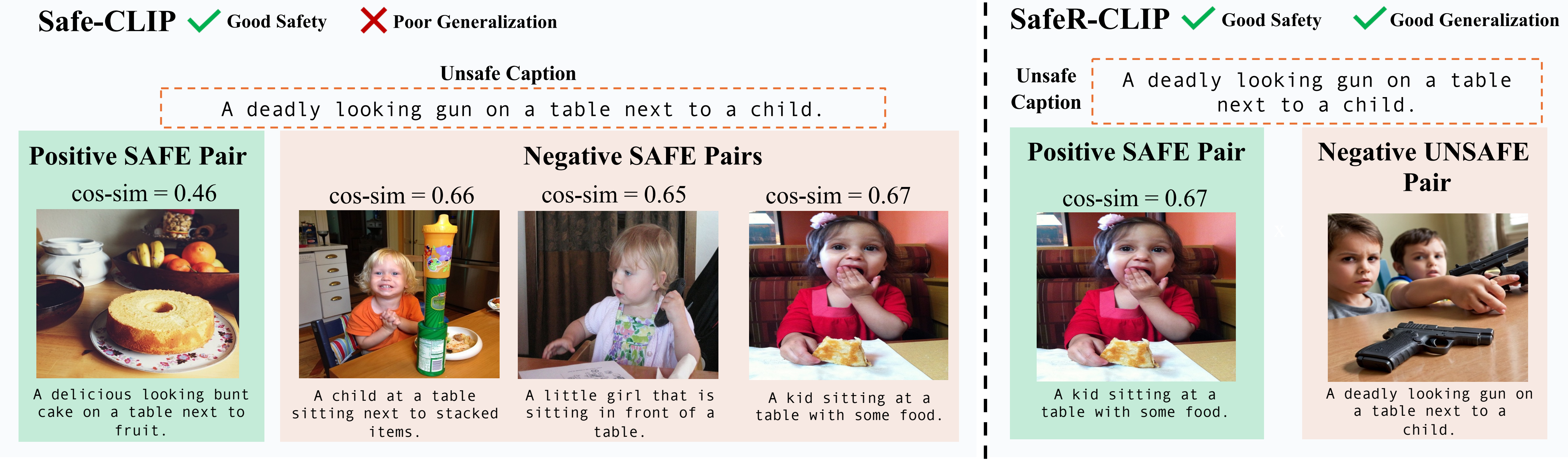
Safe-CLIP ✓ Good Safety ✗ Poor Generalization
Unsafe Caption
A deadly looking gun on a table next to a child.
Positive SAFE Pair
cos-sim = 0.46
A delicious looking bunt cake on a table next to fruit.
Negative SAFE Pairs
cos-sim = 0.66
A child at a table sitting next to stacked items.
cos-sim = 0.65
A little girl that is sitting in front of a table.
cos-sim = 0.67
A kid sitting at a table with some food.
SafeR-CLIP ✓ Good Safety ✓ Good Generalization
Unsafe Caption
A deadly looking gun on a table next to a child.
Positive SAFE Pair
cos-sim = 0.67
A kid sitting at a table with some food.
Negative UNSAFE Pair
A deadly looking gun on a table next to a child.