

## **Appendix A: Complete Evaluation Instruments**

This document provides the complete evaluation instruments used in the educational evaluation of the BattleInTheSky project, as referenced in Section 6 of the article "BattleInTheSky: A Gamified Expert System for Accessible and Ethical AI Education Using Raspberry Pi Pico W." These instruments assess learning, motivation, perceived usefulness, and artifact quality among 15 Mechatronics Engineering students at a Mexican university in Spring 2025.

### **1. Pre-Test and Post-Test (Individual)**

**Purpose:** To measure students' understanding of expert systems and related AI concepts before and after the BattleInTheSky project, quantifying learning gains.

**Format:**

- 20 questions total:
  - 10 Multiple Choice (3 points each, 30 points total)
  - 5 Short Answer (8 points each, 40 points total)
  - 5 Problem-Solving (11 points each, 55 points total, adjusted to 30 points for a 100-point total)

**Scoring:**

- Total: 100 points
- Multiple Choice: Correct (3 points) or incorrect (0 points)
- Short Answer and Problem-Solving:
  - High (full points): Complete, accurate, well-explained
  - Medium (partial points): Partially correct or incomplete, minor errors
  - Low (few points): Largely incorrect or incomplete, poor understanding
  - Zero (0 points): No answer or irrelevant

**Key Concepts Covered:**

- Definition and components of expert systems (knowledge base, inference engine, user interface)
- Rule-based systems and IF-THEN logic
- Forward and backward chaining
- AI applications in embedded systems
- Basic programming concepts
- Ethical considerations in AI (bias, transparency)

## Questions:

Multiple Choice (3 points each):

1. Which component of an expert system contains the domain knowledge? a) User interface b) Knowledge base c) Inference engine d) Database Answer: b
2. What type of reasoning does forward chaining use? a) Data-driven b) Goal-driven c) Rule-driven d) Logic-driven Answer: a
3. In an IF-THEN rule, what does the "IF" part represent? a) The action b) The condition c) The conclusion d) The result Answer: b
4. Which component interprets and applies knowledge to make decisions? a) Knowledge base b) User interface c) Inference engine d) Explanation facility Answer: c
5. What is the main advantage of rule-based systems? a) Ability to learn from data b) Explainability of decisions c) Handling of uncertainty d) Processing large amounts of data Answer: b
6. Which is NOT a characteristic of embedded systems? a) Limited resources b) Real-time operation c) Unlimited processing power d) Dedicated function Answer: c
7. What programming language is commonly used with Raspberry Pi Pico? a) C++ b) Java c) MicroPython d) JavaScript Answer: c
8. What does GPIO stand for? a) General Purpose Input Output b) Graphical Processing Input Output c) Global Positioning Input Output d) General Programming Input Output Answer: a
9. Which is a potential ethical concern in AI development? a) Increased automation b) Algorithmic bias c) Improved efficiency d) Faster processing Answer: b
10. What is algorithmic transparency? a) The speed of an algorithm b) The ability to understand how an algorithm makes decisions c) The cost of developing an algorithm d) The complexity of an algorithm Answer: b

Short Answer (8 points each):

1. Explain the role of the inference engine in an expert system. Expected Answer: The inference engine applies rules and facts from the knowledge base to deduce solutions or make decisions, using techniques like forward or backward chaining.
2. Describe a real-world application of an expert system. Expected Answer: Examples include medical diagnosis (e.g., MYCIN), financial advising, or industrial control systems, using knowledge to perform tasks.
3. What are the potential ethical concerns when developing an expert system? Expected Answer: Concerns include data bias, lack of decision transparency, job displacement, and accountability for errors.

4. Explain the difference between a knowledge base and a database. Expected Answer: A knowledge base stores rules and facts for reasoning, often symbolically; a database stores structured data for retrieval.
5. Describe the basic structure of an IF-THEN rule. Expected Answer: An IF-THEN rule has a condition (IF) and an action (THEN); if the condition is true, the action is executed.

Problem-Solving (11 points each):

1. Write an IF-THEN rule in Python to turn on an LED connected to pin 16 of a Raspberry Pi Pico if a temperature sensor reads above 25°C (variable: temp). Expected Answer: if temp > 25:  
Code to turn on LED on pin 16  
print("Temperature is above 25 degrees, turning on LED") Evaluation: Correct syntax, logical condition, appropriate action.
2. Design a knowledge base for: "If it is raining, the ground is wet. If the sprinkler is on, the ground is wet. If the ground is wet, plants will grow." Expected Answer: RULE 1: IF raining THEN ground\_is\_wet RULE 2: IF sprinkler\_on THEN ground\_is\_wet RULE 3: IF ground\_is\_wet THEN plants\_grow Evaluation: Accurate rule representation.
3. Explain how to use a sensor and actuator in a feedback control system with IF-THEN rules, with an example. Expected Answer: A sensor (e.g., temperature) triggers an actuator (e.g., heater) via rules. Example: IF temperature < 20°C THEN turn\_on\_heater. Evaluation: Clear feedback loop and relevant example.
4. Write Python code to read a button state on a Raspberry Pi Pico and print "Button pressed" if pressed. Expected Answer:  
Code to read button state  
if button\_state == 1: # Assuming 1 is pressed print("Button pressed") Evaluation: Correct logic and conditional action.
5. Describe how an expert system could control a robot navigating an obstacle course, including rules. Expected Answer: Sensors detect obstacles; rules dictate actions (e.g., IF obstacle\_ahead THEN turn\_right). Evaluation: Logical control process and relevant rules.

## 2. Motivation Survey (Individual)

**Purpose:** To measure students' motivation levels related to the BattleInTheSky project, focusing on enjoyment, confidence, and ownership.

**Format:**

- 15 items, 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree)

- Subscales (5 items each):
  - Intrinsic Motivation: Enjoyment and interest
  - Perceived Competence: Confidence in abilities
  - Autonomy: Sense of freedom and ownership

**Scoring:**

- Each item: 1–5 points
- Subscale score: Average of 5 items
- Overall score: Average of all 15 items

**Questions:**

Intrinsic Motivation:

1. I found the BattleInTheSky project enjoyable. (1 2 3 4 5)
2. I was interested in learning more about expert systems through the game. (1 2 3 4 5)
3. The project was fun and engaging. (1 2 3 4 5)
4. I looked forward to working on the game. (1 2 3 4 5)
5. I would like to do similar projects in the future. (1 2 3 4 5)

Perceived Competence:

6. I feel confident in my ability to design expert systems. (1 2 3 4 5)
7. I was able to successfully implement the game's rules. (1 2 3 4 5)
8. I understood the AI concepts involved in the project. (1 2 3 4 5)
9. I was able to overcome challenges I encountered. (1 2 3 4 5)
10. I feel my AI skills have improved. (1 2 3 4 5)

Autonomy:

11. I had enough freedom to design the game in my own way. (1 2 3 4 5)
12. I felt a sense of ownership over my project. (1 2 3 4 5)
13. I was able to make my own decisions about the game's design. (1 2 3 4 5)
14. I felt my creativity was encouraged. (1 2 3 4 5)
15. I had control over how I approached the project. (1 2 3 4 5)

**Answer Format:**

- Students circle a number from 1 to 5:
  - 1 = Strongly Disagree
  - 2 = Disagree
  - 3 = Neutral
  - 4 = Agree

- 5 = Strongly Agree

### **3. Usefulness Survey (Individual)**

Purpose: To measure how useful students perceived the BattleInTheSky project as a learning tool, assessing clarity, relevance, engagement, and skill development.

#### ***Format:***

- 10 items, 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree)
- 2 open-ended questions

#### ***Scoring:***

- Likert items: 1–5 points, averaged per student
- Open-ended: Free-text responses, analyzed thematically

#### ***Key Areas:***

- Clarity of concepts
- Relevance to studies
- Comparison to other methods
- Engagement
- Skill development

#### ***Questions:***

Likert Items:

1. The BattleInTheSky project helped me understand the practical applications of expert systems. (1 2 3 4 5)
2. The game improved my programming skills. (1 2 3 4 5)
3. I would recommend this project to other students. (1 2 3 4 5)
4. The game was a valuable learning tool. (1 2 3 4 5)
5. The instructions for the project were clear. (1 2 3 4 5)
6. The project helped me to think critically. (1 2 3 4 5)
7. The game increased my interest in AI. (1 2 3 4 5)
8. The project was relevant to my studies. (1 2 3 4 5)
9. The game was easy to use. (1 2 3 4 5)
10. The project was well-organized. (1 2 3 4 5)

Open-Ended Questions:

1. What were the most helpful aspects of the BattleInTheSky project for your learning?

2. What suggestions do you have for improving the project?

Answer Format:

- Likert items: Students circle a number from 1 to 5 (same scale as Motivation Survey)
- Open-ended: Students provide written responses

#### **4. Game Artifact Analysis (Team)**

**Purpose:** To evaluate the quality of the expert system game created by each team, assessing technical and creative aspects.

**Format:**

- Rubric with 4 criteria, each scored 1–5 points (total 20 points)

**Criteria:**

- Rule Logic: Correctness, completeness, efficiency of IF-THEN rules.
- Code Clarity: Readability, organization, documentation of the code.
- Functionality: Successful implementation of the game mechanics.
- Design: Creativity and effectiveness of the game design (within the given constraints).

**Scoring Details:**

Rule Logic:

- 5: Correct, complete, efficient, well-organized
- 4: Mostly correct, minor inefficiencies
- 3: Some errors or omissions affecting logic
- 2: Significant errors, poorly organized
- 1: Largely incorrect or missing

Code Clarity:

- 5: Well-organized, readable, documented
- 4: Mostly clear, minor issues
- 3: Somewhat disorganized, difficult to read
- 2: Poorly organized, lacks documentation
- 1: Unreadable, undocumented

Functionality:

- 5: All mechanics implemented, function smoothly

- 4: Most mechanics function, minor bugs
- 3: Some mechanics missing or have significant bugs
- 2: Major functionality missing or broken
- 1: Largely non-functional

Design:

- 5: Creative, engaging, effective
- 4: Generally good, lacks some creativity
- 3: Basic, uninspired
- 2: Poor, detracts from experience
- 1: Non-existent or detrimental

Answer Format:

- Evaluator (instructor) assigns a score from 1 to 5 for each criterion based on the team's submitted game artifact.