

User Studies for Sketch2Photo

1 User Study I

This supplemental file provides the high resolution compositions and detailed statistics of the user study I. In this user study, ten subjects were selected. Nine of them are novices to photomontage (novice to both our system and Adobe Photoshop), they are all undergraduate students majoring in computer science. The other is an artist professional in Adobe Photoshop. We split the nine novices into three groups of equal size. Group A is provided with the Adobe Photoshop; group B is provided with our hybrid blending tools, which has a drag & drop interface for interactive composition; group C is provided with our complete system. The artist is also provided with the Adobe Photoshop. Each group of subjects is given 20 minutes instruction of the provided tools, then two composition tasks are given to them.

We perform two tasks, where the subjects are required to generate a composition according to a verbal description. We measure the quality of each composition by a subjective score ranging from 1 to 5. Scores: 5 = very good, photo-realistic; 4 = good, but with slightly noticeable artifacts; 3 = obvious artifacts, but still acceptable; 2 = very obvious artifacts, not quite acceptable; 1 = very bad, unacceptable. Each composition is presented to five evaluators (not selected as subjects) and their average score is used. **The evaluators are unaware of the three groups, and hence also do not know which composition comes from which group.** We list composition results and statistics below. All the high resolution compositions are provided with the comments and scores from evaluators.

1.1 Task1: Throwing Frisbee

In this task, the subjects have to generate an image according to a verbal description of ‘a person throws a frisbee under a tree, and his/her dog jumps for it; the person is on the left and the dog is on the right’. The subjects are required to generate a result as good as they can achieve.

1.2 Task2: Fishing Bear

In this task, the subjects have to generate an image according to a verbal description of ‘a bear is catching a jumping salmon in a river; the bear is on the left and the fish is on the right’. The subjects are given 30 minutes to generate the composition.

	Subjects:	A1	A2	A3	Pro	B1	B2	B3	C1	C2	C3
	Average score:	1.8	2.8	3.4	4.8	4.2	4.0	4.8	4.2	4.6	4.2
Task I	Total time (min):	53	47	83	18	63	55	74	47	42	45
	Interaction time (min):	40+13	39+8	65+18	12+6	57+6	50+5	70+4	5+3	3+1	5+0.6
	Average score:	1.4	2.6	2.2	4.0	3.2	3.2	2.6	4.8	4.6	4.6
Task II	Total time (min):	30	26	30	22	30	28	29	28	26	30
	Interaction time (min):	20+10	18+8	15+15	15+7	27+3	24+4	25+4	2+2	2+0.5	3+3

Table 1: Statistics from the user study I. Each column is the data for one subject. ‘Pro’ indicates the professional artist. A1-3, B1-3 and C1-3 are the three subjects in each group. For each task, we show its results in three rows. The first row is the average composition quality from five evaluator. The second row is the overall time spent to generate the composition. We also present the user interaction time in the third row. This interaction time is split into the time on manual image search and the time on interactive composition for group A and B. It is split into the time on contour sketch and interactive refinement for group C.



Figure 1: Composition A1. Comments and scores: obviously fake, 1; obvious artifacts at the boundaries of scene items, 2; scene items are like floating scraps of paper, 2; composition quality is too low, 1; objects are too small, 3.



Figure 2: Composition A2. Comments and scores: artifacts are too obvious, 2; artifacts on woman and dog's leg, 3; woman's right foot is missing, 3; scene items are not clean, 3; artifacts on the legs and feet, 3.

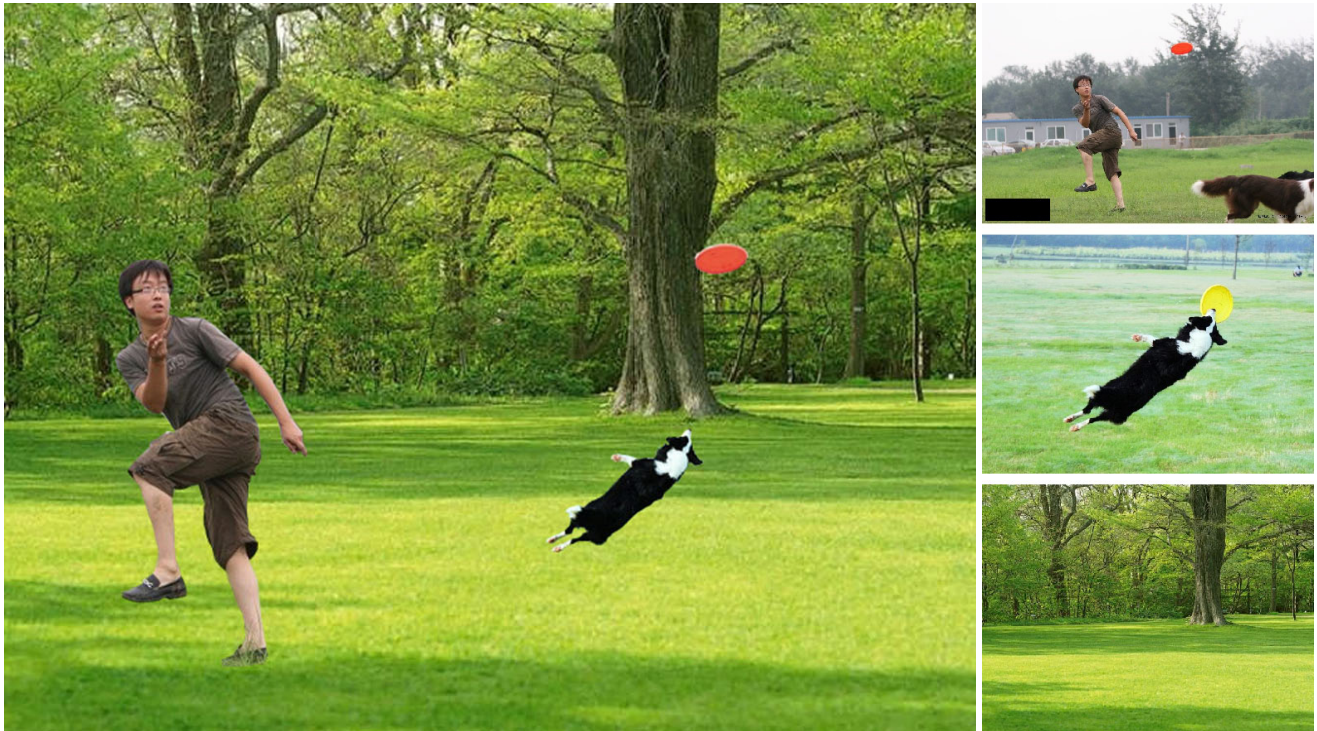


Figure 3: Composition A3. Comments and scores: illumination of the scene items does not match the background, 3; the boundaries of the the scene items are a little too hard, 4; scene items look like floating, 3; appearances of the scene items are inconsistent, 3; the dog looks a little fake, maybe because there is no shadow, 4.

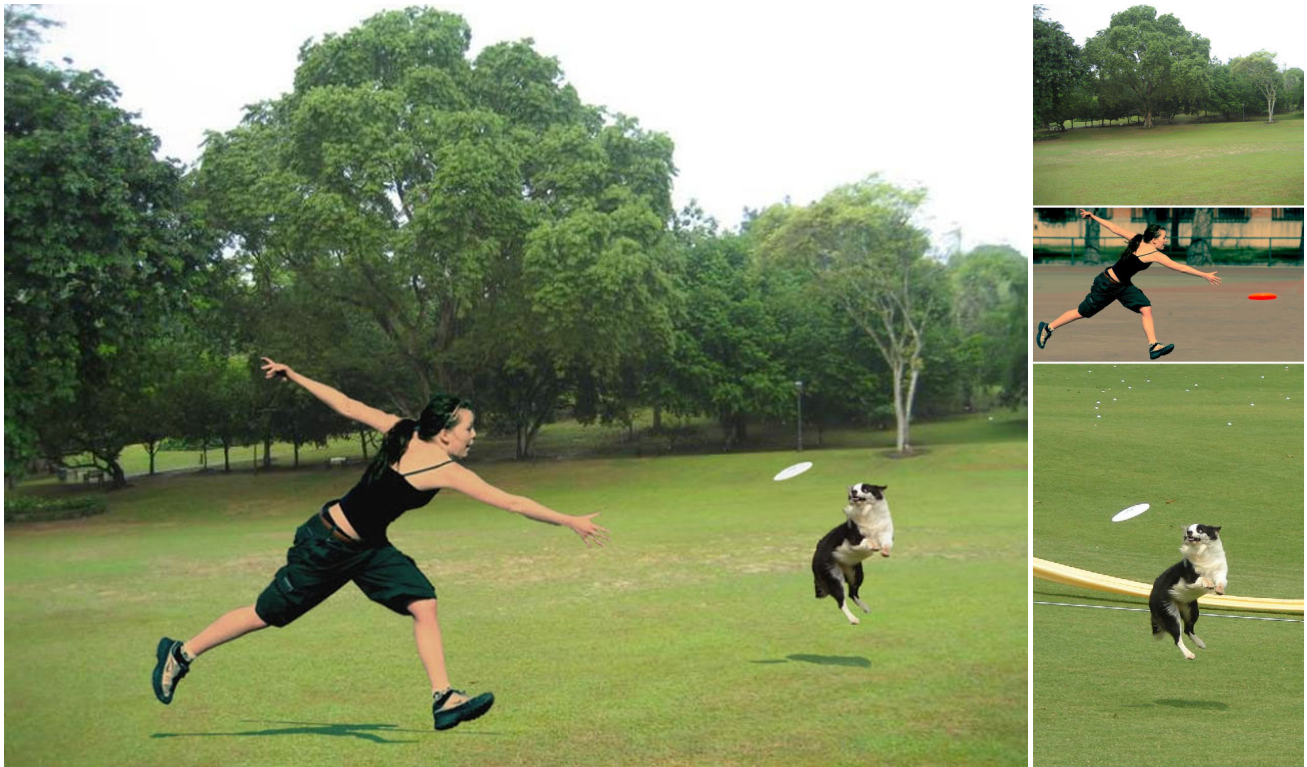


Figure 4: Composition Pro. Comments and scores: artistic rather than realistic, shadows look fake, 4; 5; 5; 5; 5.



Figure 5: Composition B1. Comments and scores: dog has no shadow, 4; hand is missing, 4; dog has no shadow, 4; 5; boy's left hand is missing, 4.

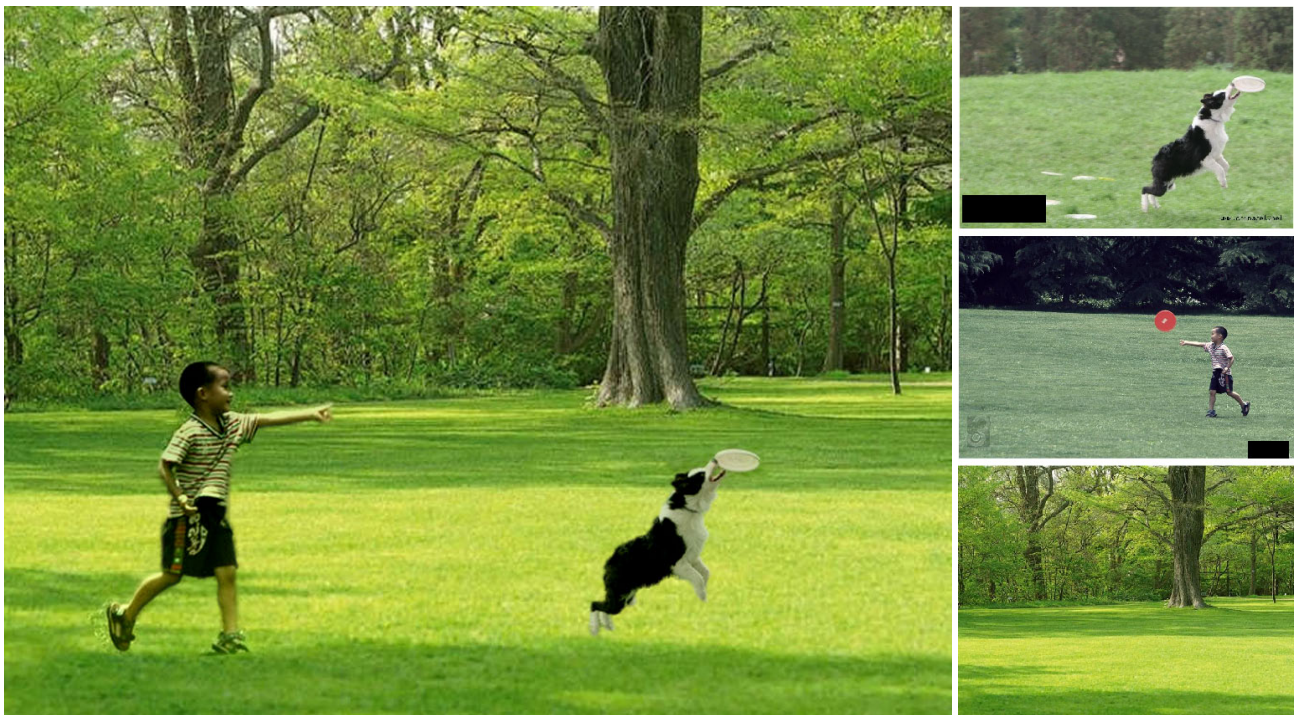


Figure 6: Composition B2. Comments and scores: artifacts around boy's head, colors of scene items are unreal, 3; boy's foot has artifacts, 4; boy's boundary is somehow blurred, 4; boy looks a little strange, 4; the interaction among scene items is the best of all, 5.



Figure 7: Composition B3. Comments and scores: the girl is not looking at the dog, 4; 5; 5; 5; 5.



Figure 8: Composition C1. Comments and scores: appearance of the man feels unreal, 3; the frisbee is a little strange, 4; the man looks like float in air, 4; 5; 5.

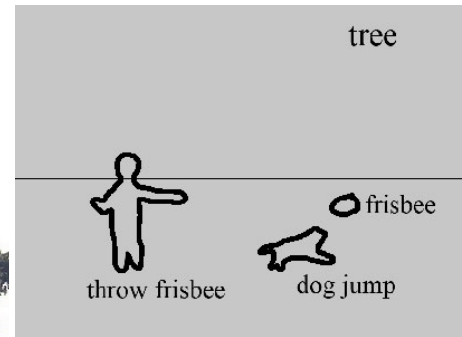


Figure 9: Composition C2. Comments and scores: The girl and the dog has inconsistent shadow, 4; 5; 5; disk is not so real without motion blur, 4; 5.

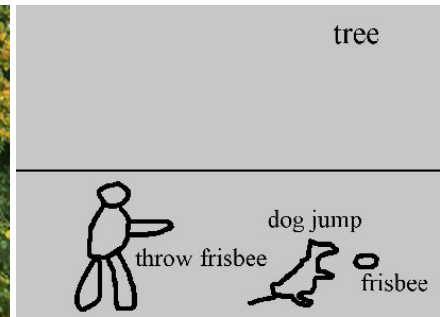


Figure 10: Composition C3. Comments and scores: the dog is overexposed, no shadows, 4; 5; there should be shadows on the ground, 4; the leg of girl is a little blurred; the feet of the girl are missing, 4.



Figure 11: Composition A1. Comments and scores: floating bear, dead fish, 1; the fish is squashed, 2; the bear is like a piece of paper, the fish is very strange, 1; very fake, 1; the bear's back is broken, the fish has a strange shape, a polar bear shouldn't appear here, 2.



Figure 12: Composition A2. Comments and scores: the bear looks like being glowing, 2; the bear is unreal, although the fish is good, 3; the boundary of bear is too blurred, 3; the foot of the bear is missing, 2; bear looks strange, 3.



Figure 13: Composition A3. Comments and scores: the composition is too rough, 1; the scene item's boundaries are too hard, 3; the spray under the fish is very fake, 2; composition artifacts are too obvious, 2; the spray is fake, the brown bear shouldn't appear in ice, it should be a polar bear, 3.



Figure 14: Composition Pro. Comments and scores: the edge of bear seems fake, 3; artifacts on bear's boundary, 4; the spray under fish has some artifacts, 4; 5; the fish is too ugly and shouldn't belong to here, it looks unnatural, 4.



Figure 15: Composition B1. Comments and scores: the bear is like walking on the water, 2; bear is floating on the water, fish has no spray, 3; there should be some spray under fish, 4; bear should make some spray, moreover, it's unnatural to be a polar bear, 3; polar bears don't belong to here, but the composition is ok, 4.



Figure 16: Composition B2. Comments and scores: the scene is obviously unreal, and the size of the fish is seriously wrong compared to the bear, 2; an interesting image but the fish is improperly large, 4; the fish looks unnatural, 4; the fish has obvious artifacts, and it's too large, the bear shouldn't be a polar bear, 2; polar bear is improper, 4.



Figure 17: Composition B3. Comments and scores: the bear is like walking on the water, the composition is very obvious, 2; the bear is floating, the fish has no spray, 3; the fish is obviously fake, 3; the bear and fish has no affect to the water, it's very fake, 2; polar bears don't belong to here, the fish is fake, 3.



Figure 18: Composition C1. Comments and scores: the composition is good, the bear's perspective may be a little incorrect, 4; 5; 5; 5; 5.



Figure 19: Composition C2. Comments and scores: the fish should make some spray, 4; 5; the fish is not real, 4; 5; 5.



Figure 20: Composition C3. Comments and scores: the fish is a little unnatural, 4; the fish is not real, 4; 5; 5; 5.

2 User Study II

This supplemental file provides the high resolution compositions and detailed statistics of the user study II. In this user study, we tested if our system can successfully generate novel scenes. Four subjects were selected, all novices to our system. After 20 minutes of instruction, one of the subjects was asked to provide 15 image composition tasks. The other three subjects used our system to generate these compositions. Note subjects may change the text labels after viewing acquired images, for example, add or remove verbs or prepositions. Each composition was evaluated in the same way as user study I. Among all the 45 compositions (15 tasks \times 3 subjects), 35 (77.8%) were considered to be successful (average score ≥ 3). Note some tasks in this user study have overlapped scene items, Their layer is decided by the drawing order of the sketch, scene items drawn later overlay the former ones. User can also adjust the position of scene items for refinement. We list composition results and statistics below.

Tasks:	airplane	cat	circus	dinosaur	elephant	gorilla	Kung Fu	nebula	robots	room	shark	shepherd	slide	soccer	witch
Subject 1:	4.4	3.2	4	2.2	2	3.6	3.6	4	4.2	3.8	4.4	4	2.2	3.6	2.8
Subject 2:	5	4	2.6	3.8	4.2	2.2	4.4	4	3.2	3	3.6	4.4	2.2	4.4	3.4
Subject 3:	4.6	2.6	3	1.8	4.2	4.2	3.2	3.8	4	4.2	3.8	4	4	2.2	3.4

Table 2: Statistics from the user study II. Each row is the data for one subject.



Figure 21: Composition Task: Airplane, Subject 1. Scores: 5; 5; 5; 4; 3.



Figure 22: Composition Task: Airplane, Subject 2. Scores: 5; 5; 5; 5; 5.



Figure 23: Composition Task: Airplane, Subject 3. Scores: 5; 5; 5; 4; 4.



Figure 24: Composition Task: Cat, Subject 1. Scores: 3; 3; 2; 5; 3.



Figure 25: Composition Task: Cat, Subject 2. Scores: 4; 5; 4; 3; 4.



Figure 26: Composition Task: Cat, Subject 3. Scores: 2; 2; 2; 4; 3.

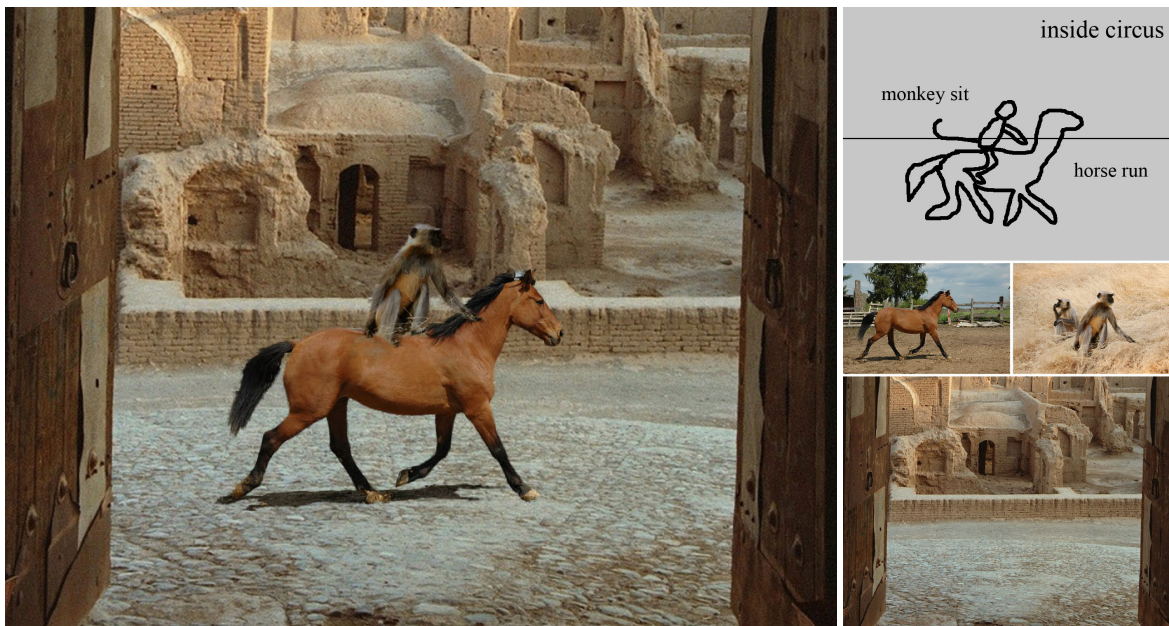


Figure 27: Composition Task: Circus, Subject 1. Scores: 3; 4; 5; 4; 4.

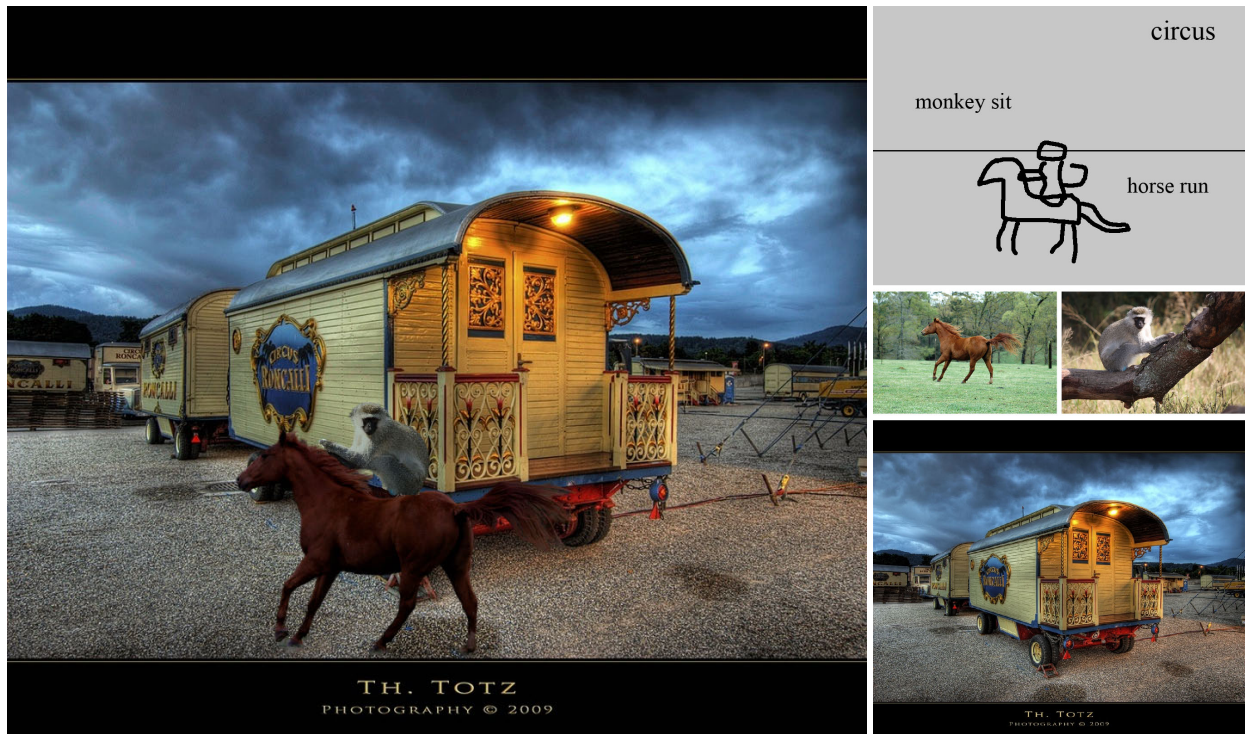


Figure 28: Composition Task: Circus, Subject 2. Scores: 3; 1; 2; 3; 4.



Figure 29: Composition Task: Circus, Subject 3. Scores: 3; 2; 3; 3; 4.



Figure 30: Composition Task: Dinosaur, Subject 1. Scores: 2; 1; 2; 3; 3.



Figure 31: Composition Task: Dinosaur, Subject 2. Scores: 3; 3; 4; 4; 5.



Figure 32: Composition Task: Dinosaur, Subject 3. Scores: 1; 1; 2; 2; 3.



Figure 33: Composition Task: Elephant, Subject 1. Scores: 2; 2; 1; 2; 3.



Figure 34: Composition Task: Elephant, Subject 2. Scores: 4; 4; 4; 4; 5.

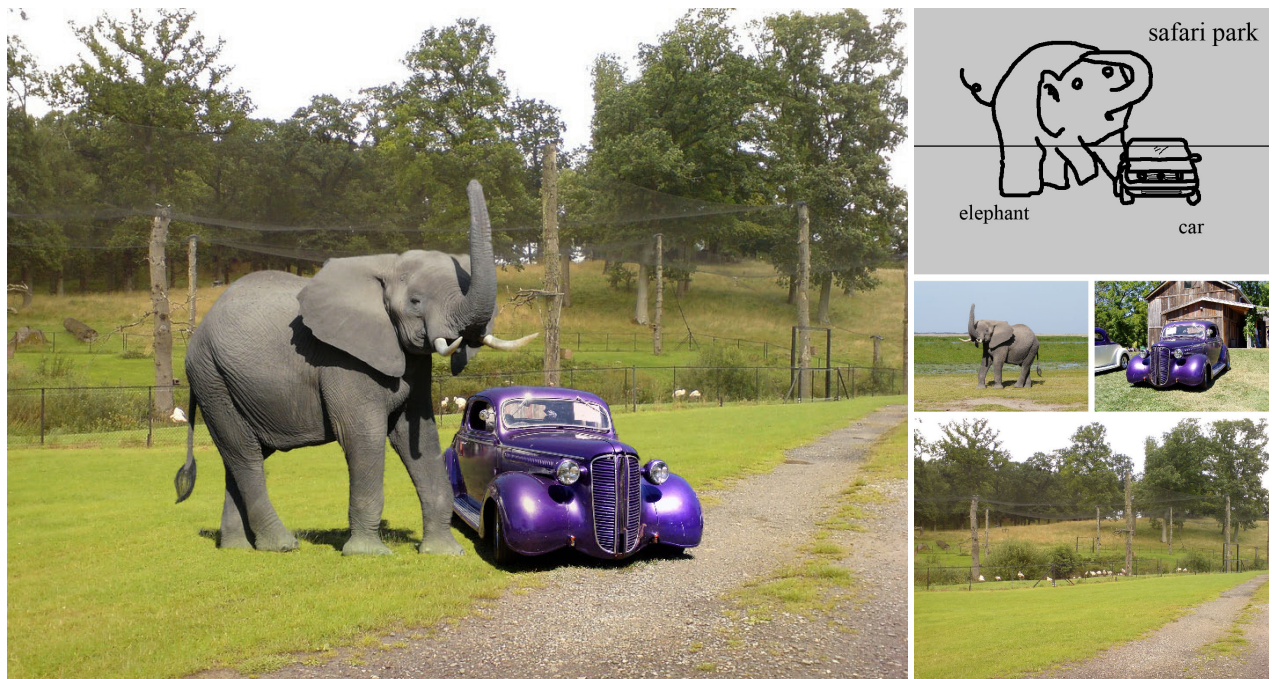


Figure 35: Composition Task: Elephant, Subject 3. Scores: 5; 5; 3; 3; 5.

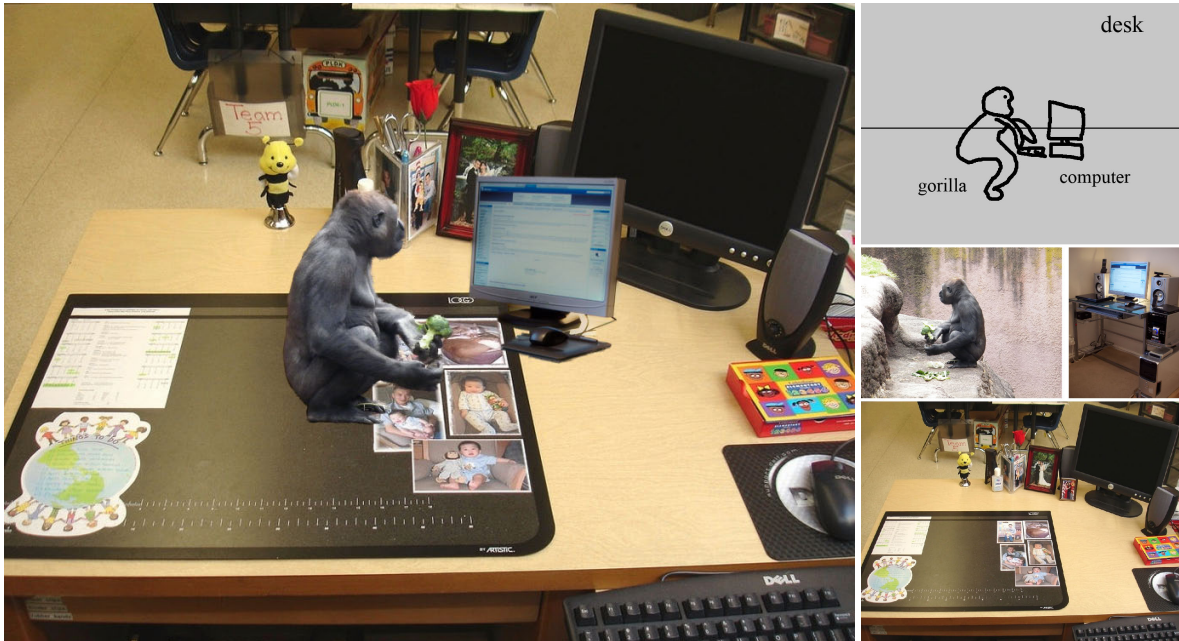


Figure 36: Composition Task: Gorilla, Subject 1. Scores: 3; 4; 4; 2; 5.



Figure 37: Composition Task: Gorilla, Subject 2. Scores: 2; 3; 2; 2; 2.

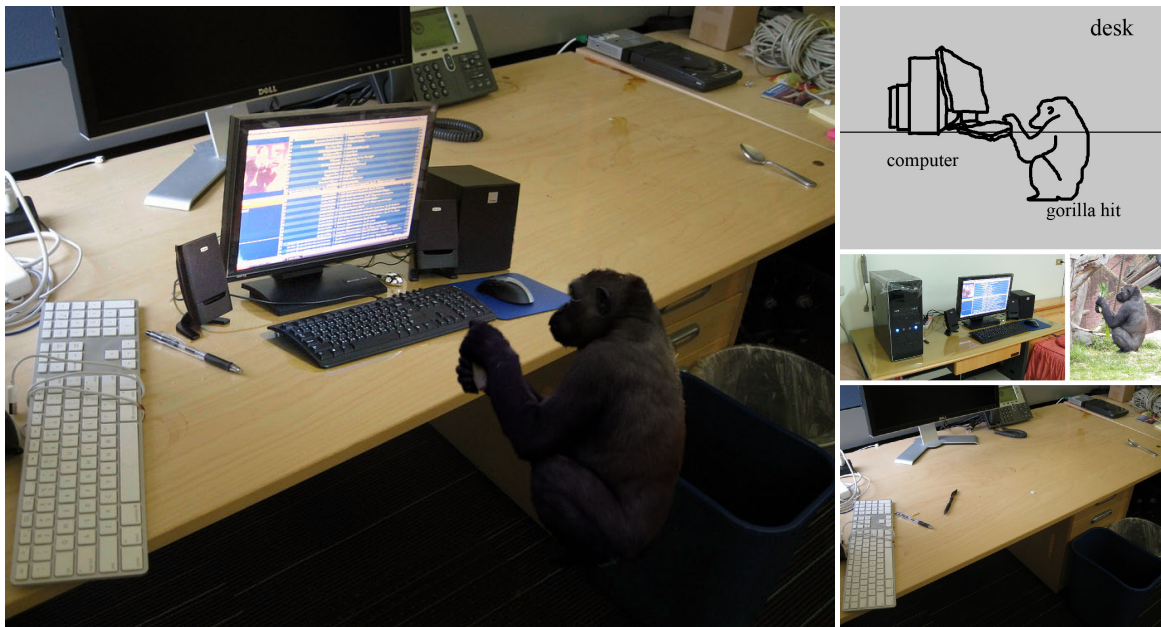


Figure 38: Composition Task: Gorilla, Subject 3. Scores: 4; 4; 4; 4; 5.



Figure 39: Composition Task: Kung Fu, Subject 1. Scores: 4; 3; 3; 3; 5.



Figure 40: Composition Task: Kung Fu, Subject 2. Scores: 5; 4; 4; 4; 5.



Figure 41: Composition Task: Kung Fu, Subject 3. Scores: 4; 3; 3; 3; 3.

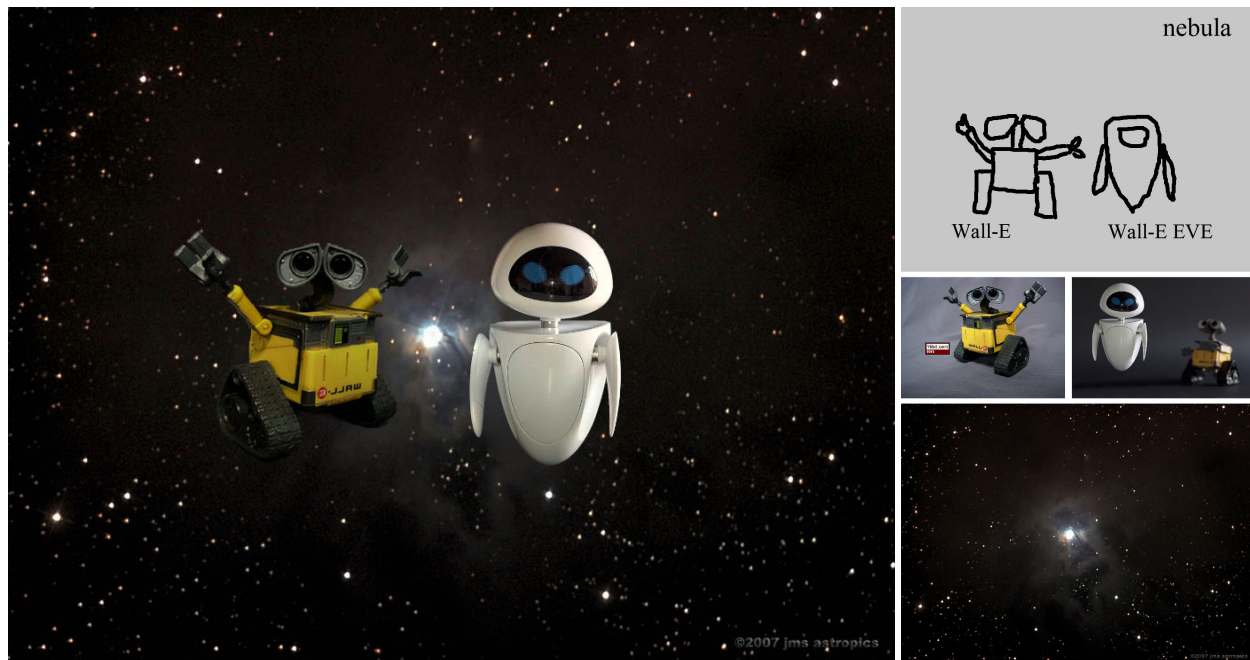


Figure 42: Composition Task: Nebula, Subject 1. Scores: 4; 4; 4; 3; 5.

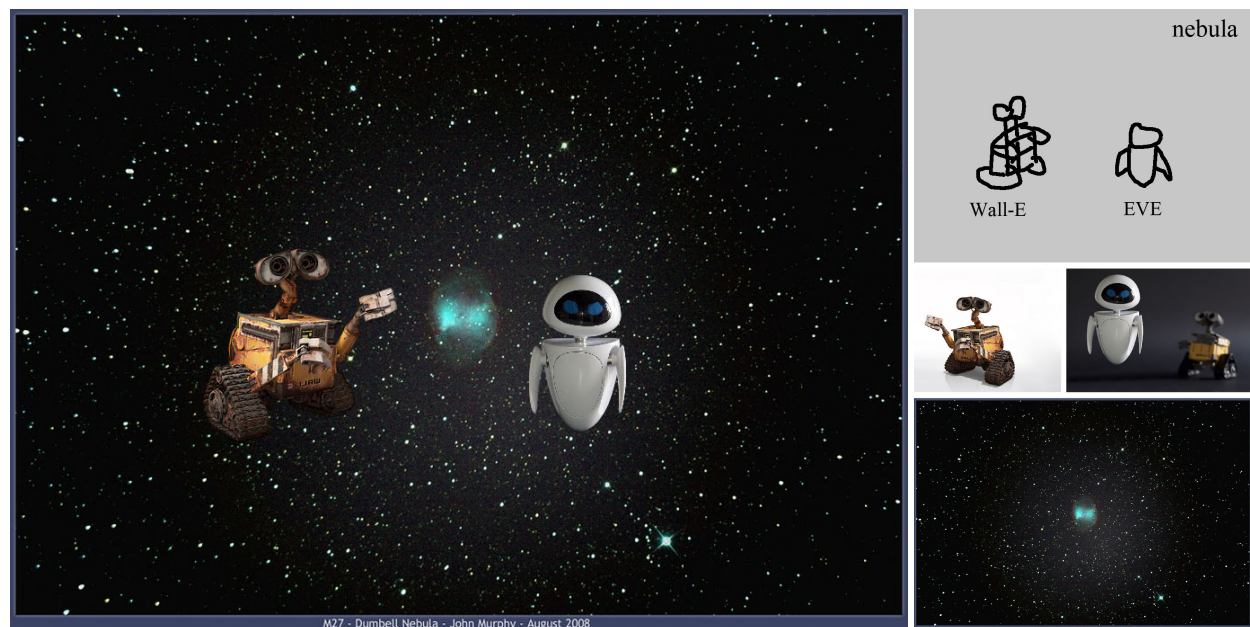


Figure 43: Composition Task: Nebula, Subject 2. Scores: 4; 4; 4; 3; 5.

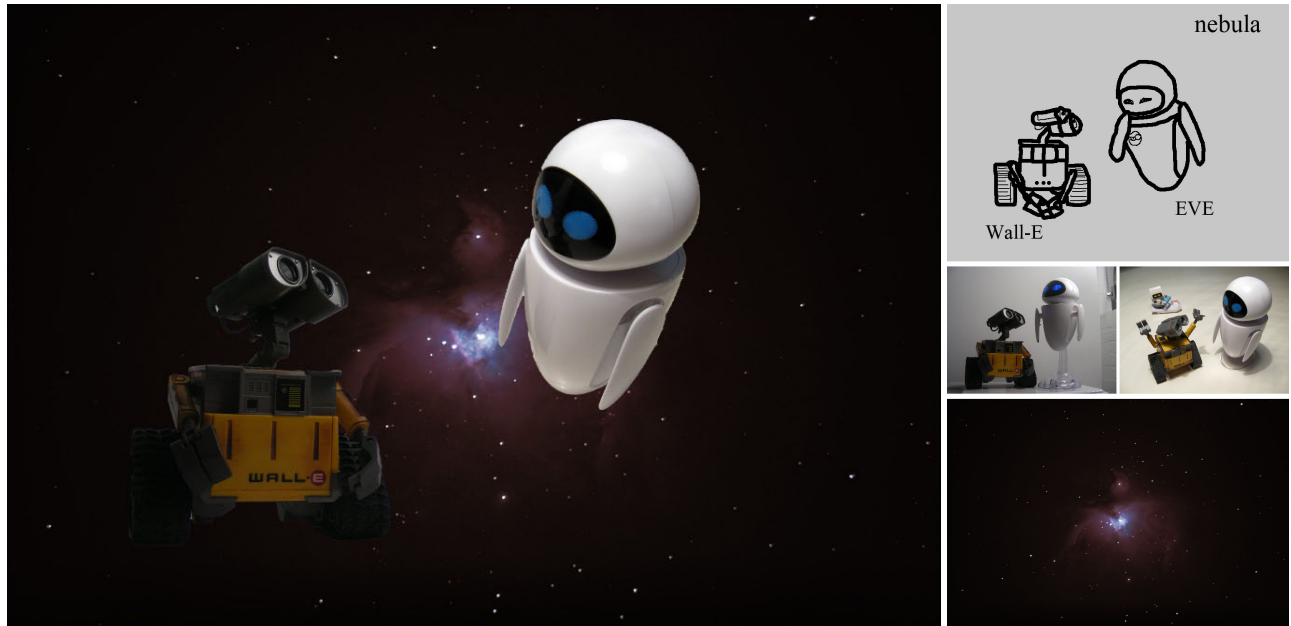


Figure 44: Composition Task: Nebula, Subject 3. Scores: 4; 4; 4; 2; 5.



Figure 45: Composition Task: Robots, Subject 1. Scores: 4; 5; 4; 4; 4.

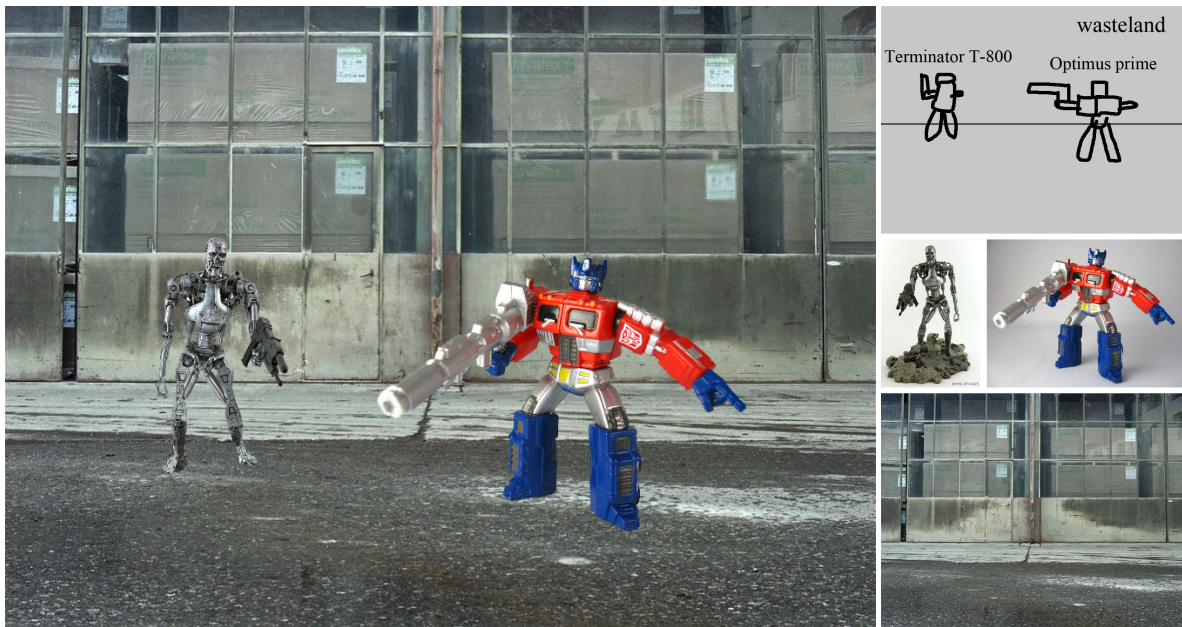


Figure 46: Composition Task: Robots, Subject 2. Scores: 3; 2; 3; 3; 5.

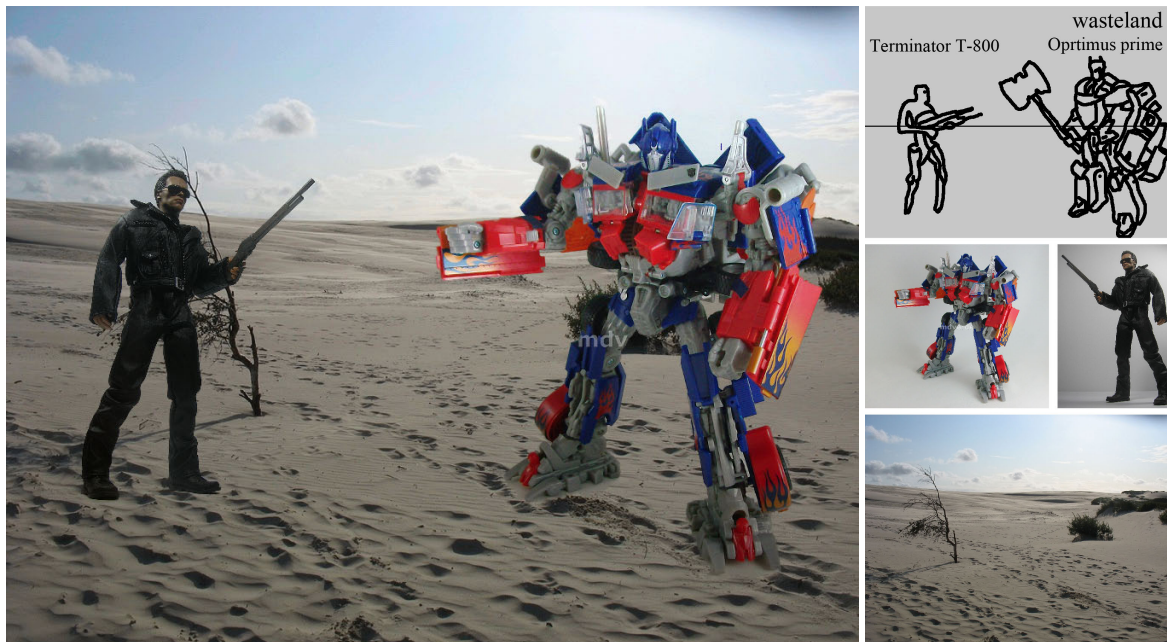


Figure 47: Composition Task: Robots, Subject 3. Scores: 4; 4; 4; 3; 5.

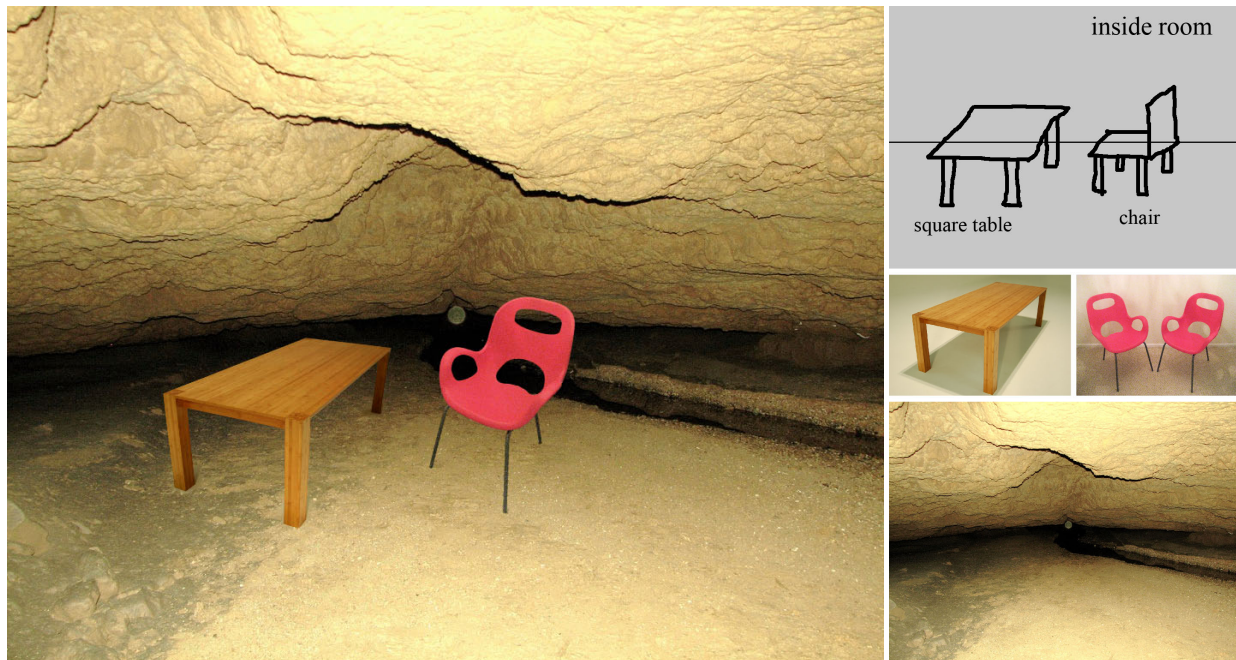


Figure 48: Composition Task: Room, Subject 1. Scores: 4; 5; 4; 2; 4.



Figure 49: Composition Task: Room, Subject 2. Scores: 2; 4; 3; 3; 3.



Figure 50: Composition Task: Room, Subject 3. Scores: 3; 5; 5; 3; 5.



Figure 51: Composition Task: Shark, Subject 1. Scores: 3; 5; 5; 4; 5.



Figure 52: Composition Task: Shark, Subject 2. Scores: 3; 4; 4; 4; 3.



Figure 53: Composition Task: Shark, Subject 3. Scores: 3; 4; 4; 3; 5.



Figure 54: Composition Task: Shepherd, Subject 1. Scores: 4; 4; 3; 4; 5.



Figure 55: Composition Task: Shepherd, Subject 2. Scores: 4; 5; 4; 4; 5.



Figure 56: Composition Task: Shepherd, Subject 3. Scores: 4; 4; 3; 4; 5.



Figure 57: Composition Task: Slide, Subject 1. Scores: 1; 3; 2; 2; 3.



Figure 58: Composition Task: Slide, Subject 2. Scores: 1; 3; 3; 2; 2.



Figure 59: Composition Task: Slide, Subject 3. Scores: 3; 5; 4; 4; 4.



Figure 60: Composition Task: Soccer, Subject 1. Scores: 5; 3; 4; 2; 4.



Figure 61: Composition Task: Soccer, Subject 2. Scores: 4; 4; 5; 4; 5.



Figure 62: Composition Task: Soccer, Subject 3. Scores: 2; 2; 2; 2; 3.



Figure 63: Composition Task: Witch, Subject 1. Scores: 2; 1; 5; 3; 3.



Figure 64: Composition Task: Witch, Subject 2. Scores: 4; 2; 4; 4; 3.



Figure 65: Composition Task: Witch, Subject 3. Scores: 3; 4; 4; 3; 3.