- **What does it mean to say that "trust is a construct" in the context of AI? How does this perspective influence the design of trustworthy AI systems?**
    - Sample Answer:

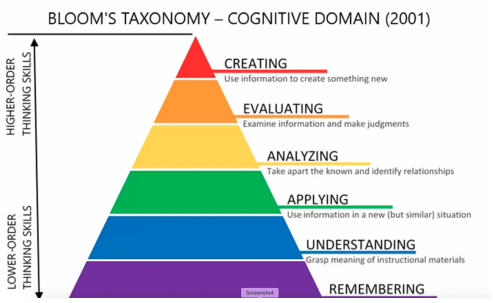
- **What is one challenge with relying solely on human feedback to align large language models?**
    - Sample Answer:

- **In what ways does the Turing Test fall short as a comprehensive measure of artificial intelligence?**
    - *Answer:*

- **Which of the following is a limitation of the Turing Test in evaluating artificial intelligence?**
    - A. It requires physical embodiment of the AI.
      B. It measures only the speed of responses.
      C. It focuses on output indistinguishability rather than actual understanding.
      D. It assesses the AI's ability to perform calculations.
    - **Answer:**

- **Which of the following is a key challenge in applying financial-style risk frameworks to AI systems?**
    - A. AI systems are always deterministic
      B. AI risks are harder to quantify and may change over time
      C. Financial systems are not regulated
      D. AI systems do not require governance structures
      Answer:
- **Rebecca Hwa used the example sentence from an advertisement (in a magazine about computer systems from the 80's) "A computer that understands you like your mother." Why is this sentence ambiguous, and what does this tell us about challenges in AI language understanding?**
    - Answer:

- Alexa Joubin talked about Bloom's Taxonomy of cognitive thought (shown at the right). As LLM's become more integrated in education, how does this taxonomy help think about student assessment?
    - **Answer:**



- Multiple Choice: AI nationalism refers to:
    - A. The use of AI in war
      B. Prioritizing domestic AI development for geopolitical advantage
      C. National-level bans on generative models
      D. Only using AI for public good
    - Answer:

- How might AI nationalism challenge the development of globally trustworthy AI systems
      Sample Answer:

- "In class, we've discussed various 'guardrails'—technical or procedural measures used to make AI systems more trustworthy. Name two specific guardrails and briefly explain how each contributes to trustworthiness."
      Sample Answer: Any 2 of the below

- One current fear of AI is that it will create a crisis of trust/confidence crisis because we no longer trust that words/documents are not written by humans. Please name one or more previously technology changes that generated similar crises.

  Answer: