1



color necessary gender match

2



color redundant gender match

3



color redundant gender mismatch

4

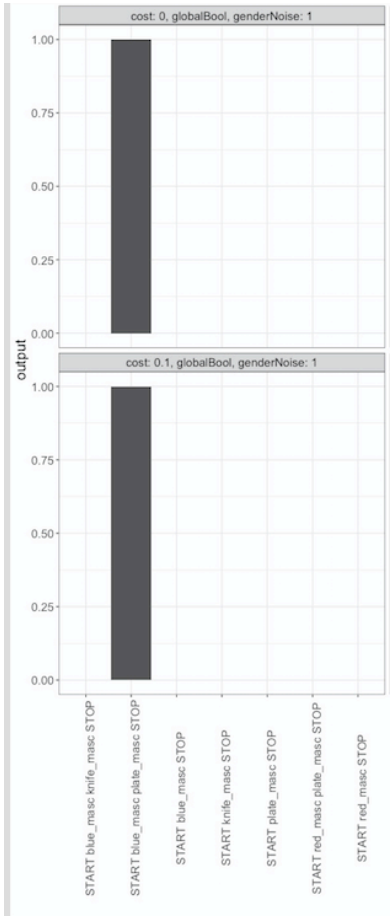


color necessary gender mismatch

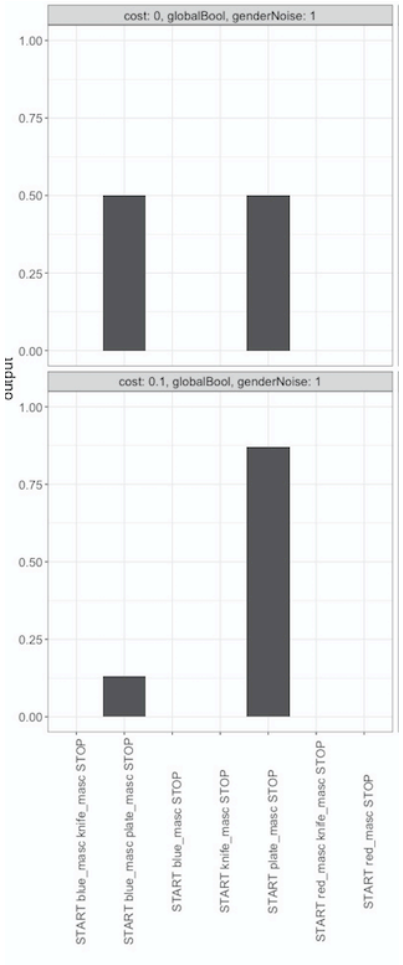
- Color redundancy: compare over all mention of color across 2 and 3
- Noun omission: compare 3 versus 4
- Baseline: look at 1

Global Boolean

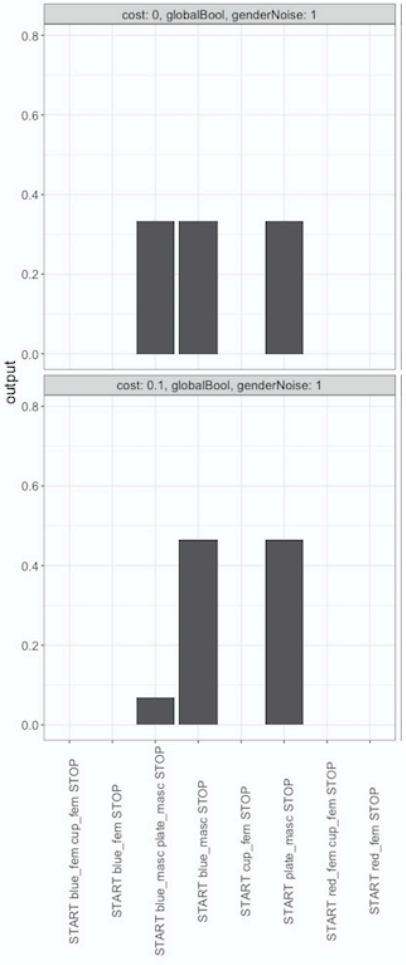
Scene 1
Color necessary
Gender match



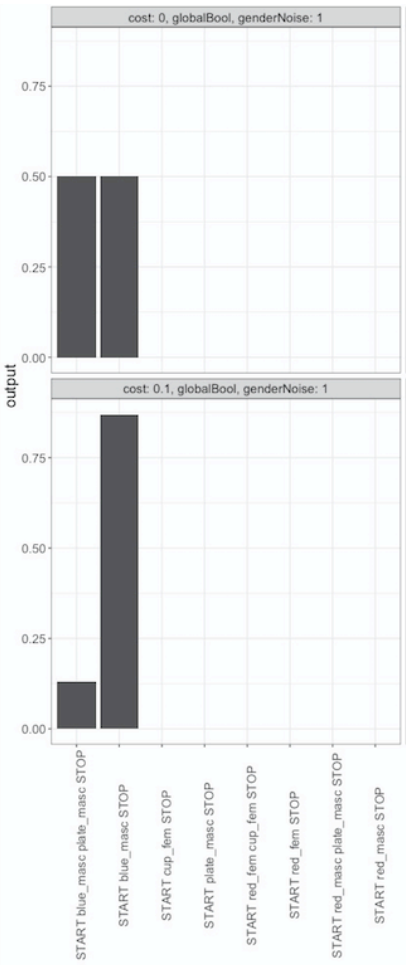
Scene 2
Color redundant
Gender match



Scene 3
Color redundant
Gender mismatch



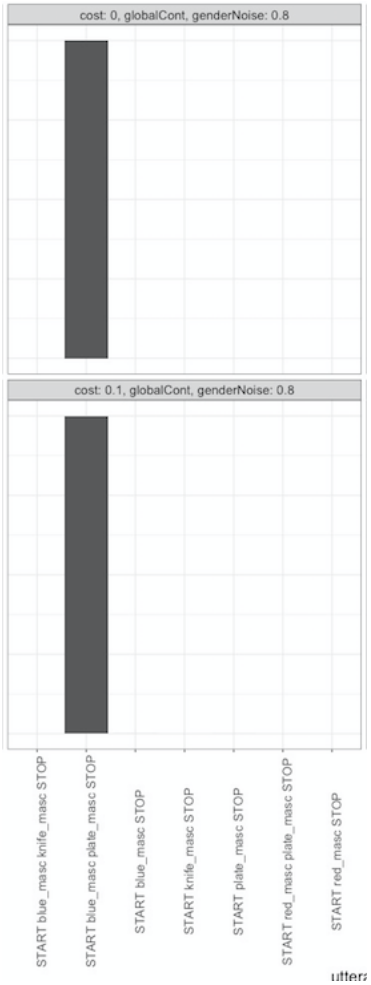
Scene 4
Color necessary
Gender mismatch



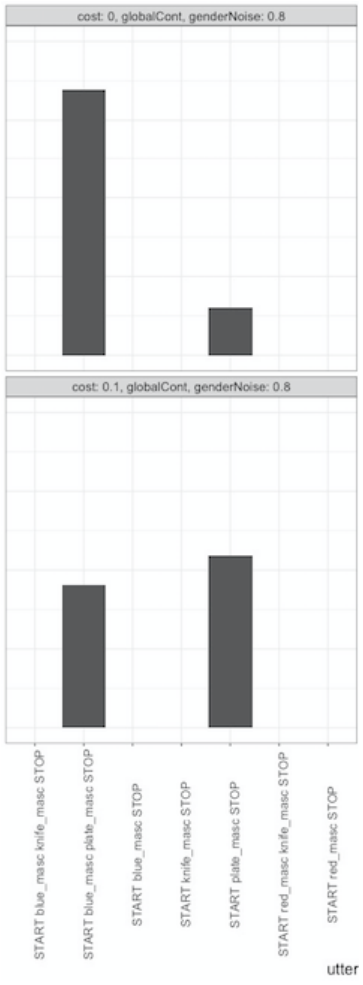
- Color overmodification (2vs 3 and 2 vs 4): more use of color adjective (combination of blue and blue plate) for gender mismatch than gender match conditions
- Noun omission (3 vs 4): higher rates of noun omission in color necessary than color redundant condition

Global Continuous

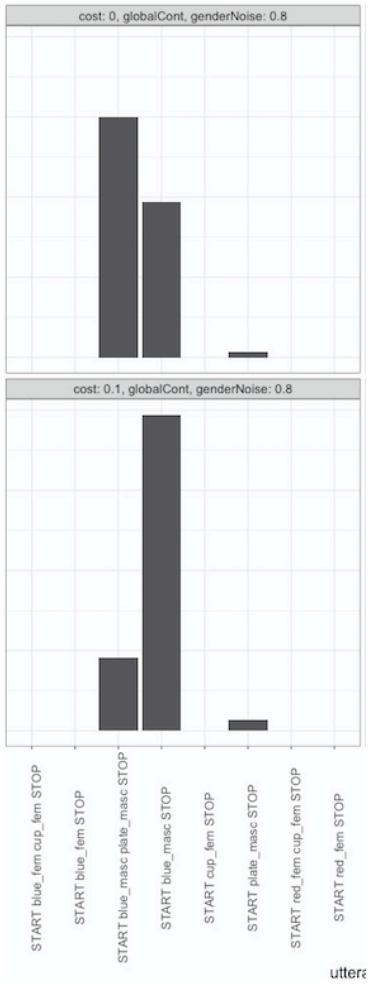
Scene 1
Color necessary
Gender match



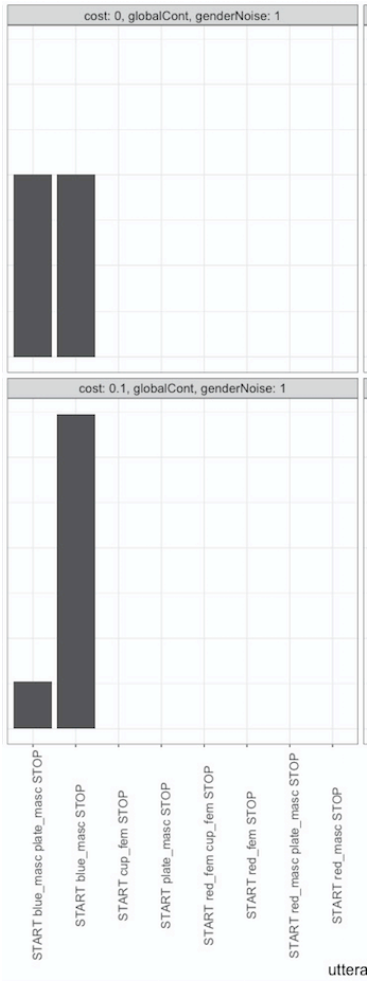
Scene 2
Color redundant
Gender match



Scene 3
Color redundant
Gender mismatch



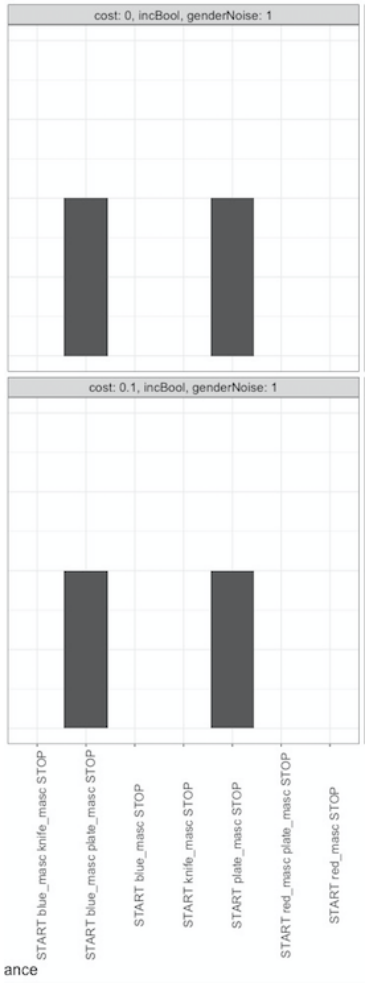
Scene 4
Color necessary
Gender mismatch



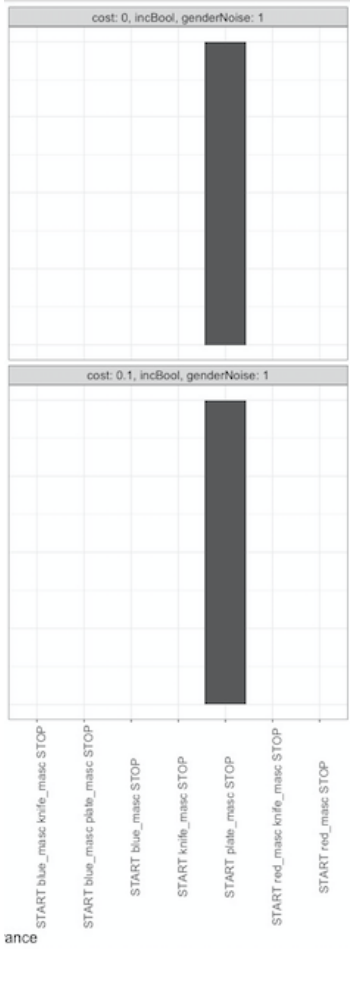
- Color overmodification (2 vs 3): more use of color adjective (combination of blue and blue plate) for gender mismatch than gender match conditions IF there is cost
- Noun omission (3 vs 4): higher rates of noun omission in color necessary than color redundant condition

Incremental boolean

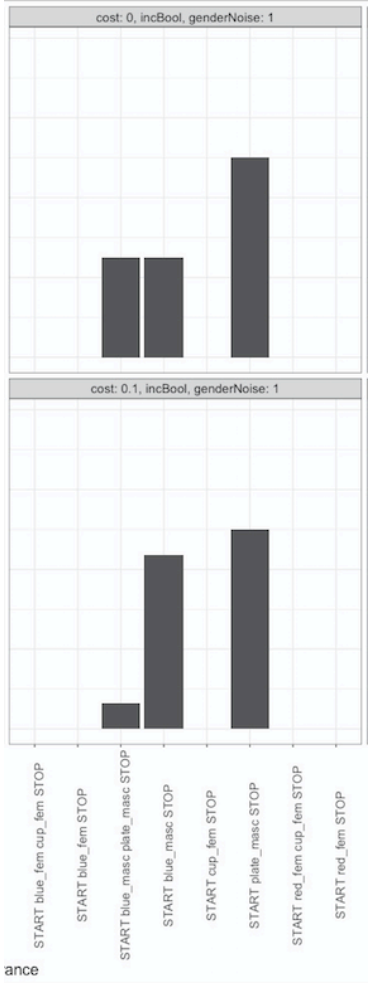
Scene 1
Color necessary
Gender match



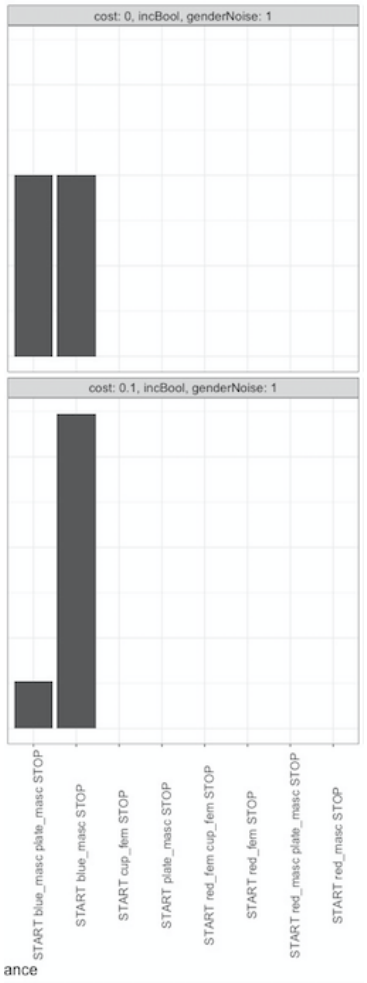
Scene 2
Color redundant
Gender match



Scene 3
Color redundant
Gender mismatch



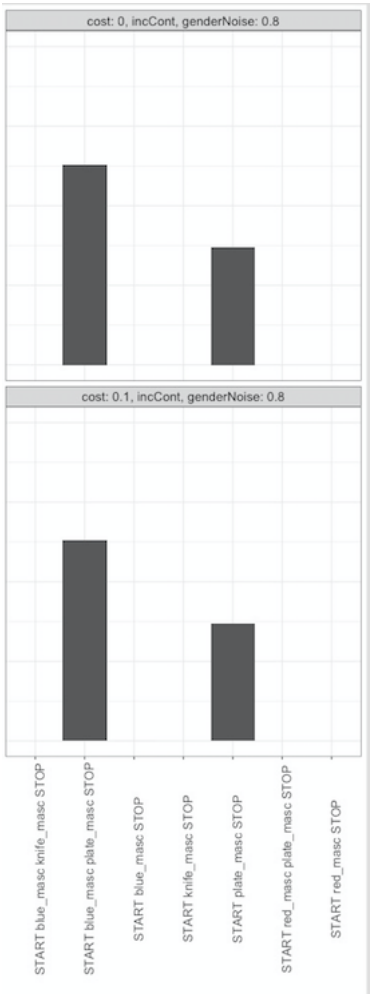
Scene 4
Color necessary
Gender mismatch



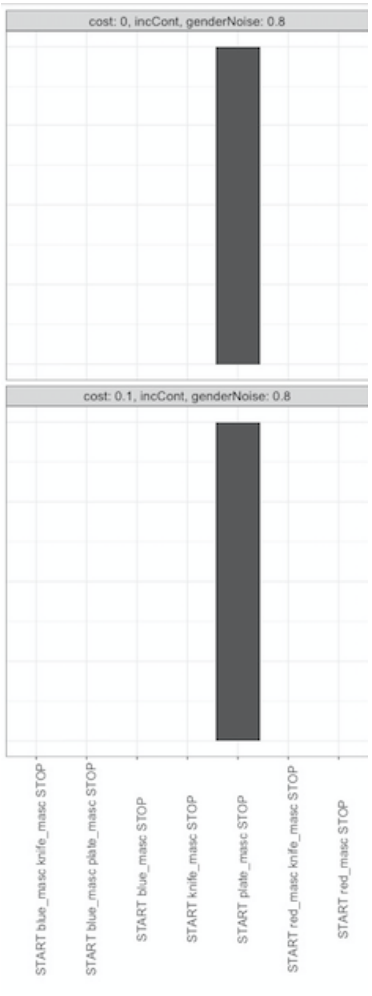
- Scene 1 is really strange → shouldn't we be seeing a 100% rate for "blue_plate". How did we reason about this?
- Color overmodification (2 vs 3) more color adjective overmodification for gender mismatch than gender match
- Noun omission (3 vs 4): higher rates of noun omission in color necessary than color redundant condition

Incremental Continuous

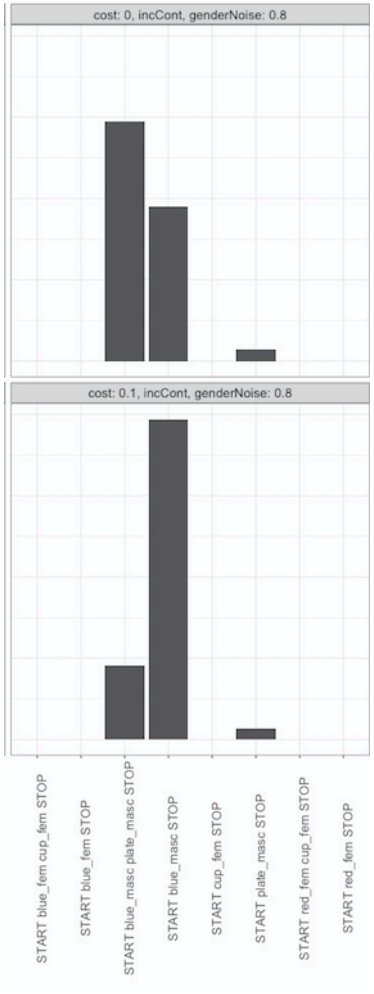
Scene 1
Color necessary
Gender match



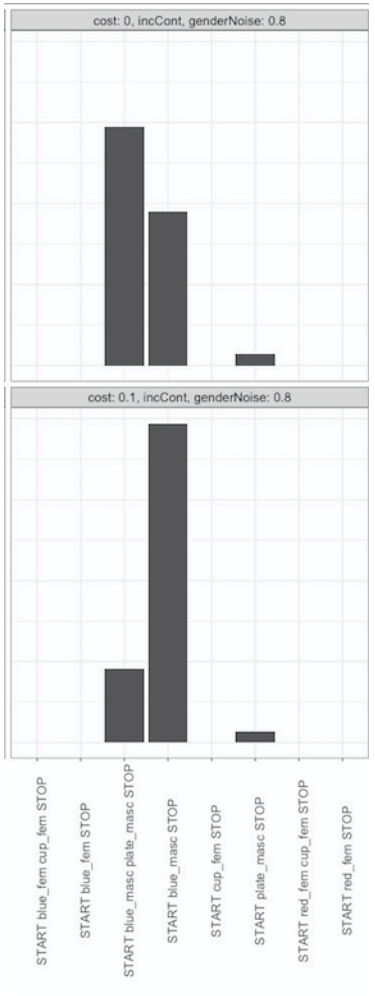
Scene 2
Color redundant
Gender match



Scene 3
Color redundant
Gender mismatch



Scene 4
Color necessary
Gender mismatch



- Again, why are we seeing just “plate” for scene 1???
- Color overmodification (2 vs 3): more color adjective overmodification for gender mismatch than gender match
- Noun omission (3 vs 4): higher rates of noun omission in color necessary than color redundant condition

Summary

The reason making these predictions is so difficult is that predictions can switch depending on which parameter values we assume.

However as I see it, the hypotheses are as follows:

If gender is used pragmatically, we expect higher rates of use of adjectives in the color redundant gender match scene (2) than in the color redundant gender mismatch scene (3), regardless of model.

When comparing scenarios that allow for noun omission (3 and 4), which both happen to be gender mismatch scenarios, we expect higher rates of noun omission in color necessary than color redundant conditions for all models.