

Effects on projectivity ratings by Embedding Operator and Trigger — Data Analysis

Hofmann, deMarneffe, Tonhauser

December 6, 2022

Contents

1	Introducing the dataset	1
2	Statistical Summaries and Graphs	4
2.1	Projectivity rating by operator	4
2.2	Distributions of projectivity rating by verb with means and 95% bootstrapped confidence interval:	5
2.3	Distributions, means, and CIs for projectivity ratings by verb, for each operator:	6
2.4	Means and confidence intervals for projectivity rating by verb and operator:	7
2.5	Projectivity ratings by operator, for each verb (Verb profiles)	8

1 Introducing the dataset

```
> str(df)

'data.frame': 57160 obs. of 9 variables:
 $ workerid      : int  1 1 1 1 1 1 1 1 1 1 ...
 $ content       : chr  "charley" "danny" "emily" "emma" ...
 $ short_trigger : chr  "acknowledge" "hear" "reveal" "discover" ...
 $ ai_block      : chr  "block1" "block1" "block1" "block1" ...
 $ ai            : num  0.98 0.99 0.99 0.99 0.98 0.98 1 0.99 0.99 0.99 ...
 $ projective    : num  0.3 0.98 0.01 0.99 0.98 0.99 0.01 0.01 0.27 0.01 ...
 $ verb         : chr  "acknowledge" "hear" "reveal" "discover" ...
 $ op            : chr  "q" "q" "q" "q" ...
 $ exp_block     : int  1 1 1 1 1 1 1 1 1 1 ...

> df$workerid <- as.factor(df$workerid)
> length(levels(df$workerid))

[1] 2682
```

The dataset consists of 57160 observations from 2682 participants (recruited on the online platforms Prolific and Amazon Mechanical Turk), across 12 experiments.

We are interested in how highly participants rate speaker commitment to the truth of an embedded complement clause, coded as projective on a real-numbered sliding scale between 0 – 1.

The complement clause was embedded under an attitude verb, which in turn was embedded under an entailment-cancelling operator. Our fixed effects factors manipulate the following:

1. The choice of attitude verb (coded as verb)
2. The entailment-cancelling operator (coded as op)

The levels for our fixed effects factors are the following:

```

> df$verb <- as.factor(df$verb)
> levels(df$verb)

[1] "acknowledge" "admit"      "announce"    "be_annoyed" "be_right"
[6] "confess"     "confirm"    "demonstrate" "discover"    "establish"
[11] "hear"        "inform"     "know"        "pretend"     "prove"
[16] "reveal"      "say"        "see"         "suggest"     "think"

> length(levels(df$verb))

[1] 20

> df$op <- as.factor(df$op)
> levels(df$op)

[1] "c" "m" "n" "q"

> length(levels(df$op))

[1] 4

```

We are interested in the effect on projective of verb and op, as well as their interaction, corresponding to a 20×4 factorial design, yielding

```

> length(levels(df$verb))*length(levels(df$op))

[1] 80

```

conditions.

We have 20 items, corresponding to the content of the complement clause.

```

> df$content <- as.factor(df$content)
> levels(df$content)

[1] "charley" "danny"   "emily"   "emma"    "frank"   "grace"
[7] "isabella" "jackson" "jayden"  "jon"     "josh"    "josie"
[13] "julian"  "mary"    "mia"     "olivia"  "owen"    "sophia"
[19] "tony"    "zoe"

> length(levels(df$content))

[1] 20

```

We have roughly 36 observations by item and condition. This is an approximate number, because the op manipulation is a between-studies manipulation, and the number of participants differs by experiment:

```

> # n observations
> length(df[,1])

[1] 57160

> # observations by item
> length(df[,1])/length(levels(df$content))

[1] 2858

> table(df$content)

charley  danny  emily  emma  frank  grace isabella  jackson

```

2858	2858	2858	2858	2858	2858	2858	2858
jayden	jon	josh	josie	julian	mary	mia	olivia
2858	2858	2858	2858	2858	2858	2858	2858
owen	sophia	tony	zoe				
2858	2858	2858	2858				

```
> # observations by verb
```

```
> length(df[,1])/length(levels(df$verb))
```

```
[1] 2858
```

```
> table(df$verb)
```

acknowledge	admit	announce	be_annoyed	be_right	confess
2858	2858	2858	2858	2858	2858
confirm	demonstrate	discover	establish	hear	inform
2858	2858	2858	2858	2858	2858
know	pretend	prove	reveal	say	see
2858	2858	2858	2858	2858	2858
suggest	think				
2858	2858				

```
> # observations by operator
```

```
> length(df[,1])/length(levels(df$op))
```

```
[1] 14290
```

```
> table(df$op)
```

c	m	n	q
14400	14680	14340	13740

```
> # observations by item and condition
```

```
> length(df[,1])/length(levels(df$content))/
```

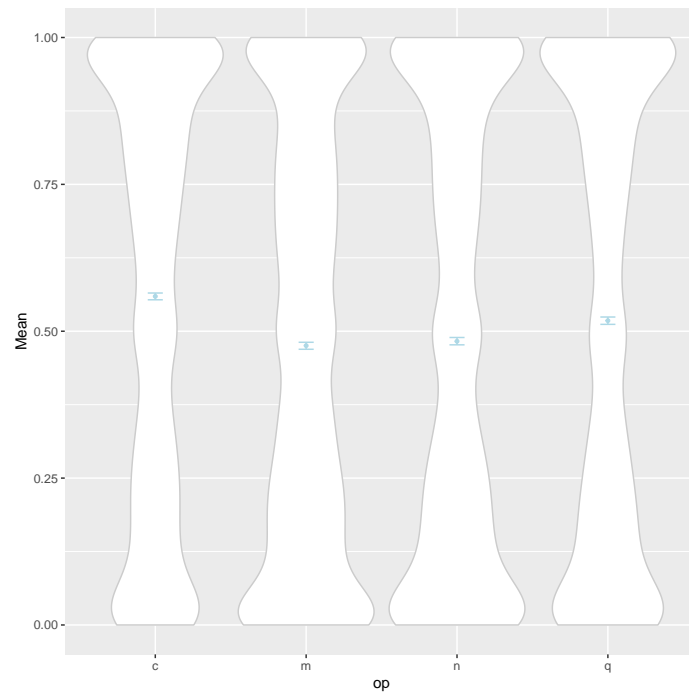
```
+ (length(levels(df$verb))*length(levels(df$op)))
```

```
[1] 35.725
```

2 Statistical Summaries and Graphs

2.1 Projectivity rating by operator

Distribution of projectivity ratings by operator with means and 95% bootstrapped confidence intervals.

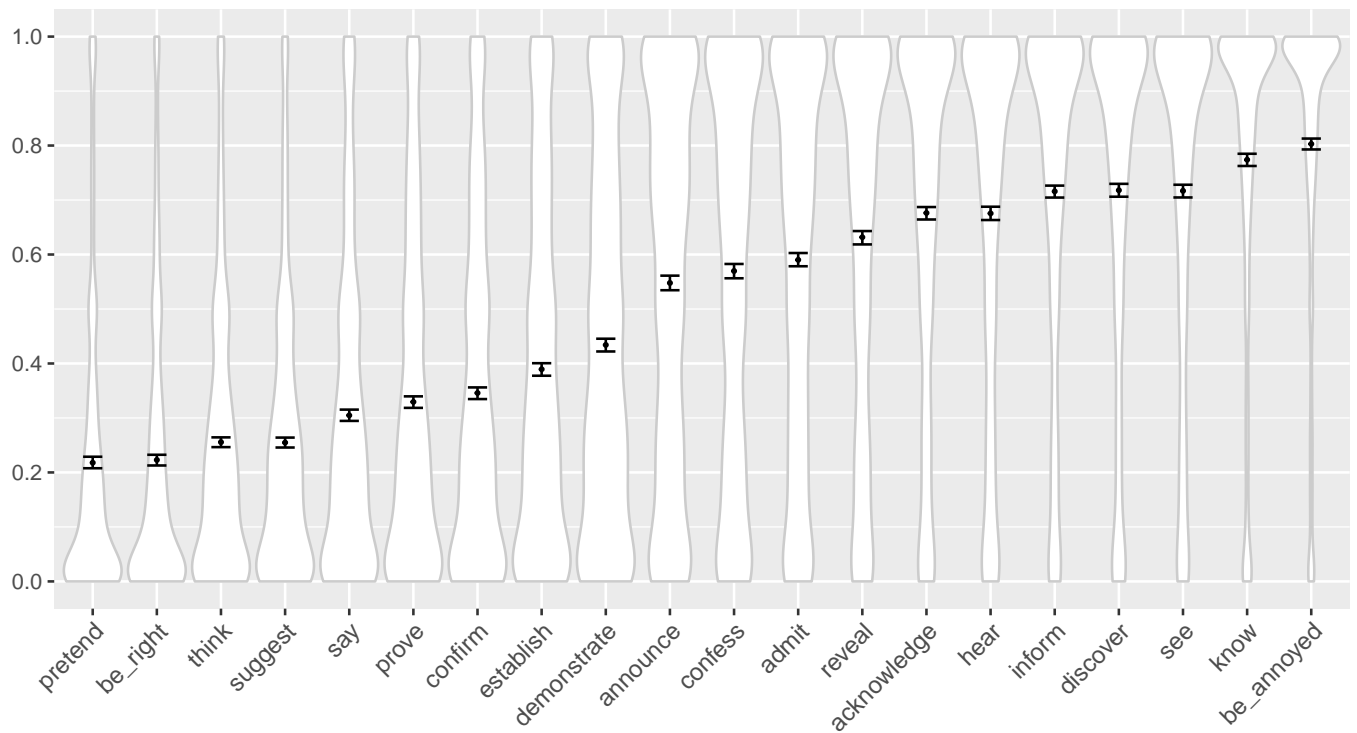


The following generalizations emerge:

- Conditionals have the highest projectivity ratings
- Projectivity ratings for questions are higher than those for modals and negation, but lower than those for conditionals
- Modals and negation have the lowest projectivity ratings
- The ratings for negation look a little higher than for modals, but error bars overlap

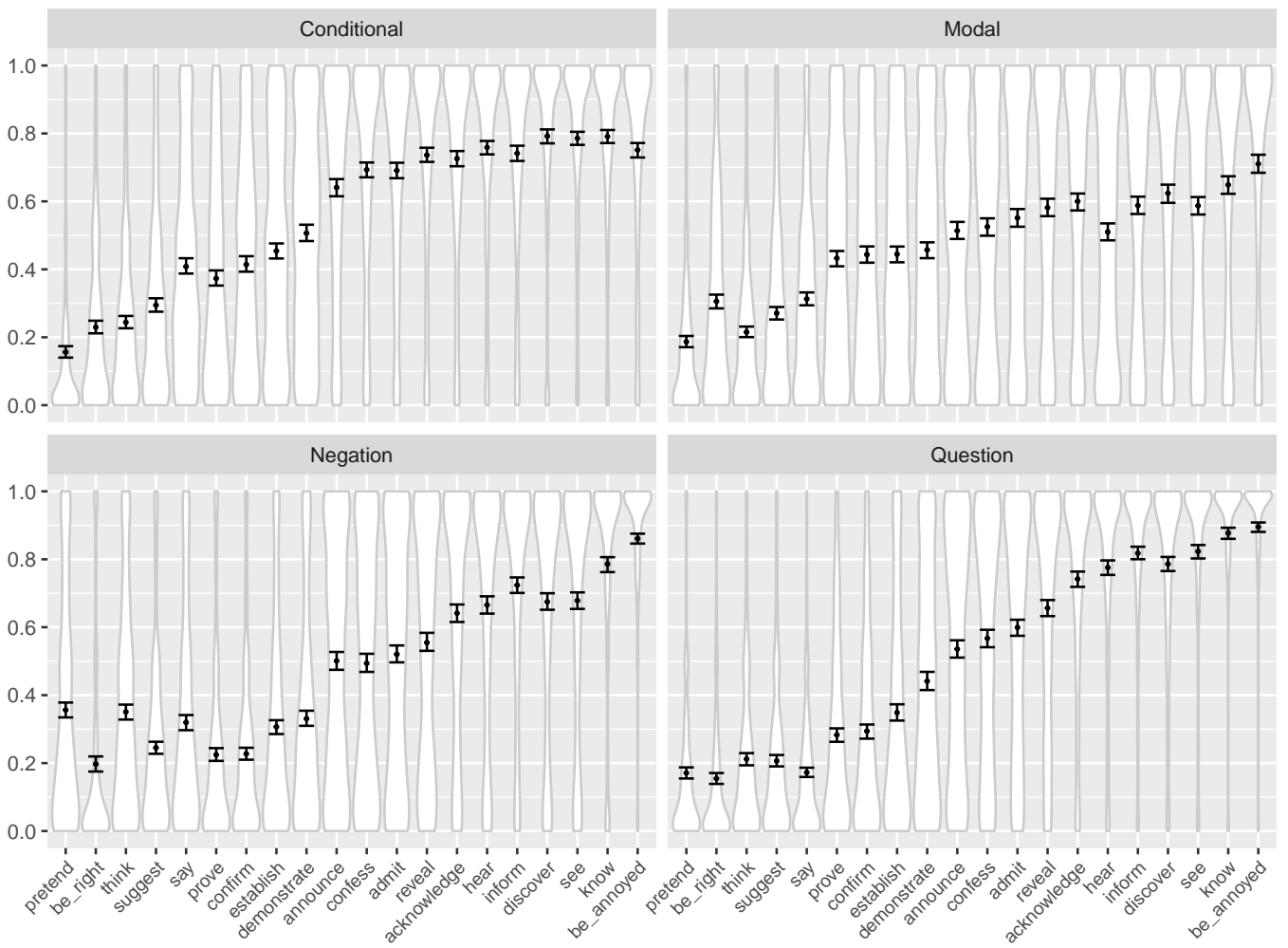
Although these differences appear to be significant, they are quite small.

2.2 Distributions of projectivity rating by verb with means and 95% bootstrapped confidence interval:

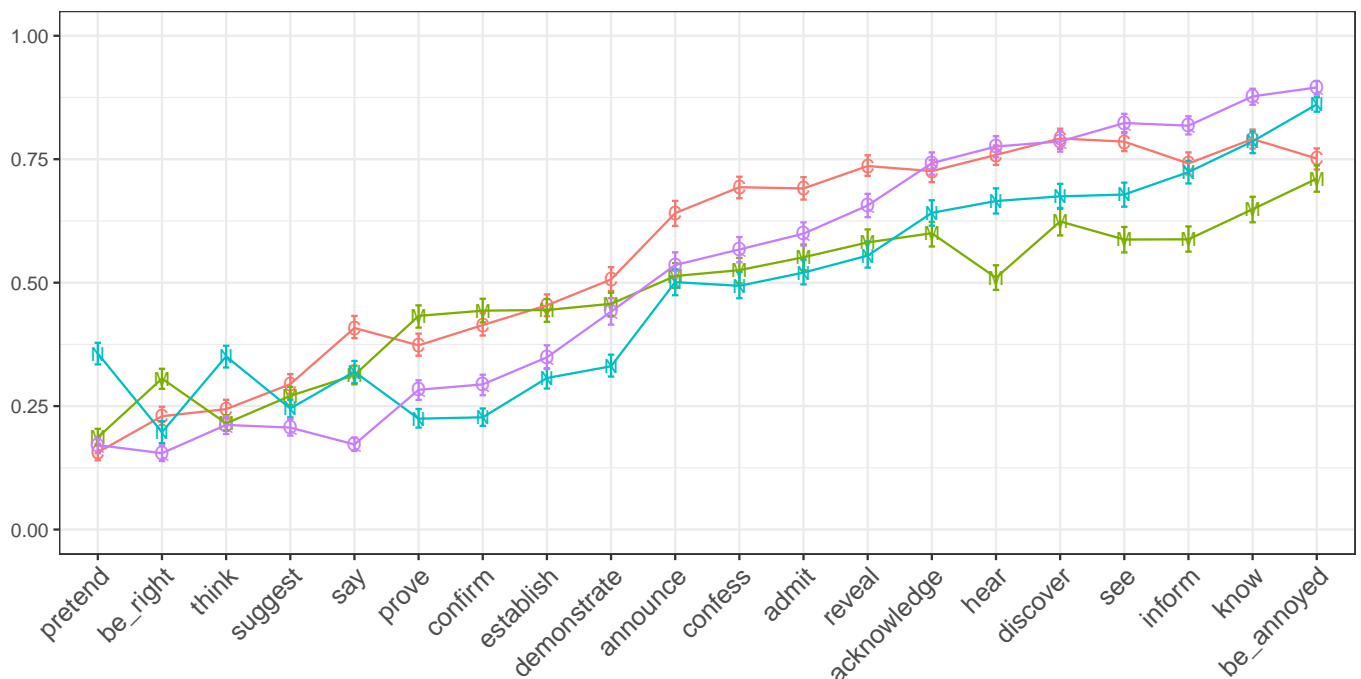


- We see gradual differences in projectivity between verbs

2.3 Distributions, means, and CIs for projectivity ratings by verb, for each operator:



2.4 Means and confidence intervals for projectivity rating by verb and operator:

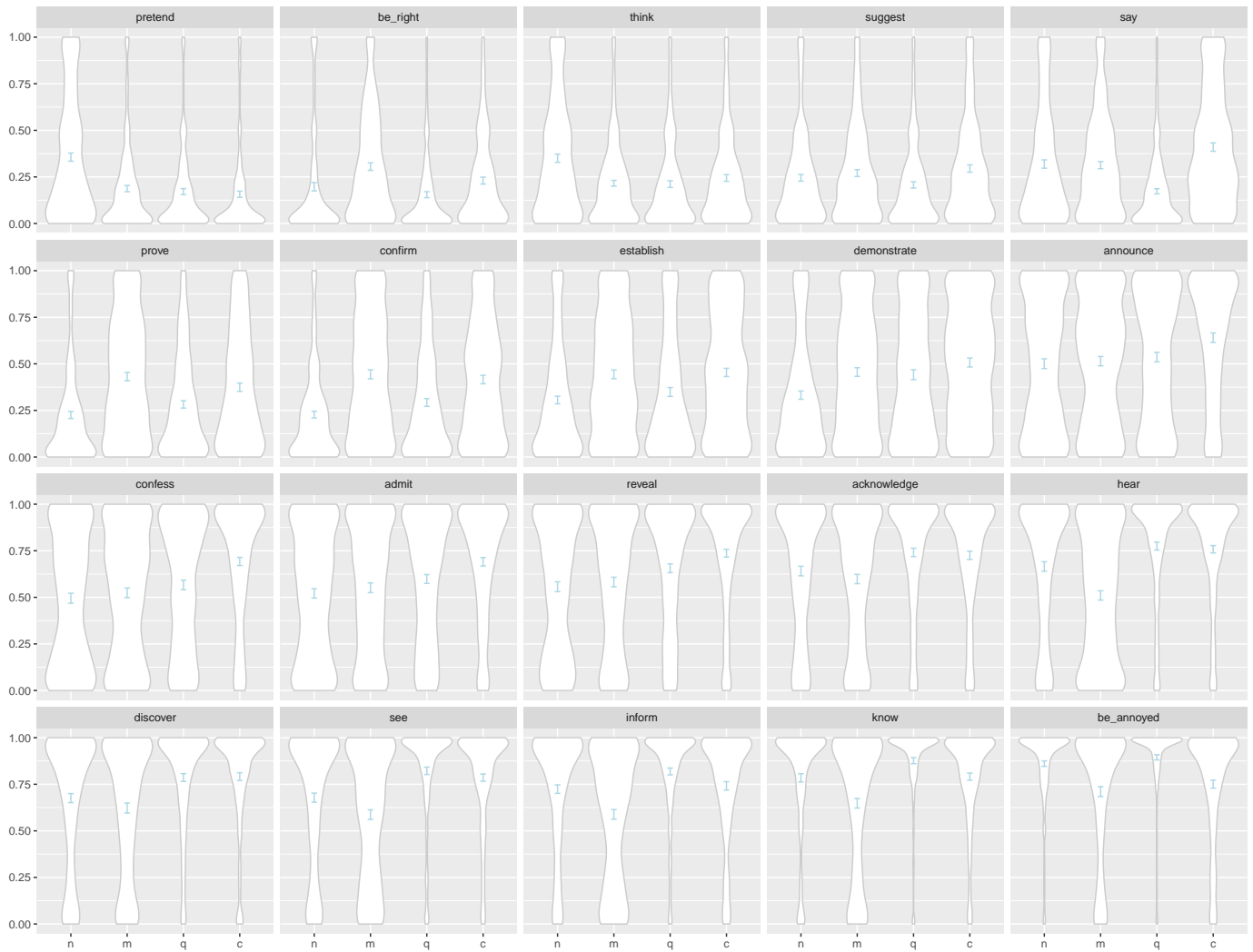


- We see interactions between verb and operator
- Two verbs show highest projectivity under negation: the anti-veridical *pretend*, and the non-veridical *think*. These are verb with relatively low overall projectivity.
- More projective verbs ('*announce*' and above) have $C > M$
- Highly projective verbs ('*hear*' and above) have $N > M$
- We do not see any group of verbs that could be characterized as 'semi-factive' in the sense of Karttunen.
 - Specifically, *discover* does not follow the predicted pattern: It is not more projective under negation, but most projective in conditionals and questions.
 - For Karttunen's 'factives', no difference between operators is expected.
 - Kajsa Djärv about this distinction: cognitive predicates are semi-factive, and emotives are factive. The pattern suggested by karttunen is also not found here (if anything, it might be the other way round)
- For more generalizations, let's look at the same information plotted differently: By-operator projectivity for each verb

2.5 Projectivity ratings by operator, for each verb (Verb profiles)

Version 1: With distributions of projectivity ratings as violin-plot

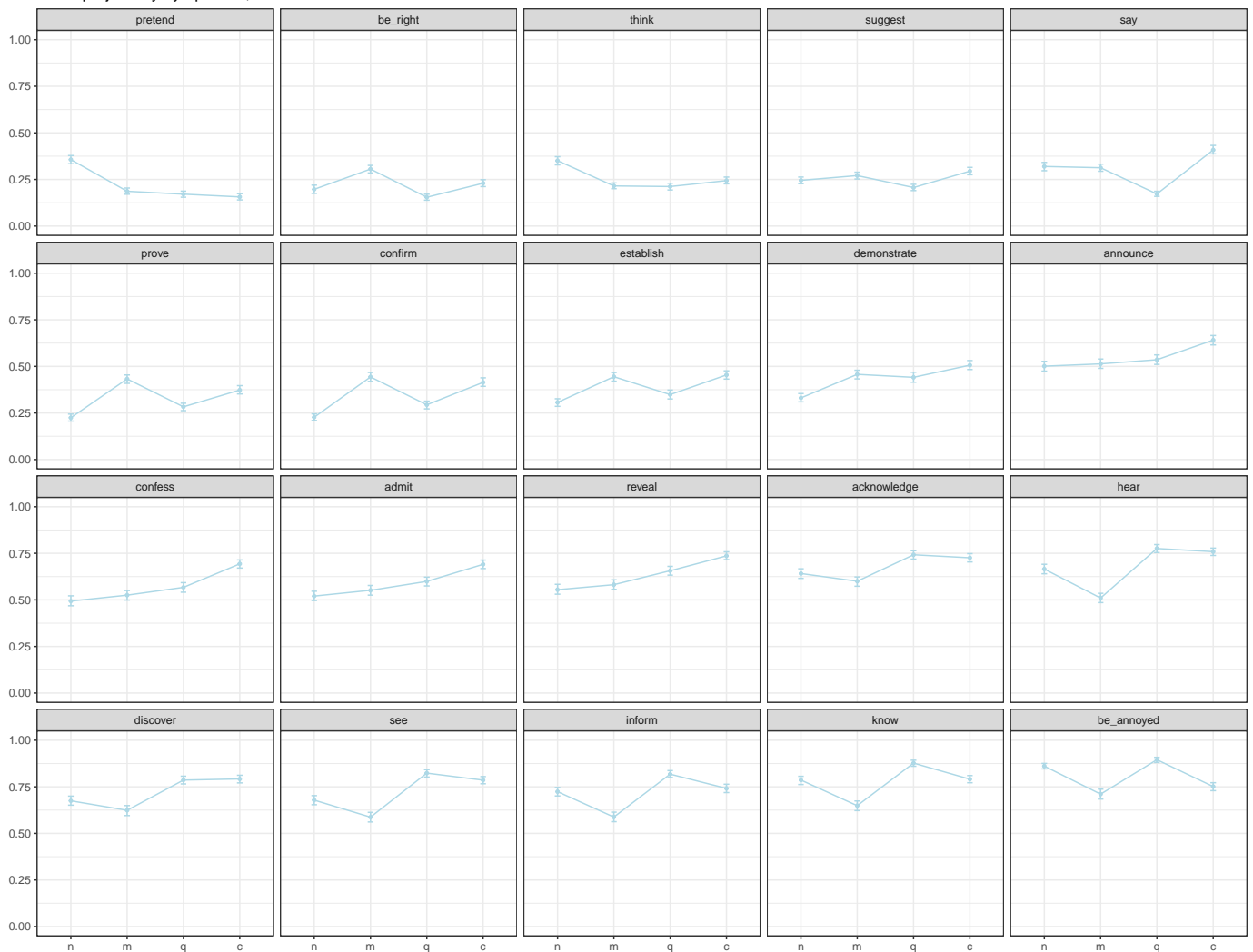
Mean projectivity by operator, for each verb



Some more generalizations: We find different ‘profiles’ for different verbs, of how embedding operators affect projectivity based on the verb. Groups of verbs show similar profiles:

Version 2: With lines to show profiles

Mean projectivity by operator, for each verb



- *pretend, think*: **anti-veridical profile**
N > M, Q, C, overall low projectivity
- *acknowledge, hear, inform, see, discover, know, be annoyed*: **'factive' profile**
Q > N, C > M, overall high projectivity (it may be possible to find further subgroups here)
- *prove, confirm, establish, demonstrate, (announce), confess, admit, reveal*: **veridical profile**
M, C > Q > N, overall med-lo to med-hi projectivity
- *be right, suggest*: **reportative profile**
M, N, C > Q

Maybe these can have better names, not trying to suggest that verbs can neatly divided in factive v non-factive, but potentially this class / profile is what prompted intuitions in previous literature, and naming in this tradition could make sense, but can be changed depending on our rhetoric, of course.

Mean projectivity by operator, for each verb

