Processing of Scalar Implicatures and Question Under Discussion

October 23, 2019

1 Methods

Participants

Participants (US IP Addresses, prior approval rating $> 95\%$) were recruited on Amazon Mechanical Turk and were paid \$2.3.

	Before Exclusions			After Exclusions		
	all QUD	any QUD	no QUD	all QUD	any QUD	no QUD
Experiment 1	50	50	50	41	35	38
Experiment 2	50	50	50	37	36	38
Experiment 3	200	200	200	fill	fill	fill
Experiment 4	400	400	400	281	256	313

Procedure and Materials

At the beginning of the experiment, participants in each group were presented with a different cover story. These cover stories were designed to establish an implicit QUD.

all QUD   Participants in all QUD group read a story explaining them that they are at a candy store and testing a row of special gumball machines. These gumball machines have 13 gumballs in their upper chambers and when they distribute gumballs, you see a certain number of gumballs move to the lower chamber. These machines also say how many gumballs you got ("You got X gumballs"), but are sometimes faulty in their report. The store worker's boss has threatened to fire him if the gumball machines are left empty and he cannot see the machines from the register and only tell how full they are by their statements. Participants were asked to help the store worker by telling him if this statement is right or wrong by pressing the yes or no key.

Is the machine empty? → Did I get all of the gumballs?

any QUD   In the any QUD group, participants read the same story except, in their story, the store worker's boss has threatened to fire him if the gumball machines get jammed.

Is the machine jammed? → Did I get none of the gumballs?

no QUD   Participants in no QUD group didn't read a cover story but instead read instructions that explained that they were going to see a gumball machine filled with gumballs and after a few seconds certain number of gumballs will move to its bottom chamber and the machine will say how many gumballs you got. Participants were asked to say if this statement right or wrong by pressing the yes or no key.

Is the statement correct? → No implicit QUD

After reading the cover story, participants went through a scripted demonstration that showed various scenarios and the store worker's reaction to their responses. In the all QUD condition, when all of gumballs dropped from the

upper chamber to the lower chamber and the machine reported "You got all of the gumballs", participants were asked to press the yes key and they read that they let the store worker know that the machine is empty and he therefore knows that he needs to refill the machine. They also read that one time someone got all of the gumballs and was told "You got all of the gumballs", they pressed the no key and the store worker was fired because he didn't refill the machine. In the any QUD condition, participants got no gumballs and the machine reported "You got none of the gumballs". Participants were asked to press the yes key and read that they let the store worker know that the machine is not distributing gumballs and that he needs to fix the machine. They read that one time someone got 0 gumballs and was told "You got none of the gumballs", they pressed the no key and the store worker was fired because he didn't repair the machine. In the no QUD condition, the only feedback that participants got was a sentence that said whether they agreed or disagreed with the statement.

To ensure that participants paid attention to the cover story, participants in all QUD and any QUD conditions were asked a multiple-choice question about when the store worker will be fired. For the all QUD group the correct answer was "when the machines are empty" and for the any QUD group the correct answer was "when the machines jam". When participants answered this question incorrectly they were presented with the cover story again and went through the demonstration. Halfway through the experiment, participants were asked to answer this multiple-choice question again. This time, when they answered incorrectly they were only asked to try again. This was done to prevent the decay of the implicit QUD over time.

There were 4 practice trials with "You got all of the gumballs" and "You got none of the gumballs" statements. On 2 of these trials, the statement was correct, and on 2 of them it was incorrect.

After the practice trials, there were 72 experimental trials. The number of gumballs in the lower chamber (0 to 13 gumballs) and the quantifier in the statement (some, none, all) were varied. On 32 of the trials the expected answer was yes, and on 32 of the trials the expected answer was no. The remaining 8 trials were occurrences of the critical trial and the main focus of this experiment. On these trials, all thirteen of the gumballs dropped to the lower chamber and participants heard the statement "You got some of the gumballs". This statement is true when it's interpreted semantically as "You got some and possibly all of the gumballs" but when it's interpreted pragmatically as "You got some but not all of the gumballs", it's false.

INSERT TABLE HERE - STIMULI LIST

Experiments differed from each other in the following ways:

Experiment 1. In the critical condition, participants heard "You got some gumballs".

Experiment 2. In the critical condition, participants heard "You got some of the gumballs".

Experiment 3. Participants were prescreened for age. Only participants aged between 18 and 25 and participants older than 45 were recruited. In the critical condition, participants heard "You got some of the gumballs".

Experiment 4. Participants had 4 seconds to respond after they heard the statement. In the critical condition, participants heard "You got some of the gumballs"

Experiment 5. Participants had 4 seconds to respond after they heard the statement. In the critical condition, participants heard "You got some gumballs".

2 Results

Exclusions

The exclusion criteria and the number of participants excluded with each criterion are given below:

1. non-native English speakers

2. participants who get at least two practice trials wrong

3. participants who get the audio check wrong more than one once

4. participants who get the second comprehension question wrong more than twice

5. participants with accuracy of lower than 85% on non-critical trials

6. response times that are 2.5 standard deviations above the mean for that condition or that are faster than the onset of the quantifier.

	Experiment 1	Experiment 2	Experiment 3*	Experiment 4
1	0	3	13	37
2	32	24	23**	93**
3	0	0	1	5
4	0	1	0	15
5	4	7	26	200***
total	**36**	**35**	**62**	**350**

*7 participants were excluded because their age didn't match the age they reported in prescreening.
**Criterion was changed to only exlude participants who got more than two practice trials wrong.
**Criterion was changed to only exclude participants with accuracy lower

than 85% on non-critical trials with quantifiers "some", "all", "none", and numbers below 6.

Judgements

In all three experiments, in the critical condition, participants in the all QUD group gave more pragmatic responses than the participants in the any QUD group and participants in the any QUD group gave more pragmatic responses than the no QUD group.
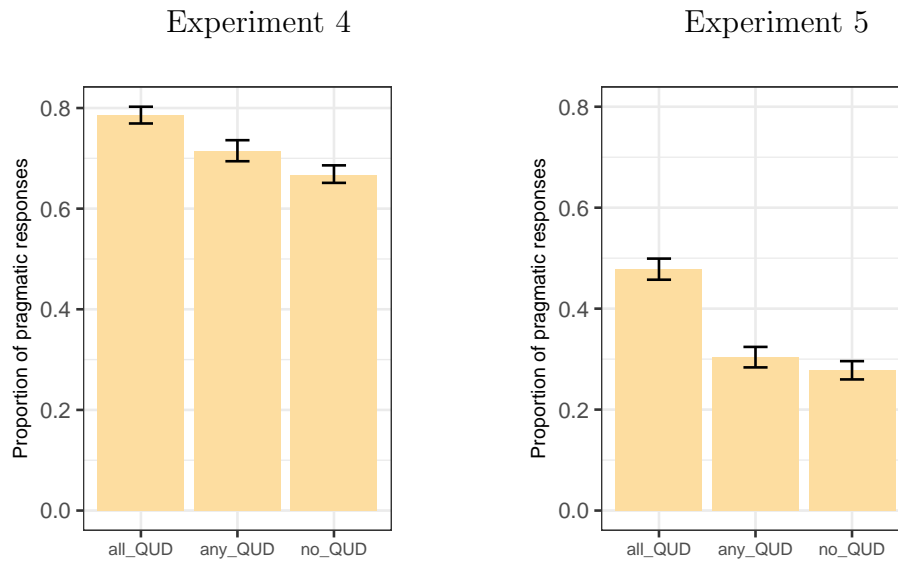

Figure 1: Proportion of pragmatic responses on critical trials for the three experiments

cover story → implicit QUD → likelihood of pragmatic responses → scaler implicature

A mixed effects logistic regression predicting response type with random by-

participants intercepts from fixed effects of QUD found a main effect of QUD such that there are more pragmatic (NO) responses for all QUD compared to any QUD and no QUD.

Experiment 4

```
Fixed effects:
                          Estimate Std. Error z value             Pr(>|z|)
(Intercept)               -5.2821     0.5325  -9.920 < 0.0000000000000002 ***
Answer.conditionany_QUD    1.4246     0.5498   2.591             0.009564 **
Answer.conditionno_QUD     2.0248     0.5277   3.837             0.000125 ***
---
```

Experiment 5

```
Fixed effects:
                          Estimate Std. Error z value             Pr(>|z|)
(Intercept)               -5.2821     0.5325  -9.920 < 0.0000000000000002 ***
Answer.conditionany_QUD    1.4246     0.5498   2.591             0.009564 **
Answer.conditionno_QUD     2.0248     0.5277   3.837             0.000125 ***
---
```

No context $\rightarrow$ more uncertainity $\rightarrow$ more bimodal $\rightarrow$ statistical test for bimodality???
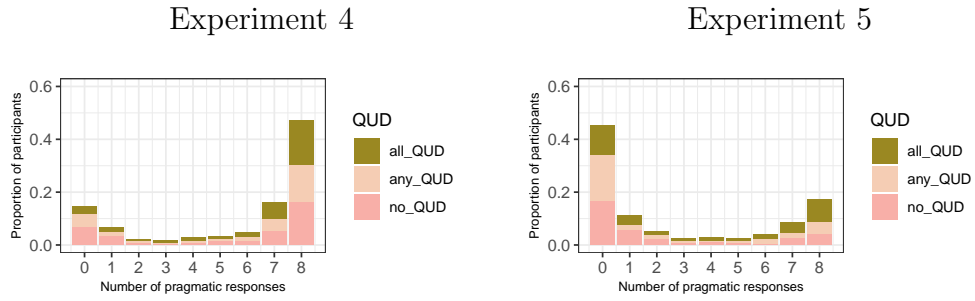

Figure 2: Distribution of participants over number of pragmatic responses on critical trials

Response Times

The QUD affected the response time for semantic and pragmatic responses differently. The more relevant the alternative, the faster the pragmatic responses became and the slower the semantic responses became.
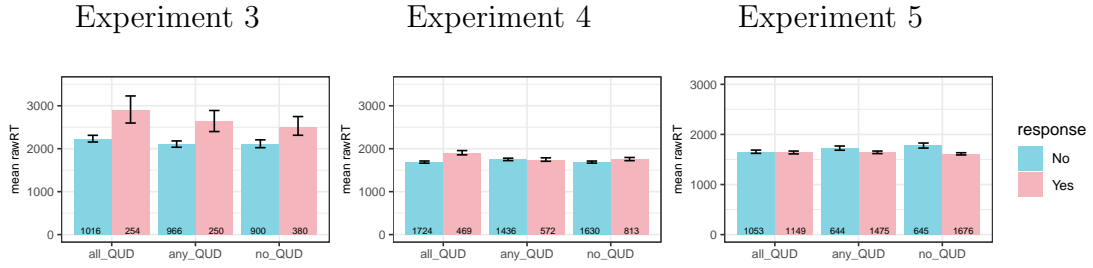

Figure 3: Mean response times for semantic and pragmatic responses on critical trials as a function of QUD

A mixed effects linear regression model with random by-participant intercepts predicting log-transformed response time from fixed effects of QUD, response type and their interaction, revealed a main effect of response type such that NO responses were always faster than YES responses (response type was centered before entering the analysis).

Experiment 3

```
Fixed effects:
                             Estimate Std. Error       df t value            Pr(>|t|)
(Intercept)                   7.64706    0.02061 468.82920 371.125 < 0.0000000000000002 ***
Answer.conditionany_QUD      -0.04627    0.02946 468.25786  -1.571              0.1169
Answer.conditionno_QUD       -0.06288    0.02915 470.08781  -2.157              0.0315 *
ckey                          0.14992    0.03826 2172.32325  3.918            0.000092 ***
Answer.conditionany_QUD:ckey  0.02457    0.05446 2167.16503   0.451              0.6519
Answer.conditionno_QUD:ckey  -0.02393    0.05257 1824.43899  -0.455              0.6489
---
```

Experiment 4

```
Fixed effects:
                              Estimate  Std. Error         df t value             Pr(>|t|)
(Intercept)                   7.423645    0.011042  848.873171 672.324 < 0.0000000000000002 ***
Answer.conditionany_QUD       0.002145    0.015932  842.005407   0.135               0.8929
Answer.conditionno_QUD       -0.022961    0.015206  845.505595  -1.510               0.1314
ckey                          0.117455    0.018031 5074.120647   6.514     0.0000000000802 ***
Answer.conditionany_QUD:ckey -0.102870    0.025684 4556.375058  -4.005     0.0000629682595 ***
Answer.conditionno_QUD:ckey  -0.062684    0.024409 4339.537644  -2.568               0.0103 *
---
```

Experiment 5

Responder Type

As a post hoc analysis, participants were divided into two groups, semantic
and pragmatic responders. Semantic responders are participants who gave
semantic responses more than half of the time (in critical condition) and
pragmatic responders are participants who gave pragmatic responses more
than half of the time. Participants who gave the same number of pragmatic
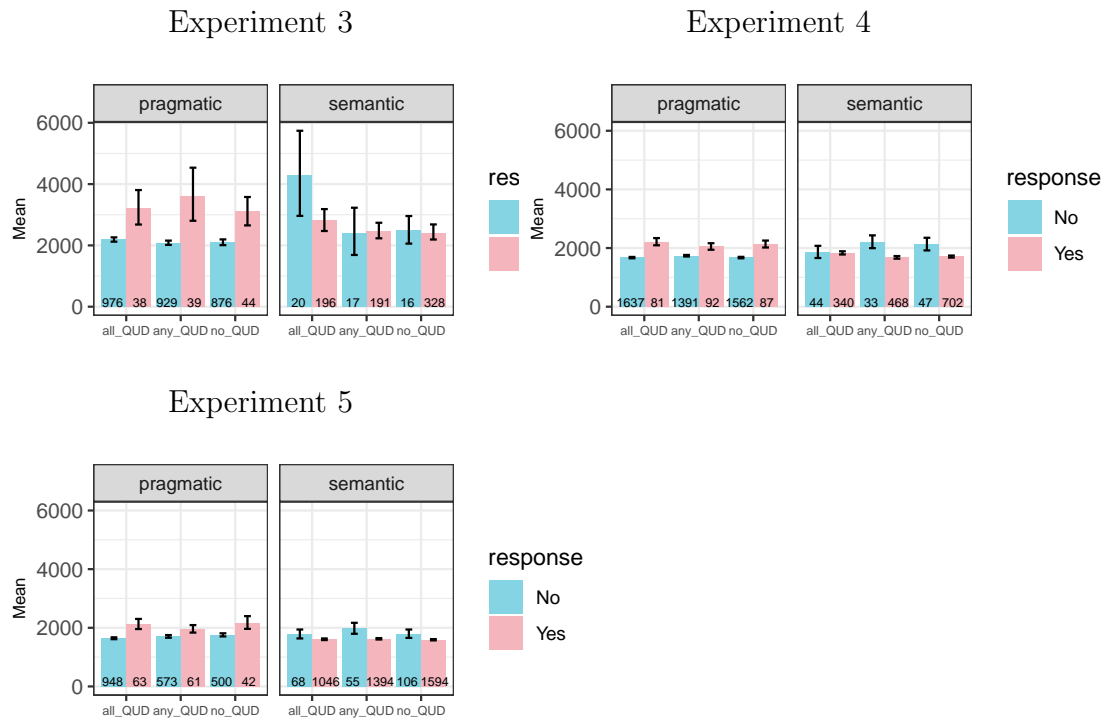and semantic responses were excluded from this visualization and analysis.

Figure 4: Mean response times for semantic and pragmatic responders (inconsistent responders excluded)