
Language-Based Similarity	20%

Automated Testing – QA Metrics

QA Metrics	Weight
Domain Knowledge	35%
Keyword-Based Similarity	35%
Language-Based Similarity	30%

Total Questions	Total Questions
185 Questions	200 Questions
## Bugs	## Bugs
49 Bugs detected	78 Bugs detected