

# VQA-E Dataset User Study Manual

# Web Tool

- URL: <http://172.21.150.216:5000>
- Read the instructions and click 'Start Evaluation' button
- You will be assigned with a personal url:
  - i.e <http://172.21.150.216:5000/<uid>>
  - Please don't go to other user's url.
  - If you accidentally close the window, you can resume the assessment by input your personal url : <http://172.21.150.216:5000/<uid>>
  - You will assess about 100 questions/answer/explanation sets.
  - You will need around 15-20 min to complete the assessment.
- Click 'Next' button to go to next one
- Click ← in your browser to go to previous one

# Explanation Quality Assessment

- Read image, Q(question), A(answer)
- Give scores to E(explanation) in four quality metrics
  - Fluent:
    - Is this E a fluent English sentence/phrase? A fluent E should be correct in grammar and idiomatic in wording.
  - Correct:
    - Is this E is correct according to the image content?
  - Relevant:
    - Is this E relevant to the question? If an E is relevant, you should be able to infer the answer from E.
  - Complementary:
    - Is this E complementary to the Q&A? If an E is complementary, it should provide complementary information that describes the details related to the question and answer.

# Explanation Quality Assessment

- Scores (1~5)
  - 1-Very Poor
  - 2-Poor
  - 3-Barely Acceptable
  - 4-Good
  - 5-Very Good

# Examples

**Not correct.** The man is not playing tennis.



Evaluation # 8 | Q-Question, A-Answer, E-Explanation. [Help](#)

Q: What is the man doing?

A: playing frisbee

E: A man is playing tennis in the park.

Please evaluate E (explanation) in following aspects.

Fluent:

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Correct:

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Relevant:

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Complementary:

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Next

# Examples

**Not relevant.** E is not relevant to question. We cannot infer the answer from E.

**Not complementary.** E does not provide related details to the Q&A.



Evaluation # 7 | Q-Question, A-Answer, E-Explanation. [Help](#)

Q: Is the TV on or off?

A: on

E: A living room filled with furniture on top of a wooden floor.

Please evaluate E (explanation) in following aspects.

Fluent:

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Correct:

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Relevant:

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Complementary:

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Fluent (1~5): whether the E is a fluent sentence on grammar.

Next

# Examples

**Relevant.** E is relevant to question. E tells that the donut is on a plate.

**Neutrally complementary.** E does not provide many complementary details about answer.



Evaluation # 3 | Q-Question, A-Answer, E-Explanation. [Help](#)

Q: Where is the donuts?

A: plate

E: A table with a donut on a plate

Please evaluate E (explanation) in following aspects.

Fluent:

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Correct:

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Relevant:

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Complementary:

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Next

# Examples

**Relevant.** E is relevant to question. We can infer the answer from E.

**Complementary.** E gives related details about the answer (Not eating dessert but eating pizza).



Evaluation # 17 | Q-Question, A-Answer, E-Explanation, [Help](#)

Q: Is this family eating dessert?

A: no

E: Picture of family eating pizza at table and in living room

Please evaluate E (explanation) in following aspects.

Fluent:

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Correct:

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Relevant:

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Complementary:

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Next



# Examples

**Relevant.** E is relevant to question. We can infer the answer from E.

**Complementary.** E gives details of the answer (because the man wears a boston jersey).



Evaluation # 11 | Q-Question, A-Answer, E-Explanation [Help](#)

Q: What team does the man play for?

A: boston

E: A man wearing a Boston jersey walks with a bat.

Please evaluate E (explanation) in following aspects.

Fluent:

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☒ 5

Correct:

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Relevant:

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Complementary:

☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5

Next