

Ability	Task Type	Video Source (Secs)	Example
Content Understanding	Object Existence	People & Bog (148)	<i>Are there any stuffed animals visible in the video?</i> (A) No stuffed animals. (B) ... (C) ... (D) ...
	Object Count	Howto & Style (60)	<i>How many pink shoes appear in the video from 0 to 14 seconds?</i> (A) Four pink shoes. (B) ... (C) ... (D) ...
	Action Count	Comedy (71)	<i>How many different actions does it take for the man with ... hit by a broom?</i> (A) four. (B) ... (C) ... (D) ...
	Detailed Description	People & Blogs (758)	<i>What are the detailed steps taken in the video to prepare the dish?</i> (A) Heating oil in a purple pan, sautéing vegetables (B) ... (C) ... (D) ...
	Brief Description	Pets & Animals (35)	<i>What happens at the video?</i> (A) A person extends their hand towards two Siberian Huskies. (B) ... (D) ...
	Event Description	Education (105)	<i>How does the speaker describe the transformation brought by the women ... ?</i> (A) They saw themselves as victims and didn't take action. (B) ... (D) ...
	Event Sequence	Film & Animation (172)	<i>Can you list the events in the order they occur in the video?</i> (A) A woman in a white dress searches for 'Daniel' [.....] (B) ... (D) ...
	OCR	News & Politics (10)	<i>What is written on the screen at the start of the video?</i> (A) ALBINO COBRA (B) ... (C) ... (D) ...
Temporal Understanding	Action Recognition	People & Blogs (35)	<i>What action does the person with ... perform in the early part of the video?</i> (A) painting daisies on a canvas (B) ... (C) ... (D) ...
	Action Sequence	Gaming (90)	<i>Describe the sequence of actions performed from the beginning to the end ...</i> (A) Green-haired man watches, groups march, blue [.....] (B) ... (D) ...
	Action Location	Science & Technology (28)	<i>At what time ... person's hands start using pliers to work with a metal wire?</i> (A) from 6 to 8 seconds (B) ... (C) ... (D) ...
	Event Location	Pets & Animals (110)	<i>When does the individual in the blue shirt hold a ... puppy inside the vehicle?</i> (A) Around 75-85 seconds (B) ... (C) ... (D) ...
	Object Location	Travel & Events (10)	<i>Which second does the woman start placing squash on the plate?</i> (A) 10th second (B) ... (C) ... (D) ...
	Temporal Location	Film & Animation	<i>How does the brick building visually change throughout the video?</i> (A) It changes color and texture. (B) ... (C) ... (D) ...
	Object Temporal Relation	(98)	
Spatial Understanding	Object Spatial Location	Education (61)	<i>What is the bounding box coordinates of the tree at 14 second of the video?</i> (A) [0.3, 0.0, 0.71, 1.0] (B) ... (C) ... (D) ...
	Objects	Film & Animation (284)	<i>How does the pink pig character with a red dress typically hold the football?</i> (A) On her head, balancing it. (B) ... (C) ... (D) ...
	Spatial Relation	Film & Animation (127)	<i>What direction is the car moving in the video?</i> (A) towards the right (B) ... (C) ... (D) ...
	Object Spatial Tracking		
Understanding	Human Activity Analysis	Travel & Events (100)	<i>Why are various people gathered in the video?</i> (A) To visit a historical monument. (B) ... (C) ... (D) ...
Understanding	Anomaly Detection	Travel & Events (336)	<i>Which of the following violations might be present in the video?</i> (A) Sexual. (B) ... (C) ... (D) ... (E) ...
Relation Reasoning	Image-Video	News & Politics (12)	<i>What role does the person in the image portray in the video?</i> (A) A teacher (B) ... (C) ... (D) ...
	Video-Video	Sports (7 and 10)	<i>What's the relationship between video 1 and video 2?</i> (A) video 1 is before video 2 (B) ... (C) ... (D) ...
Temporal Reasoning	Event Prediction	Howto & Style (73)	<i>What will the person do next after concluding their exercise on the yoga mat?</i> (A) The person will start doing jumping jacks beside the mat. (B) ... (D) ...
	Action Prediction	Film & Animation (107)	<i>What action is expected to occur immediately after the video concludes?</i> (A) The drawing man in blue and the skunk will walk together. (B) ... (D) ...
Causal Reasoning	Why	Comedy (593)	<i>Why do the people in the last scene seem excited?</i> (A) They received unexpected gifts. (B) ... (C) ... (D) ...
Counterfactual Reasoning	Counterfactual	Sports (63)	<i>Imagine if the SUV ... had proper winter tires instead of ... , what is the result?</i> (A) It has a smoother ride on the snowy surface. (B) ... (C) ... (D) ...
Commonsense Reasoning	Commonsense	Entertainment (64)	<i>What type of fruit is being shown at the market in the video?</i> (A) Mangosteen (B) ... (C) ... (D) ...
Logical Reasoning	Logic	Science & Technology (70)	<i>What will happen right after this video comes to an end?</i> (A) A woman will be presenting a new armor design on a stage. (B) ... (D) ...