




Ground Truth

Question 1	Ans. 1	Relation	Question 2	Ans. 2
Is it evening?	No	←	Is it sunny?	Yes



Ans. None: Yes 
Ans. SQuINT: Yes 
Ans. CP-VQA: Yes 
Ans. Ours: No

Ans. None: Yes
Ans. SQuINT: Yes
Ans. CP-VQA: Yes
Ans. Ours: Yes

Ground Truth




Question 1	Ans. 1	Relation	Question 2	Ans. 2
Is the dog being friendly to the bird?	Yes	→	Is the dog biting the bird?	No

Ans. None: Yes
Ans. SQuINT: Yes
Ans. CP-VQA: Yes
Ans. Ours: Yes

Ans. None: Yes 
Ans. SQuINT: No
Ans. CP-VQA: Yes 
Ans. Ours: No

Ground Truth

Question 1	Ans. 1	Relation	Question 2	Ans. 2
Is the pizza vegetarian?	Yes	↔	Is there any meat on the pizza?	No


Ans. None: No 
Ans. SQuINT: No 
Ans. CP-VQA: No 
Ans. Ours: Yes

Ans. None: No
Ans. SQuINT: No
Ans. CP-VQA: No
Ans. Ours: No

Ground Truth


Question 1	Ans. 1	Relation	Question 2	Ans. 2
Is this a busy street?	Yes	↔	Is there a lot of traffic on the street?	Yes

Ans. None: Yes
Ans. SQuINT: Yes
Ans. CP-VQA: Yes
Ans. Ours: Yes

Ans. None: Yes
Ans. SQuINT: No 
Ans. CP-VQA: Yes
Ans. Ours: Yes

Ground Truth

Question 1	Ans. 1	Relation	Question 2	Ans. 2
Is this meal vegan?	No	←	Is there meat on the plate?	Yes


Ans. None: No
Ans. SQuINT: No
Ans. CP-VQA: No
Ans. Ours: Yes 

Ans. None: Yes
Ans. SQuINT: Yes
Ans. CP-VQA: Yes
Ans. Ours: Yes

Ground Truth




Question 1	Ans. 1	Relation	Question 2	Ans. 2
Is this a vegan dish?	Yes	→	Is there meat?	No

Ans. None: Yes
Ans. SQuINT: Yes
Ans. CP-VQA: Yes
Ans. Ours: Yes

Ans. None: Yes 
Ans. SQuINT: No
Ans. CP-VQA: No
Ans. Ours: No

Ground Truth

Question 1	Ans. 1	Relation	Question 2	Ans. 2
Is the ground damp?	Yes	↔	Is the ground wet?	Yes

Ans. None: Yes
Ans. SQuINT: No 
Ans. CP-VQA: No 
Ans. Ours: No 

Ans. None: Yes
Ans. SQuINT: Yes
Ans. CP-VQA: Yes
Ans. Ours: Yes