## Appendix

### WSC Instance: The man couldn’t lift his son because he was so weak. Answer: the man (Agent)

**Evidence and labels:**
- “However I was so weak that I couldn’t lift” → EA
- “She was so weak she couldn’t lift” → EA
- “It hurts to lift my leg and it’s kind of weak” → EP

**Stats and resolution:**
- Agent evidence strength: 97
- Patient evidence strength: 72
- Number of scraped sentences: 109
- Resolution: Agent

### WSC Instance: The man couldn’t lift his son because he was so heavy. Answer: the man (Patient)

**Evidence and labels:**
- “I couldn’t lift it. It was too heavy” → EP
- “When they arrived, 23-year-old Ethel was wearing men’s clothing. Her bedroll was so heavy she couldn’t lift”. → EP
- “With the gutter machine on a scissors lift we ran the gutter out on the roof. It was so heavy”. → EA

**Stats and resolution:**
- Agent evidence strength: 12
- Patient evidence strength: 31
- Number of scraped sentences: 38
- Resolution: Patient

### COPA Instance: I missed the bus. What happened as a RESULT?

**Option 1:** I was early to work.
**Option 2:** I was late to work.

**Answer:** Option 2

**Evidence and labels:**
- “Brenda forgot to set her alarm clock, overslept, missed her train, and was late to work” → EO2
- “Listen Donna, I don’t care if you missed the bus this morning. You have been late” → EO2
- “Well, I missed the bus again ... Last time I told my mom that the bus came early” → EO1

**Stats and resolution:**
- Agent evidence strength: 12
- Option 1 evidence strength: 4
- Option 2 evidence strength: 39
- Number of scraped sentences: 71
- Resolution: Option2

### COPA Instance: The woman felt envious of her sister. What was the CAUSE of this?

**Option 1:** Her sister was happy.
**Option 2:** Her sister got divorced.

**Answer:** Option 1

**Evidence and labels:**
- “I found that throughout her story, I was happy for her success and I never felt envious” → EO1
- “Although I was happy for my friends, I also felt envious” → EO1
- “A divorce moves at the speed of complete deadening silence, or that’s how it feels,” she said. “When we got divorced, I felt” → EO2

**Stats and resolution:**
- Agent evidence strength: 12
- Option 1 evidence strength: 36
- Option 2 evidence strength: 18
- Number of scraped sentences: 23
- Resolution: Option 1

---

Table S1: Example Resolution for a WSC pair and Two COPA Test Instances.