

Annotation Guidelines

(1) - Not helpful

Rationality

- The reply does not include or not related to the question. For example, the case that the question is asking something about education while the response is all about entertainment should be considered as no;
- The reply does not answer the question at all;
- The reply is not clear;

Completeness

- The reply doesn't satisfy every requirements in the question;
- The reply does not contain the explanation or proof of their idea or statement;

The following reply is a sample of (1) rating.

Q: You are fainted. and when you wake up, you find yourself totally naked, falling from the sky without a parachute. how would you describe the moment?

A: I would wonder what's going to hit the ground first; me or my poop.

(2) - Not very helpful

Rationality

- The reply does not include or is only slightly related to the question;
- The reply is not clear;
- The reply is very short and doesn't convey much useful information.

Completeness

- The reply doesn't satisfy most requirements in the question;
- The reply is only answering a minor part of the question;
- The reply is not directly answering the question;
- The reply does not contain the explanation or proof of their idea or statementl.

Q: It's said that we block traumatic events from our memory, what could you have forgotten?

A: I rather not remember it's blocked for a good reason.

(3) - Neither not helpful nor helpful

Rationality

- The reply is much related to the question;
- The reply is very short and doesn't convey much useful information.

Completeness

- The reply does satisfy some requirements in the question;
- The reply answers a small part of the question;
- The reply does not contain the explanation or proof of their idea or statement.

Accuracy

- The response includes terms that are hard to understand without further definition/proof;
- The response includes no additional related information apart from the question;
- The response is just piling up specific terms;
- The response contains ambiguous definition and/or explanations.

The following reply is a sample of (3) rating.

Q: What are some Movies you never get tired of watching?

A: Lotr, gladiator, Iron Man 1.2.3, harry potter(all of them)

(4) - Helpful

Rationality

- The reply is totally related to the question;
- The reply is concise and clear, the reader can get the main idea of it.
- The reply is logical.

Completeness

- The reply does satisfy most or all requirements in the question;
- The reply contains some of explanation or proof of their idea or statement.

Accuracy

- The response gives further definition/proof for terms that are hard to understand;
- The response gives some additional related information apart from the question;
- The response may still contain some ambiguous definition and/or explanations.

The following reply is a sample of (4) rating.

Q: How are gorillas and similar animals taught sign language?

A: They aren't. At least not in the way you probably believe. Gorillas don't compose sentences. They can be trained to repeat patterns and expect results, but they aren't using human language in any meaningful sense. It's similar to training your dog to bark when they want a treat. There's a great podcast with an episode about this: <https://podcasts.apple.com/us/podcast/koko-the-gorilla/id1380008439?i=1000483131905>

(5) - Very helpful

Rationality

- The reply is totally related to the question;
- The reply is either long and detailed, or concise and clear, the reader can get the main idea of it.
- The reply is logical.

Completeness

- The reply satisfies all requirements in the question;
- The reply contains detailed explanation or proof of their idea or statement.

Accuracy

- The response gives further definition/proof for all terms that are hard to understand;
- The response gives some additional related information apart from the question;
- The response contains no ambiguous definition and/or explanations.

The following reply is a sample of (5) rating.

Q: Why Positive charges always are on the upper side of cloud and negative charges below it ?

A: So a thunderstorm has two parts that form lightning, an updraft and a downdraft. These are massive columns of air moving up/down. When that much air is being moved around it is easy for dirt or sand or dust to be picked up and thrown into the storm as well. When bits of the particulates collide they start to build up a charge, this happens enough times and you can have such a large difference in charge between particles that you get lightning. Now for the question, positive charges are not always on the top actually. There are 3 types of lightning, cloud to cloud, cloud to ground, and ground to cloud. All of these decided by where the charges are. While cloud to ground is the one we tend to notice the most it is definitely not the only kind. Same thing with volcanoes as well, so much ash rubbing against each other that charges are formed and lightning goes off, however that tends to be in the smoke column so technically cloud to cloud lightning.