Narratives, Actions and Impact on Dishonesty

Despoina Alempaki (Warwick), Valeria Burdea (LMU), Xueheng Li (Sun Yat-Sen)

23rd Conference on Research on Economic Theory and Econometrics

Tinos, 10th July 2025

Introduction

- Private information at the heart of many economic activities
- People are often averse to lying (e.g., Abeler et al., 2019), but there is high heterogeneity (e.g., Gibson et al., 2013)
  - US financial investors: 7% disciplined for fraud (e.g., Egan et al., 2008)
  - Auto repair industry: over 50% unnecessary (e.g., Wolinksy 1993; 1995)

Introduction

- Private information at the heart of many economic activities
- People are often averse to lying (e.g., Abeler et al., 2019), but there is high heterogeneity (e.g., Gibson et al., 2013)
  - US financial investors: 7% disciplined for fraud (e.g., Egan et al., 2008)
  - Auto repair industry: over 50% unnecessary (e.g., Wolinksy 1993; 1995)
- People often don't act in isolation
  - Dishonesty-promoting advice increases lying in individual tasks (Leib et al., 2023)
  - Groups lie more than individuals because groups exchange more negative than positive arguments (Kocher et al., 2018)

Introduction

- Private information at the heart of many economic activities
- People are often averse to lying (e.g., Abeler et al., 2019), but there is high heterogeneity (e.g., Gibson et al., 2013)
  - US financial investors: 7% disciplined for fraud (e.g., Egan et al., 2008)
  - Auto repair industry: over 50% unnecessary (e.g., Wolinksy 1993; 1995)
- People often don't act in isolation
  - Dishonesty-promoting advice increases lying in individual tasks (Leib et al., 2023)
  - Groups lie more than individuals because groups exchange more negative than positive arguments (Kocher et al., 2018)
- Narratives: explanations shape how individuals make sense of the world (e.g., Aina, 2024; Andre, 2023; Barron and Fries, 2023; Barron et al., 2023; Benabou et al., 2020; Eliaz and Spiegler, 2020; Graeber et al., 2024; Harrs et al, 2021; Roos and Reccious, 2024; Schwartzstein and Sunderam, 2021; Shiller, 2017; 2020)
  - **Arguments used to justify different courses of action on the basis of morality** (Benabou et al., 2020)

Introduction

- Imagine you are about to submit your workplace expenses report

Introduction

- Imagine you are about to submit your workplace expenses report
- A recent discussion with a co-worker comes to mind
  - "*Exaggerating business expenses on a report is a natural response if the company underpays anyway. This is just a way of getting what one deserves*." **NEGATIVE**

Introduction

- Imagine you are about to submit your workplace expenses report
- A recent discussion with a co-worker comes to mind
  - "*Exaggerating business expenses on a report is a natural response if the company underpays anyway. This is just a way of getting what one deserves*." **NEGATIVE**
- Will be exposed to this justification have any **effect** on your **own** behaviour?

Introduction

- Imagine you are about to submit your workplace expenses report
- A recent discussion with a co-worker comes to mind
  - "*Exaggerating business expenses on a report is a natural response if the company underpays anyway. This is just a way of getting what one deserves*." **NEGATIVE**
- Will be exposed to this justification have any **effect** on your **own** behaviour?
  - "*Accurate expense reporting demonstrates integrity and fairness. The peace of mind from ethical consistency outweighs any financial gain from deception*." **POSITIVE**

This project

RQ1: What is the effect of **positive** and **negative** narratives on dishonesty?

This project

RQ1: What is the effect of **positive** and **negative** narratives on dishonesty?

RQ2: Is their effect solely driven by their **action salience**?

This project

RQ1: What is the effect of **positive** and **negative** narratives on dishonesty?

RQ2: Is their effect solely driven by their **action salience**?

RQ3: Does their influence depend on the **impact** of the immoral action?

Related Literature

- Traditional models of communication or persuasion
  - e.g., Crawford and Sobel (1982), Milgrom (1981), Kamenica and Gentzkow (2011)
- Narrative theories
  - Shiller (2017), Benabou and Tirole (2018), Benabou et al. (2020), Eliaz and Spiegler (2020), Schwartzstein and Sunderam (2021)
- Experimental work of hard information on behavior
  - e.g., Abeler et al. (2019), Erat and Gneezy (2012), Gneezy et al. (2018)
- Recent experiments on narratives
  - Barron et al. (2023), Graeber et al. (2024), Hillenbrand and Verrina (2022), Morag and Segal (2024)

Related Literature

- Traditional models of communication or persuasion
  - e.g., Crawford and Sobel (1982), Milgrom (1981), Kamenica and Gentzkow (2011)
- Narrative theories
  - Shiller (2017), Benabou and Tirole (2018), Benabou et al. (2020), Eliaz and Spiegler (2020), Schwartzstein and Sunderam (2021)
- Experimental work of hard information on behavior
  - e.g., Abeler et al. (2019), Erat and Gneezy (2012), Gneezy et al. (2018)
- Recent experiments on narratives
  - Barron et al. (2023), Graeber et al. (2024), Hillenbrand and Verrina (2022), Morag and Segal (2024)
- Narrative effects on strategic deception
  - Narratives can change the perceived lying cost of one's immoral action

Experimental Design

Experimental Design

- 2x5 between-subject varying
  - **Impact**: HIGH vs LOW
  - **Social info**: None, Positive Narratives, Negative Narratives
    Positive Actions, Negative Actions

Main task – the deception game (Erat and Gneezy, 2011)

- Computer randomly draws one number from {1, 2, 3, 4, 5, 6}

Main task – the deception game (Erat and Gneezy, 2011)

- Computer randomly draws one number from {1, 2, 3, 4, 5, 6}
- S observes this number, but R does not (common knowledge)

Main task – the deception game (Erat and Gneezy, 2011)

- Computer randomly draws one number from {1, 2, 3, 4, 5, 6}
- S observes this number, but R does not (common knowledge)
- R then chooses a number between 1 and 6

Main task – the deception game (Erat and Gneezy, 2011)

- Computer randomly draws one number from {1, 2, 3, 4, 5, 6}
- S observes this number, but R does not (common knowledge)
- R then chooses a number between 1 and 6
- Payoffs – depend on impact and on whether R's choice matches the draw
  - R knows: Option A, if choice = draw; Option B, if choice ≠ draw

Main task – the deception game (Erat and Gneezy, 2011)

- Computer randomly draws one number from {1, 2, 3, 4, 5, 6}
- S observes this number, but R does not (common knowledge)
- R then chooses a number between 1 and 6
- Payoffs – depend on impact and on whether R's choice matches the draw
  - R knows: Option A, if choice = draw; Option B, if choice ≠ draw
- S sends a message to R from

Message 1: *"The outcome from the draw is 1."*
Message 2: *"The outcome from the draw is 2."*
Message 3: *"The outcome from the draw is 3."*
Message 4: *"The outcome from the draw is 4."*
Message 5: *"The outcome from the draw is 5."*
Message 6: *"The outcome from the draw is 6."*

Impact manipulation

	S 	R 
LOW		
Option A (R choice = draw)	£1.00	£1.00
Option B (R choice ≠ draw)	£3.50	**£0.90**

Impact manipulation

LOW	S 	R 
Option A (R choice = draw)	£1.00	£1.00
Option B (R choice ≠ draw)	£3.50	**£0.90**

HIGH	S 	R 
Option A (R choice = draw)	£1.00	£1.00
Option B (R choice ≠ draw)	£3.50	**£0.10**

Social Information manipulation

- **Narratives**
  - “Before you make your decision, please read the **explanations two other participants had given for their decisions** in similar previous tasks.”

Social Information manipulation

- **Narratives**
  - “Before you make your decision, please read the **explanations two other participants had given for their decisions** in similar previous tasks.”
- **Actions**
  - “Before you make your decision, please read the **decisions two other participants had made** in similar previous tasks.”

Social Information manipulation

- **Narratives**
  - "Before you make your decision, please read the **explanations two other participants had given for their decisions** in similar previous tasks."
- **Actions**
  - "Before you make your decision, please read the **decisions two other participants had made** in similar previous tasks."
- Fix beliefs before making choice in deception game (e.g., Hillenbrand and Verrina, 2022)
  - Informed about the average choice participants made in similar studies with various payments (Gerlach et al., 2019)
    - Approximately 50% truth-telling and 50% lying

Narrative selection

- In a pre-study, participants were asked: "*Please explain in a few sentences why you made this choice as you would explain to another person*."
- Use GPT-4 and recruit a separate sample (N=240) to rate persuasiveness

Narratives

Positive (persuasiveness: ~6)

1. “The honest and accurate thing to do. Give the other person the actual outcome of the dice roll. It's what I choose to do. It matches my values.”
2. “I wanted to tell the truth and this was the real outcome. I know I could have gotten more money by lying but I didn't feel good about that and this way I still get some money and the other person does too.”
3. “I chose this option as I felt it was only fair that we should both benefit from this task equally. I felt that sending a message with the wrong outcome would be devious and dishonest. I wanted us to both benefit from the task.”
4. “I made this choice as in a situation where we both have a chance to earn equal money I feel it is only fair for us both do so, as opposed to me taking more money for myself.”

Negative (persuasiveness: ~5)

1. “I will benefit either way, but there is a possibility of benefitting more. This world doesn't often allow risk-free choices to win money, and so I've taken it. So would you.”
2. “I simply maximised my own chance of gaining more. Although morally perhaps I should choose to share and indicate the true value I feel there's not too much to gain and not enough to lose for the other person for it to make a significant dent in my self-worth if I choose the selfish option.”
3. “Because I'd make more money for very little effort. The other person is still getting some, too.”
4. “I chose this option because times are hard due to the cost of living crisis and I need to maximise the chances of receiving the most amount of money for my time. Therefore I lied to the other person hoping they would choose the wrong number.”

Decision interface – treatment differences

No info interface

Your decision

Now please choose the message that you would like to send to the other participant.

The outcome from the die roll is 1	The outcome from the die roll is 2	The outcome from the die roll is 3	The outcome from the die roll is 4	The outcome from the die roll is 5	The outcome from the die roll is 6

If you need a reminder, the two possible monetary payments are:

- **Option A**: If the other participant chooses **the actual outcome** of the die roll, **i.e. 3**, then you will receive £1 and they will receive £1.
- **Option B**: If the other participant chooses **a number different than the actual outcome** of the die roll, then you will receive £3.5 and they will receive £0.1.

Narratives interface

Positive_Narrative

Before you make your decision, please read the explanations two other participants had given for their decisions in similar previous tasks.

Participant 1:*"I wanted to tell the truth and this was the real outcome. I know I could have gotten more money by lying but I didn't feel good about that and this way I still get some money and the other person does too."*

Participant 2:*"I made this choice as in a situation where we both have a chance to earn equal money I feel it is only fair for us both do so, as opposed to me taking more money for myself."*

Your decision

Now please choose the message that you would like to send to the other participant.

The outcome from the die roll is 1	The outcome from the die roll is 2	The outcome from the die roll is 3	The outcome from the die roll is 4	The outcome from the die roll is 5	The outcome from the die roll is 6

If you need a reminder, the two possible monetary payments are:

- **Option A**: If the other participant chooses **the actual outcome** of the die roll, **i.e. 3**, then you will receive £1 and they will receive £1.
- **Option B**: If the other participant chooses **a number different than the actual outcome** of the die roll, then you will receive £3.5 and they will receive £0.1.

Negative_Narrative

Before you make your decision, please read the explanations two other participants had given for their decisions in similar previous tasks.

Participant 1:*"Because I'd make more money for very little effort. The other person is still getting some, too."*

Participant 2:*"I chose this option because times are hard due to the cost of living crisis and I need to maximise the chances of receiving the most amount of money for my time. Therefore I lied to the other person hoping they would choose the wrong number."*

Your decision

Now please choose the message that you would like to send to the other participant.

The outcome from the die roll is 1	The outcome from the die roll is 2	The outcome from the die roll is 3	The outcome from the die roll is 4	The outcome from the die roll is 5	The outcome from the die roll is 6

If you need a reminder, the two possible monetary payments are:

- **Option A**: If the other participant chooses **the actual outcome** of the die roll, **i.e. 3**, then you will receive £1 and they will receive £1.
- **Option B**: If the other participant chooses **a number different than the actual outcome** of the die roll, then you will receive £3.5 and they will receive £0.9.

Actions interface

Positive_Action

Before you make your decision, please read the decisions two other participants had made in similar previous tasks.

Participant 1 *sent a message corresponding to the actual outcome of the die roll.*

Participant 2 *sent a message corresponding to the actual outcome of the die roll.*

Your decision

Now please choose the message that you would like to send to the other participant.

The outcome from the die roll is 1	The outcome from the die roll is 2	The outcome from the die roll is 3	The outcome from the die roll is 4	The outcome from the die roll is 5	The outcome from the die roll is 6

If you need a reminder, the two possible monetary payments are:

- **Option A**: If the other participant chooses **the actual outcome** of the die roll, **i.e. 3**, then you will receive £1 and they will receive £1.
- **Option B**: If the other participant chooses **a number different than the actual outcome** of the die roll, then you will receive £3.5 and they will receive £0.9.

Negative_Action

Before you make your decision, please read the decisions two other participants had made in similar previous tasks.

Participant 1 *sent a message different than the actual outcome of the die roll.*

Participant 2 *sent a message different than the actual outcome of the die roll.*

Your decision

Now please choose the message that you would like to send to the other participant.

The outcome from the die roll is 1	The outcome from the die roll is 2	The outcome from the die roll is 3	The outcome from the die roll is 4	The outcome from the die roll is 5	The outcome from the die roll is 6

If you need a reminder, the two possible monetary payments are:

- **Option A**: If the other participant chooses **the actual outcome** of the die roll, **i.e. 3**, then you will receive £1 and they will receive £1.
- **Option B**: If the other participant chooses **a number different than the actual outcome** of the die roll, then you will receive £3.5 and they will receive £0.1.

Type elicitation (A week before)

- **Prosociality**
  - SVO (Murphy et al., 2011)
  - Incentivized, 6 allocations between self and other
    - Selfish or prosocial
- **Morality**
  - Incentivized, individual die-rolling (Gneezy, 2013)
    - High or low morality concerns
- **Image**
  - Unincentivized, 11-point answer scale from 0="not at all important" to 10="very important"
    - General (Petrischeva et al., 2022)
      - "How important to you is the opinion that others hold about you?"
    - Dimension-specific
      - Prosociality: "How important to you is that others perceive you as a prosocial person (i.e., a person that intends to benefit others with behaviours such as helping, sharing, donating, cooperating or volunteering)?"
      - Honesty: "How important to you is that others perceive you as an honest person?"

Study 1: Procedures

- Prolific, December 2023, OSF pre-registered
- UK residents
- Completion fees: £0.60 (part 1) + £0.50 (part 2)
- One out of ten senders were matched with a receiver
- ~310 senders per treatment (N=3,092)

Hypotheses

- Effect of **impact**

H1: Truth-telling frequencies in HIGH_ are higher than in LOW_

Hypotheses

- Effect of **impact**

H1: Truth-telling frequencies in HIGH_ are higher than in LOW_

- Effect of **narratives**

H2: Holding impact fixed, truth-telling frequencies will be highest in the pos_narrative treatment, followed by the no_info treatment and lowest in the neg_narrative one

Hypotheses

- Effect of **impact**

H1: Truth-telling frequencies in HIGH_ are higher than in LOW_

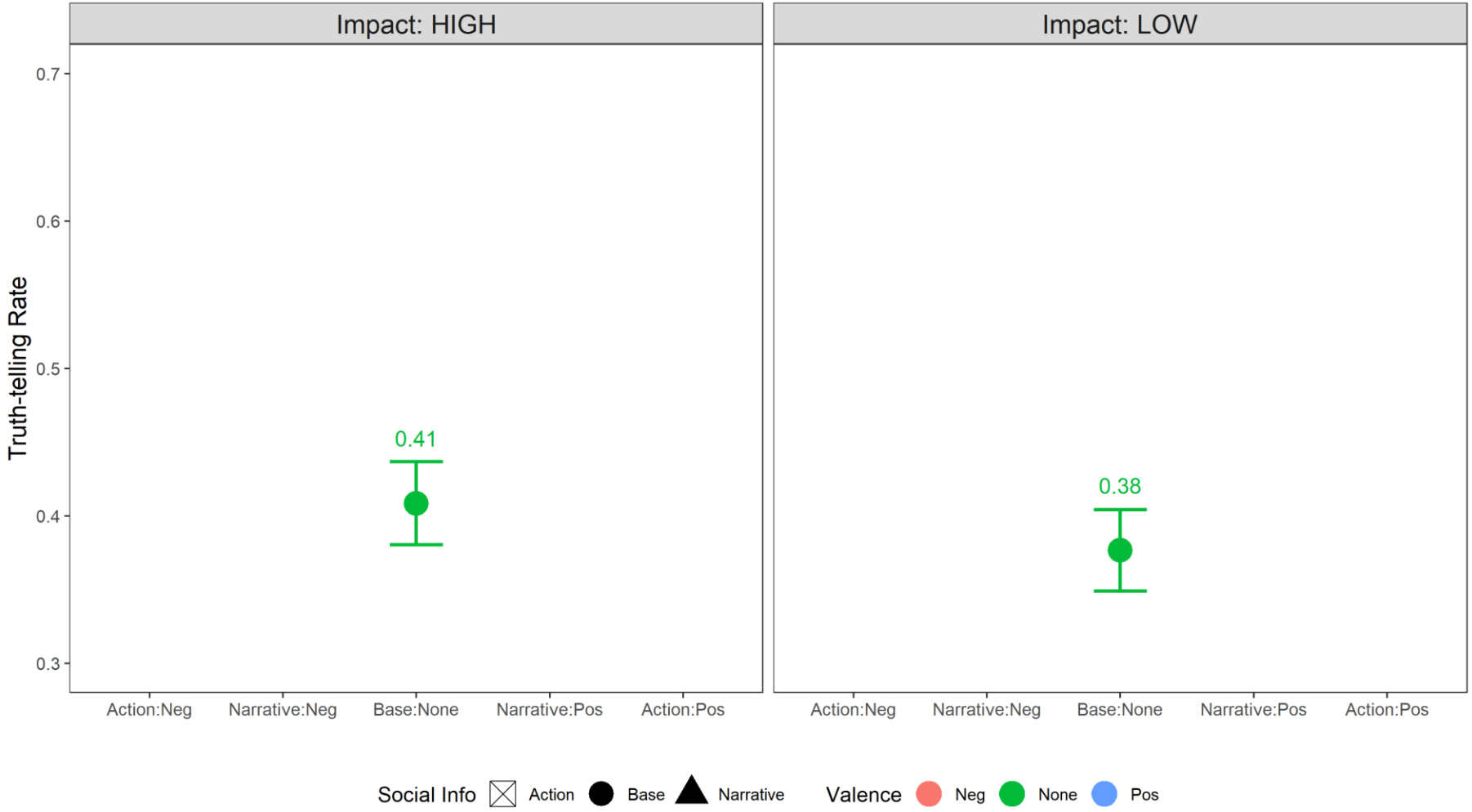
- Effect of **narratives**

H2: Holding impact fixed, truth-telling frequencies will be highest in the pos_narrative treatment, followed by the no_info treatment and lowest in the neg_narrative one

- Effect of **narratives x impact**
  - **H3:** Positive narratives increase truth-telling rates more in HIGH than in LOW
  - **H4:** Negative narratives decrease truth-telling rates more in LOW than in HIGH

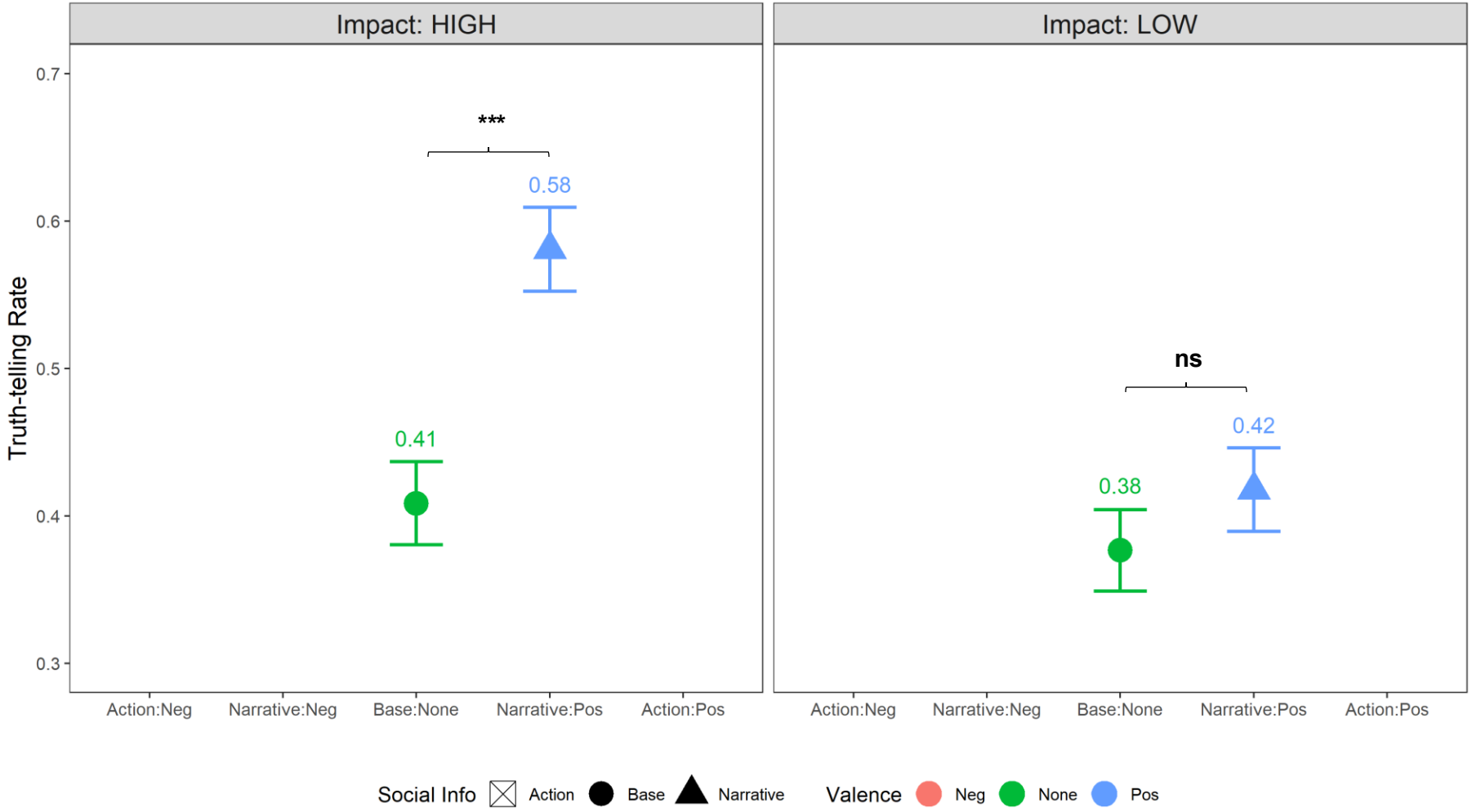
Results

Baseline levels of truth-telling



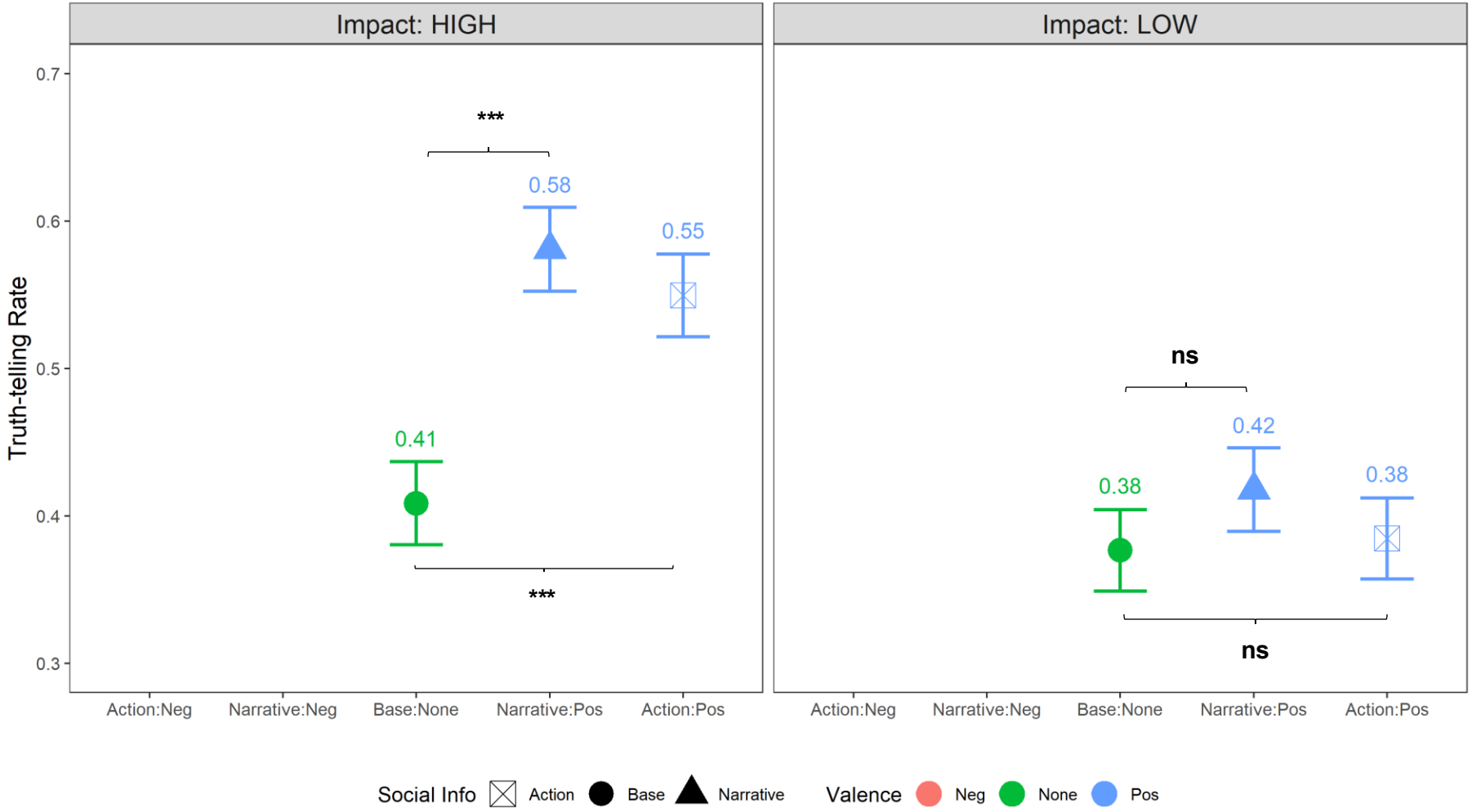
Note: p-values from Chi-squared test, *** p < 0.01, ** p < 0.05, * p < 0.10, ns p > 0.10

Positive narratives increase truth-telling only in HIGH



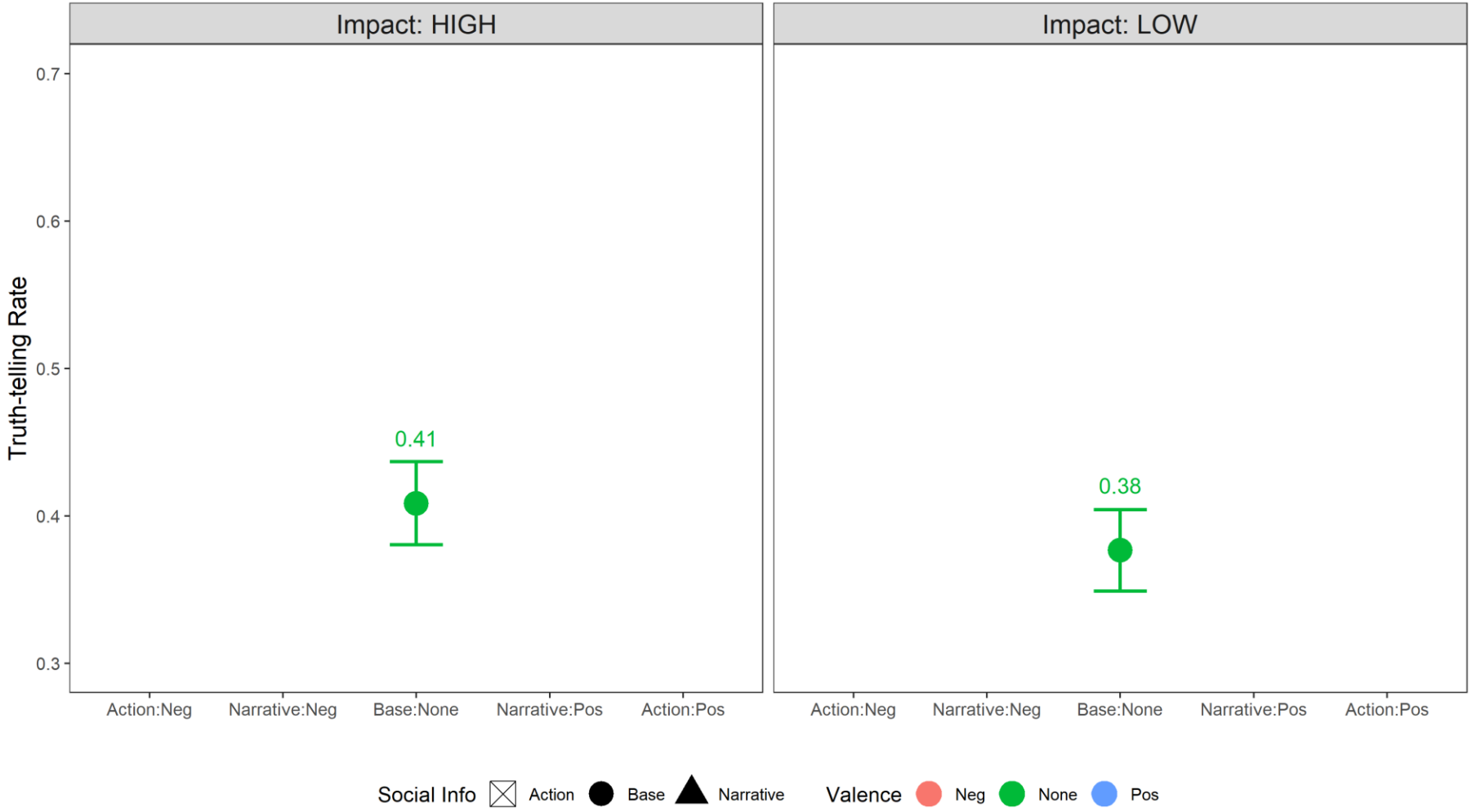
Note: p-values from Chi-squared test, *** p < 0.01, ** p < 0.05, * p < 0.10, ns p > 0.10

Positive actions increase truth-telling only in HIGH



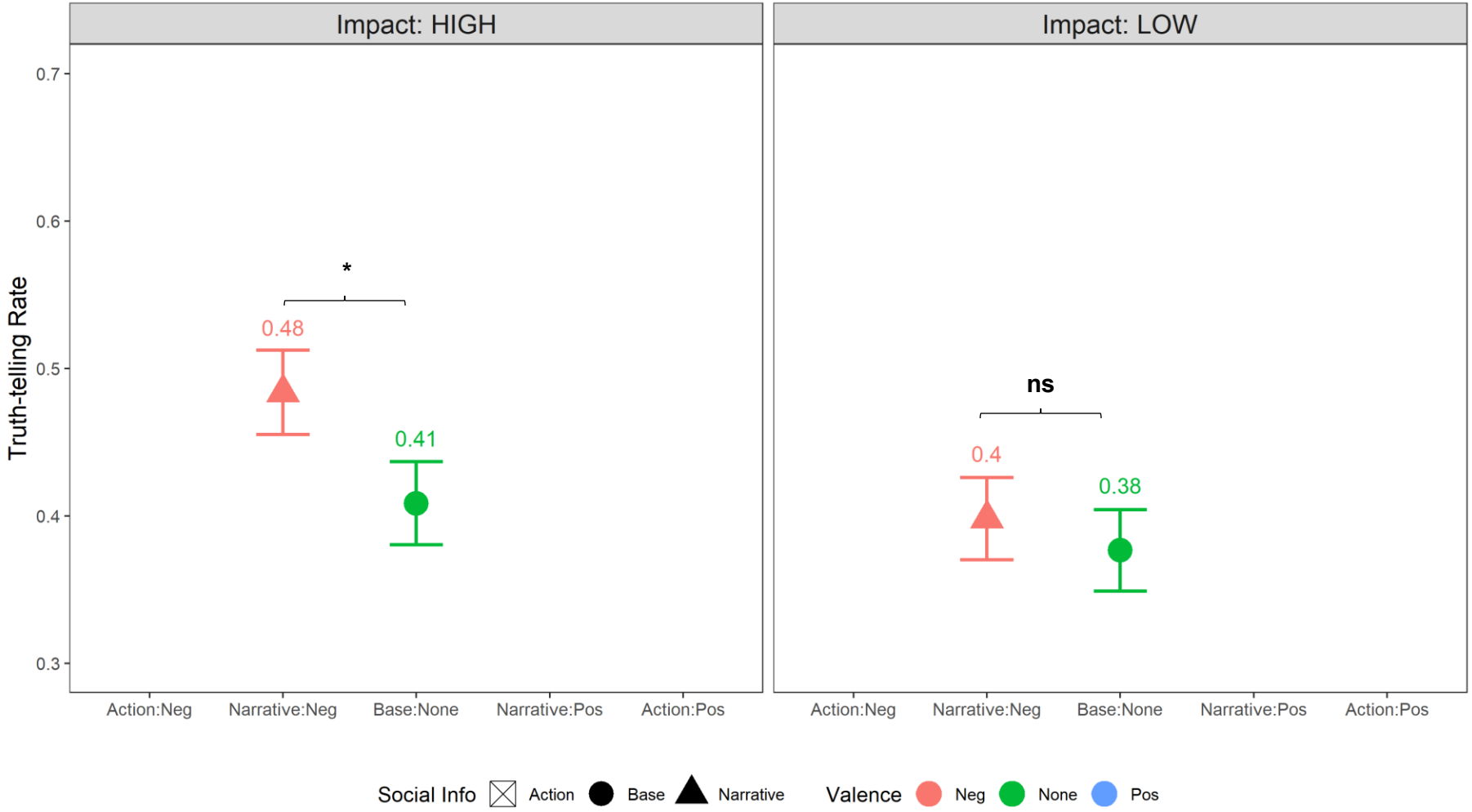
Note: p-values from Chi-squared test, *** p < 0.01, ** p < 0.05, * p < 0.10, ns p > 0.10

What about negative narratives/actions?



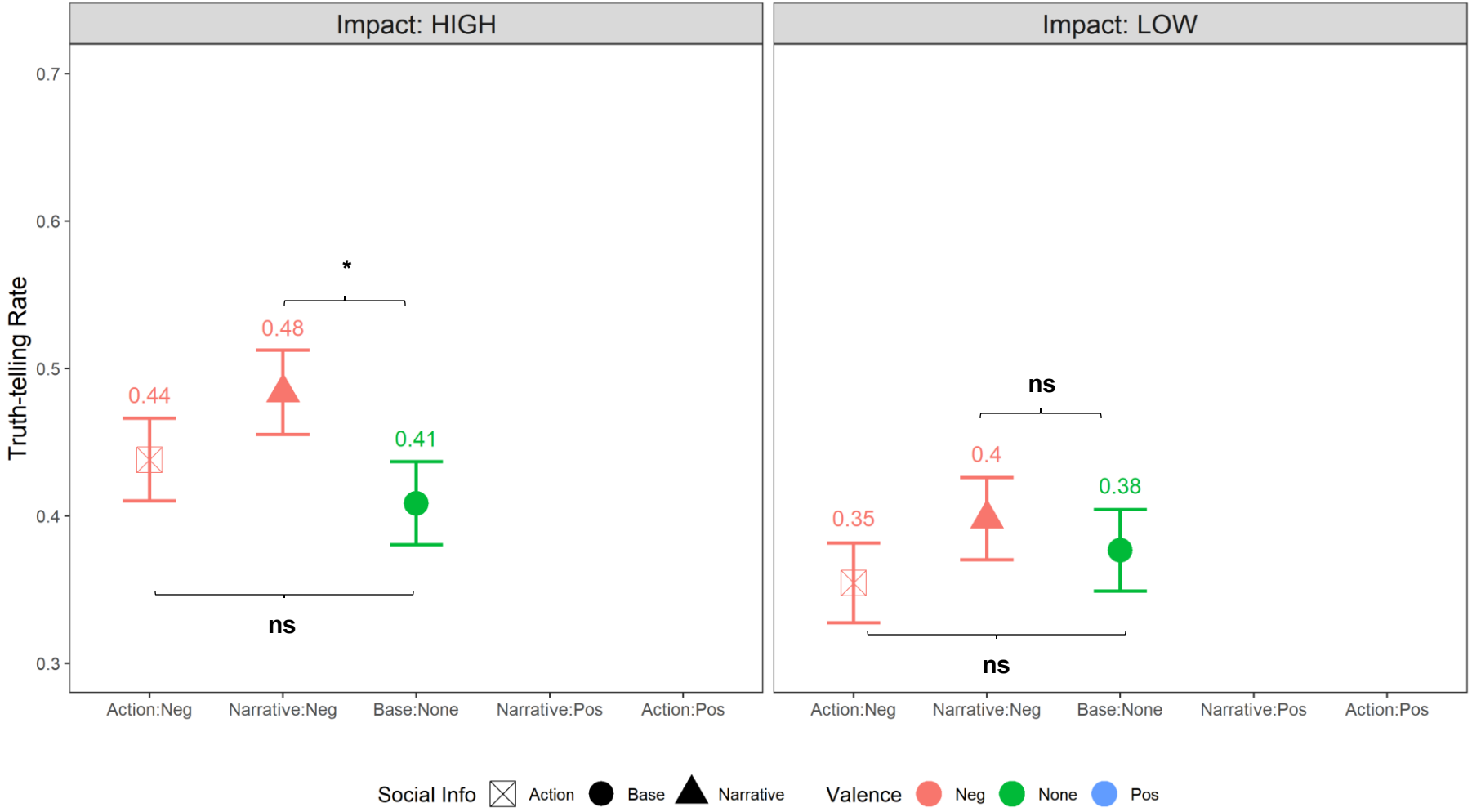
Note: p-values from Chi-squared test, *** p < 0.01, ** p < 0.05, * p < 0.10, ns p > 0.10

Negative narratives do not decrease truth-telling



Note: p-values from Chi-squared test, *** p < 0.01, ** p < 0.05, * p < 0.10, ns p > 0.10

Negative actions do not decrease truth-telling

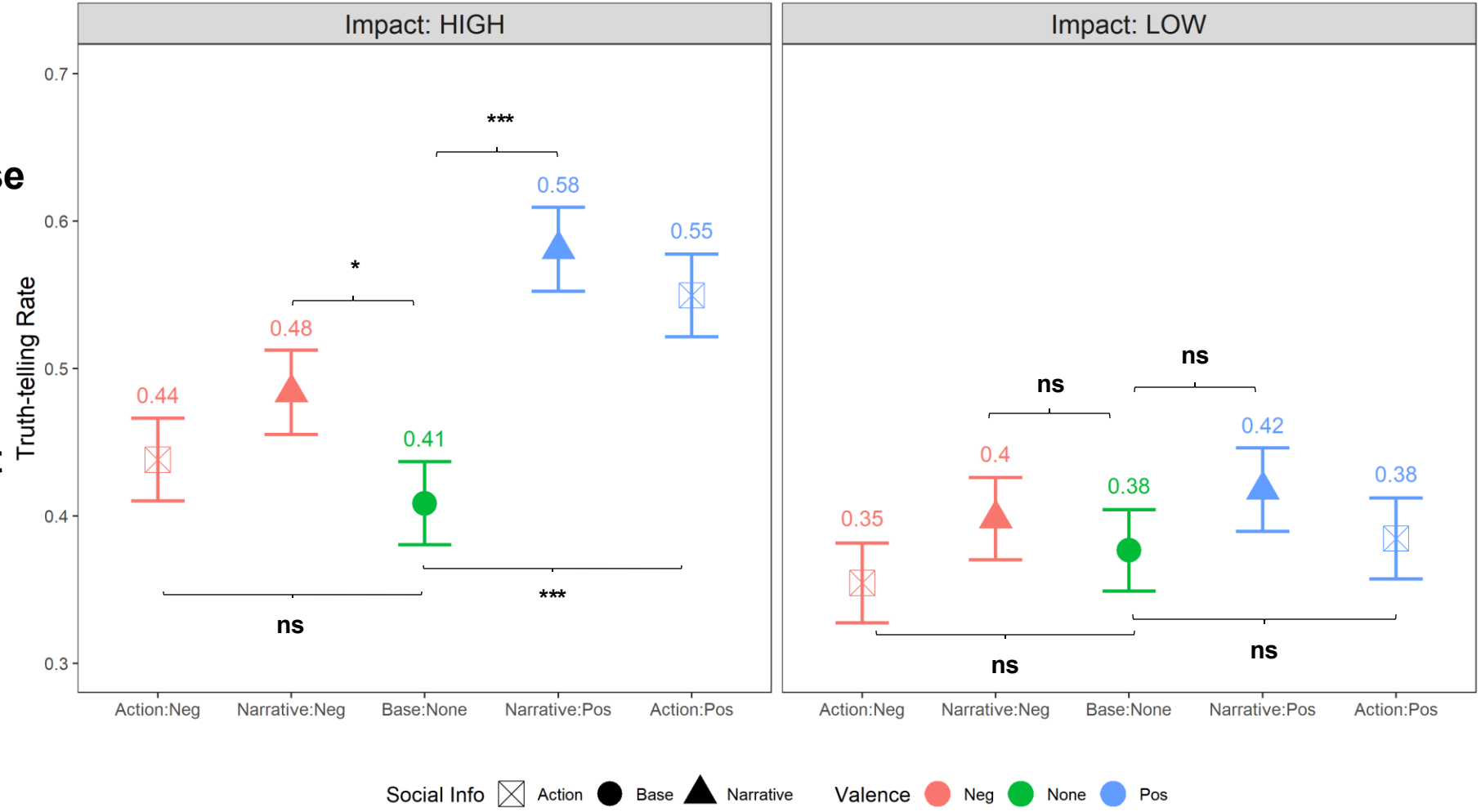


Note: p-values from Chi-squared test, *** p < 0.01, ** p < 0.05, * p < 0.10, ns p > 0.10

Sum: Differential effect of **narratives** and **impact**

1. Positive narratives **increase** truth-telling
   a) Only in HIGH
   b) Signaling of others' positive behavior

2. Negative narratives **do not decrease** truth-telling



Note: p-values from Chi-squared test, *** p < 0.01, ** p < 0.05, * p < 0.10, ns p > 0.10

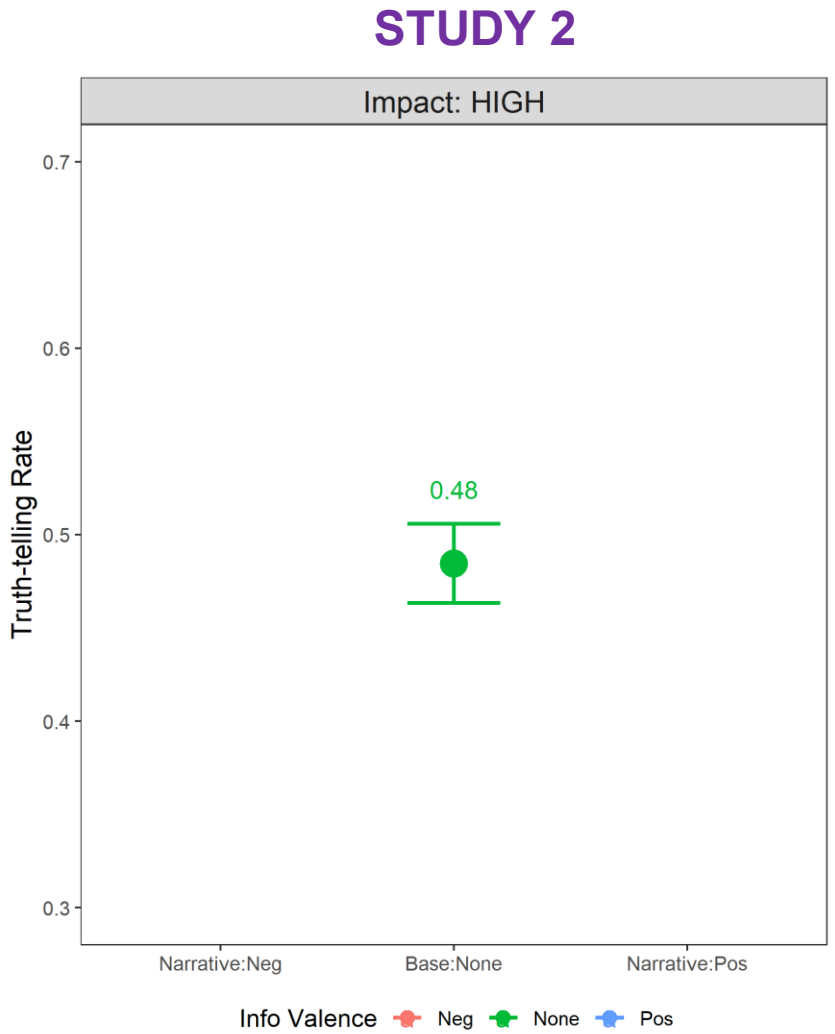
Before moving forward

- How robust are the narrative effects in the HIGH impact setting?

Study 2: Partial replication study

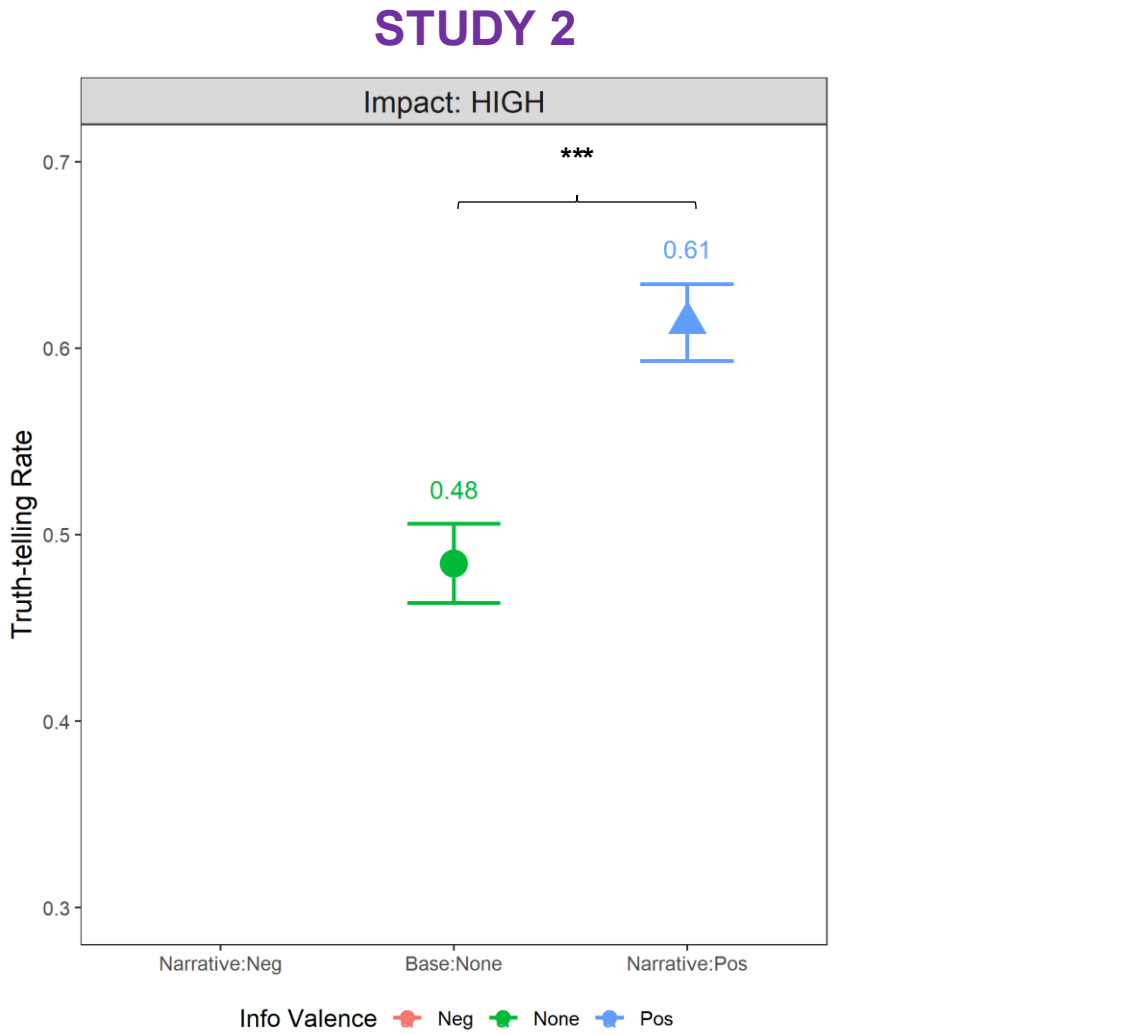
- Prolific, August 2024, OSF pre-registered
- Similar procedures as Study 1
- Highly powered: N=(1,650) - ~550 senders per treatment (before: ~310)
- 3 treatments only: HIGH-impact
  - Narratives (Positive, Negative) + Baseline
- Different from Study 1
  - Participants are forced to think about the problem for at least 15s
    - Controls for differences in how much people think about the decision problem when social information is available

Replication Results



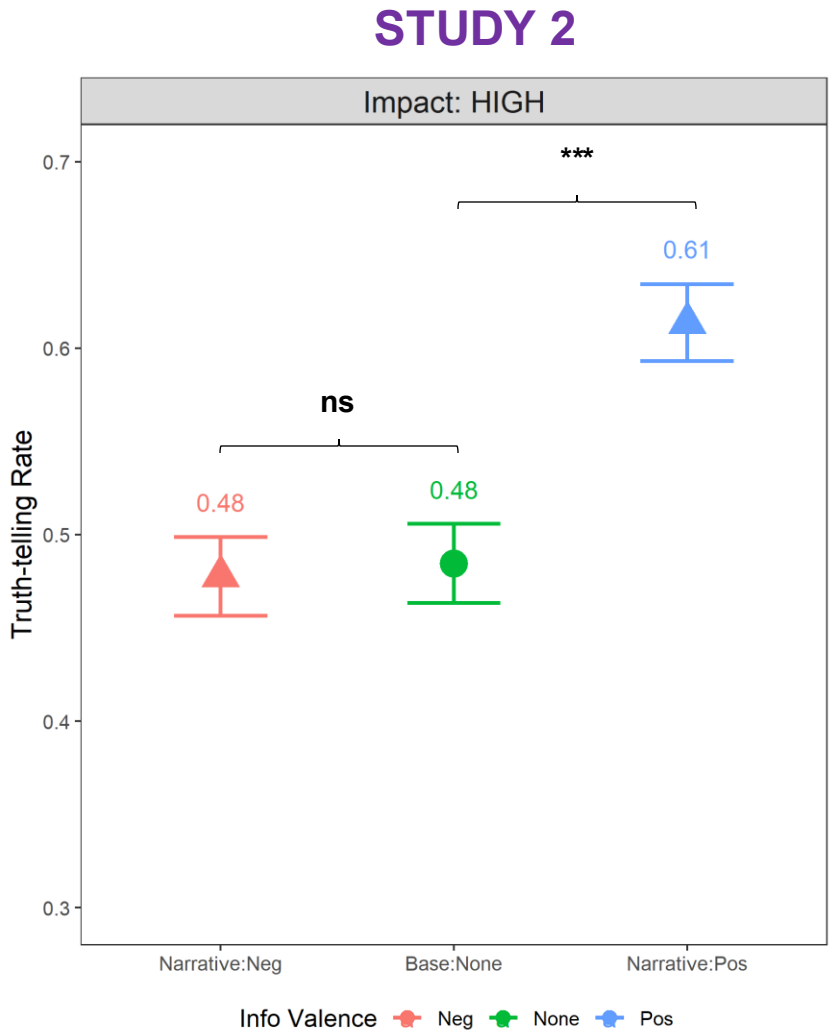
Note: p-values from Chi-squared test, *** p < 0.01, ** p < 0.05, * p < 0.10, ns p > 0.10

Positive narratives increase truth-telling



Note: p-values from Chi-squared test, *** p < 0.01, ** p < 0.05, * p < 0.10, ns p > 0.10

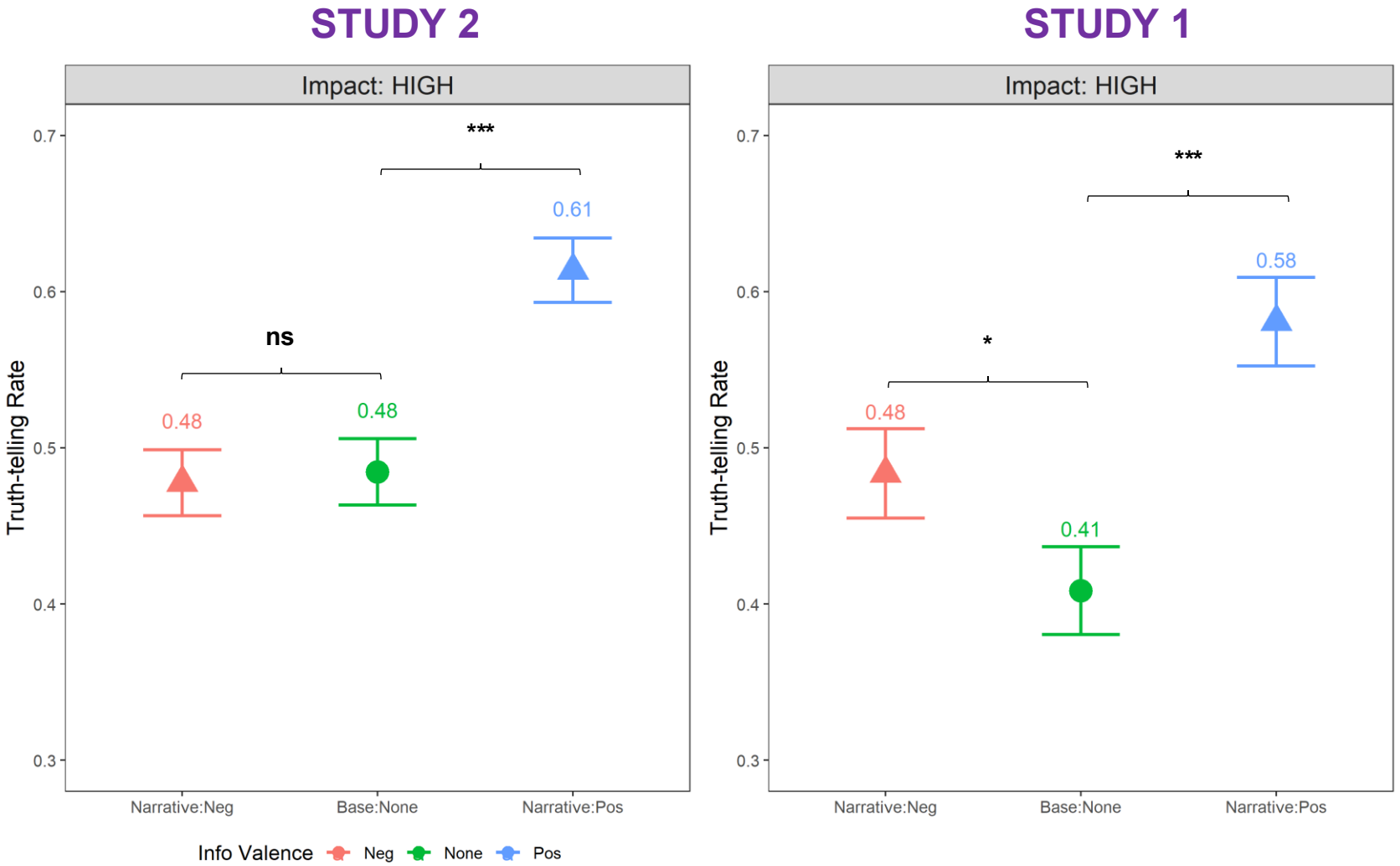
Negative narratives do not influence truth-telling



Note: p-values from Chi-squared test, *** p < 0.01, ** p < 0.05, * p < 0.10, ns p > 0.10

Combining the two studies in HIGH

- Positive narratives **increase** truth-telling
- Negative narratives **do not influence** truth-telling



Note: p-values from Chi-squared test, *** p < 0.01, ** p < 0.05, * p < 0.10, ns p > 0.10

Why are narratives differentially effective across impact levels?

- Is lying **relatively more inappropriate in HIGH** compared to LOW?
- Measure injunctive norms across impacts

Study 3: Injunctive norms across impact

- 2 treatments in each study
  - Baseline_HIGH vs Baseline_LOW
  - 500 participants per study
  - Both methods incentivized
- S3.1: Prolific, November 2024, OSF pre-registered
  - Elicit social-appropriateness via KW (Krupka and Weber, 2013)
  - “Average” normative beliefs
- S3.2: Prolific, February 2025, OSF pre-registered
  - Elicit social-appropriateness and norm multiplicity via Norm-Drawing Task (Panizza et al., 2024)
  - The same distribution of beliefs could be generated by a variety of underlying norm types
  - Beliefs about the distribution of normative beliefs
    - Draw inferences on the degree of uncertainty that underlies norm assessments

Third-party norm elicitation (Krupka-Weber, 2013)

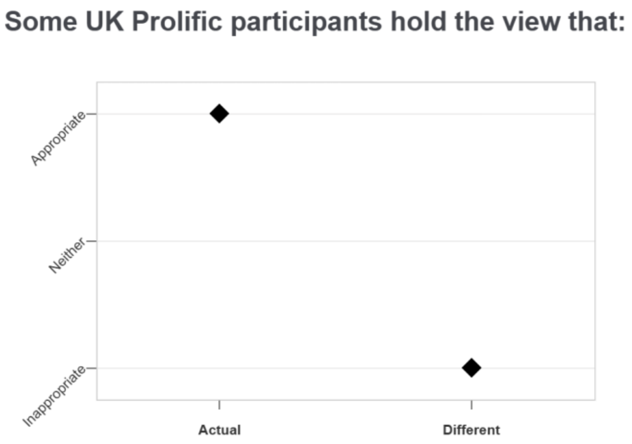
Now consider the following action:

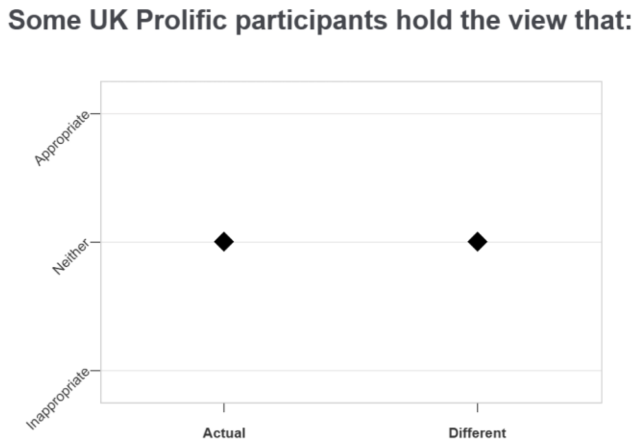
Person A sent a message corresponding to
a number different than the actual outcome of the die roll.

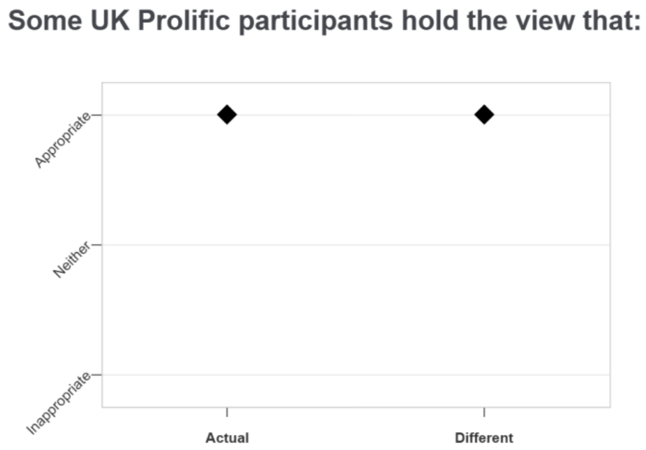
Please indicate how socially appropriate you think this action is.

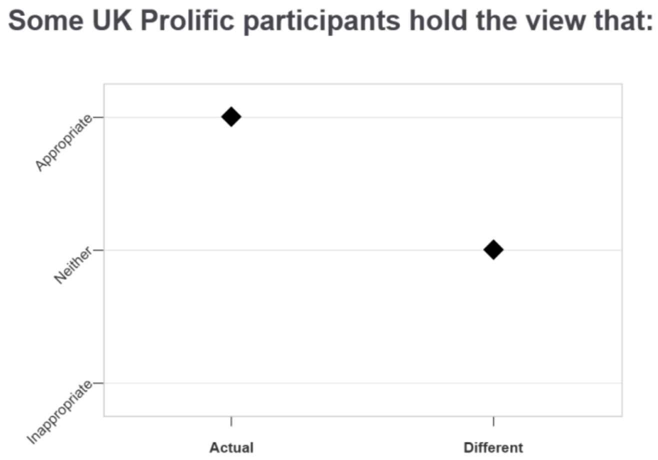
Very Socially Inappropriate	Socially Inappropriate	Somewhat Socially Inappropriate	Somewhat Socially Appropriate	Socially Appropriate	Very Socially Appropriate

Third-party norm elicitation (NDT – Panizza et al., 2024)



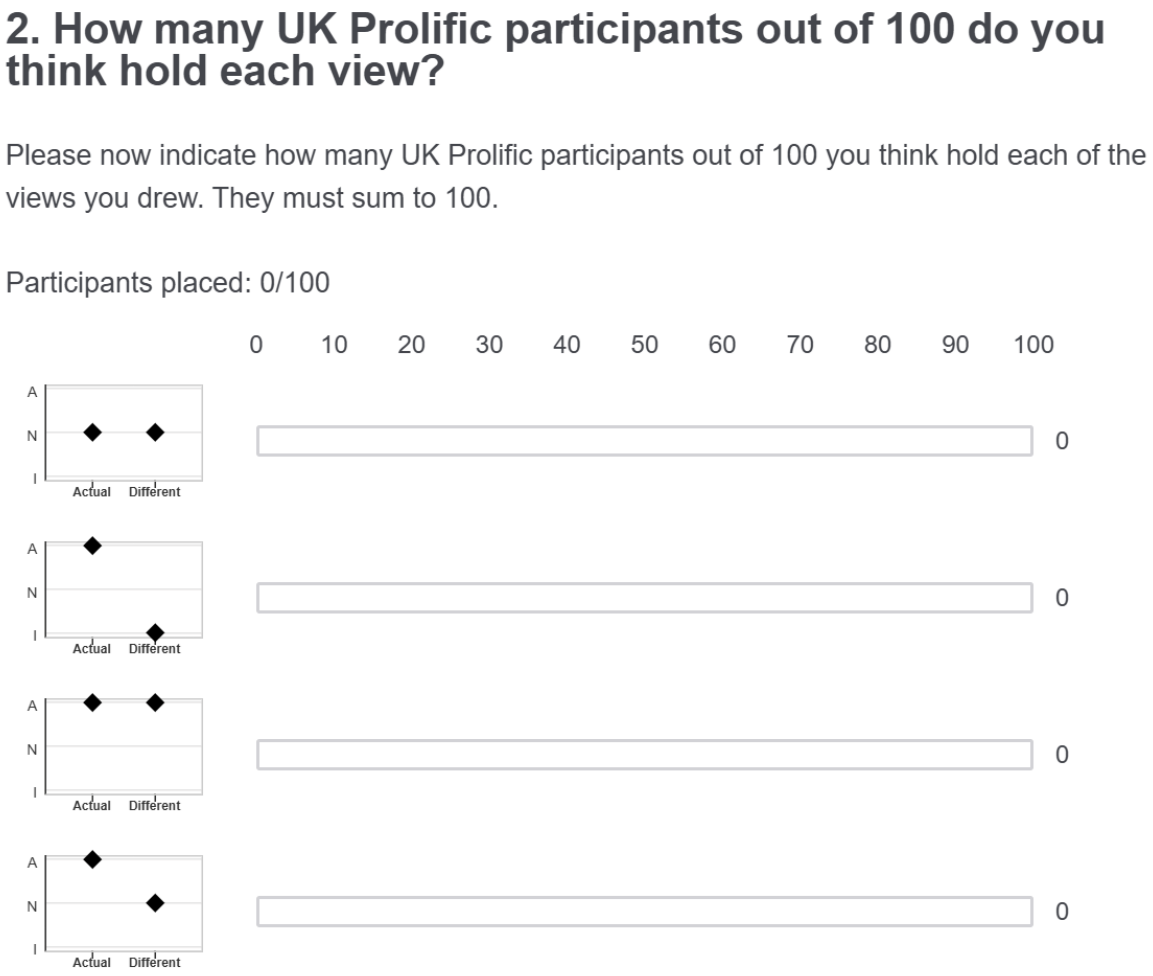






Step 1. Define the set of norms that exist

Third-party norm elicitation (NDT – Panizza et al., 2024)

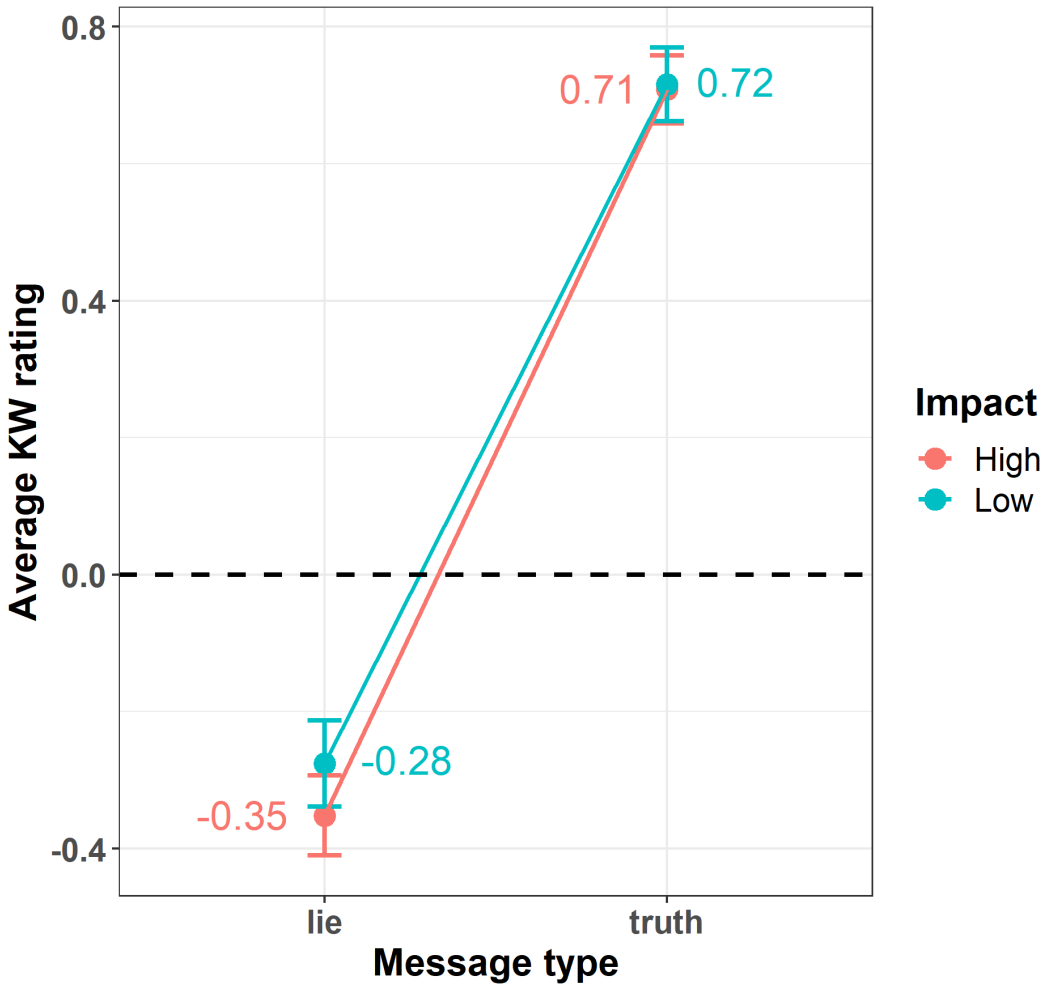


Step 2. Guess how these norms are distributed in the population

3.1 Krupka-Weber Results

- -1: Very Socially Inappropriate
- +1: Very Socially Appropriate

- Truth-telling equally Appropriate
- Lying relatively more Inappropriate in HIGH



3.2 Norm Drawing Task Results

- Stronger evidence that lying is more inappropriate in HIGH than in LOW
  - **Both actions are neither Appropriate nor Inappropriate**
    - Report View: fewer in HIGH (60%) than in LOW (73%) (Wilcoxon 2-sided test, $p = 0.01$)
  - **Truth is Appropriate and Lying is Inappropriate**
    - Distribution: in HIGH - 43%, compared to LOW - 37% (Wilcoxon 2-sided test, $p = 0.04$)
- Truth-telling norms stricter in HIGH, higher norm ambiguity in LOW

Narratives and social norms as complements

- Why do positive narratives work better in HIGH impact contexts?
  - Stronger normative consensus against lying
    - Less normative ambiguity to overcome
  - Narratives align with and reinforce existing norms
- Why do negative narratives have limited effectiveness?
  - Less persuasive in general, possibly due to competing with existing moral norms against lying
  - May actually trigger moral reflection, sometimes even increasing honesty

Conclusion

- Narratives shape honesty most when they reinforce social norms
- Positive narratives increase honesty by up to 17pp when harm is high and norms are clear
- Negative narratives rarely change truth-telling
- Effectiveness depends on harm level and normative consensus
  - Policy Design: target high-stakes via narratives for efficiency
  - Through action signaling
    - No intervention is inherently more valid than the other, **even if** they work through the same mechanism as direct action information
    - People often observe justifications rather than actions
    - Narratives more memorable (Graeber et al., 2024) and as such particularly valuable for scaling up interventions
- **Take-away: Understanding prevailing social norms is crucial for effective narrative persuasion**

Thank you!

despoina.alempaki@warwick.ac.uk

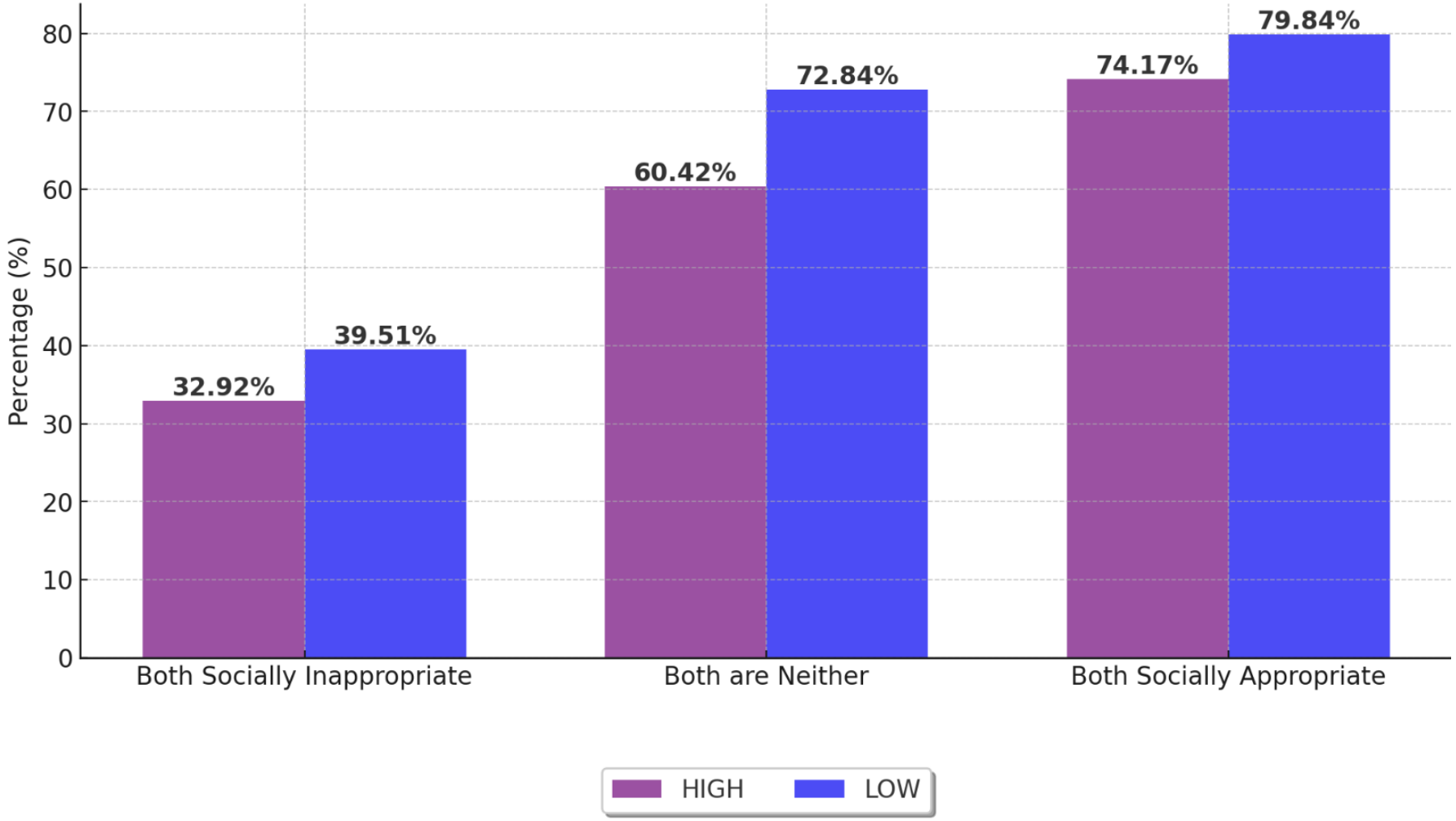
Narrative selection (1)

- In a pre-study, participants were asked: "*Please explain in a few sentences why you made this choice as you would explain to another person*."
  - Categorised as positive if justify telling the truth and negative if justify lying
- Used GPT-4 to:
  - Group them into content-categories: if it talked about 'honesty', 'fairness', 'honesty and fairness', 'self-payoff maximization' or 'other'
  - Rate their persuasiveness (scale: 1-7)
- Selected 45 positive and 45 negative narratives that varied in content and perceived persuasiveness

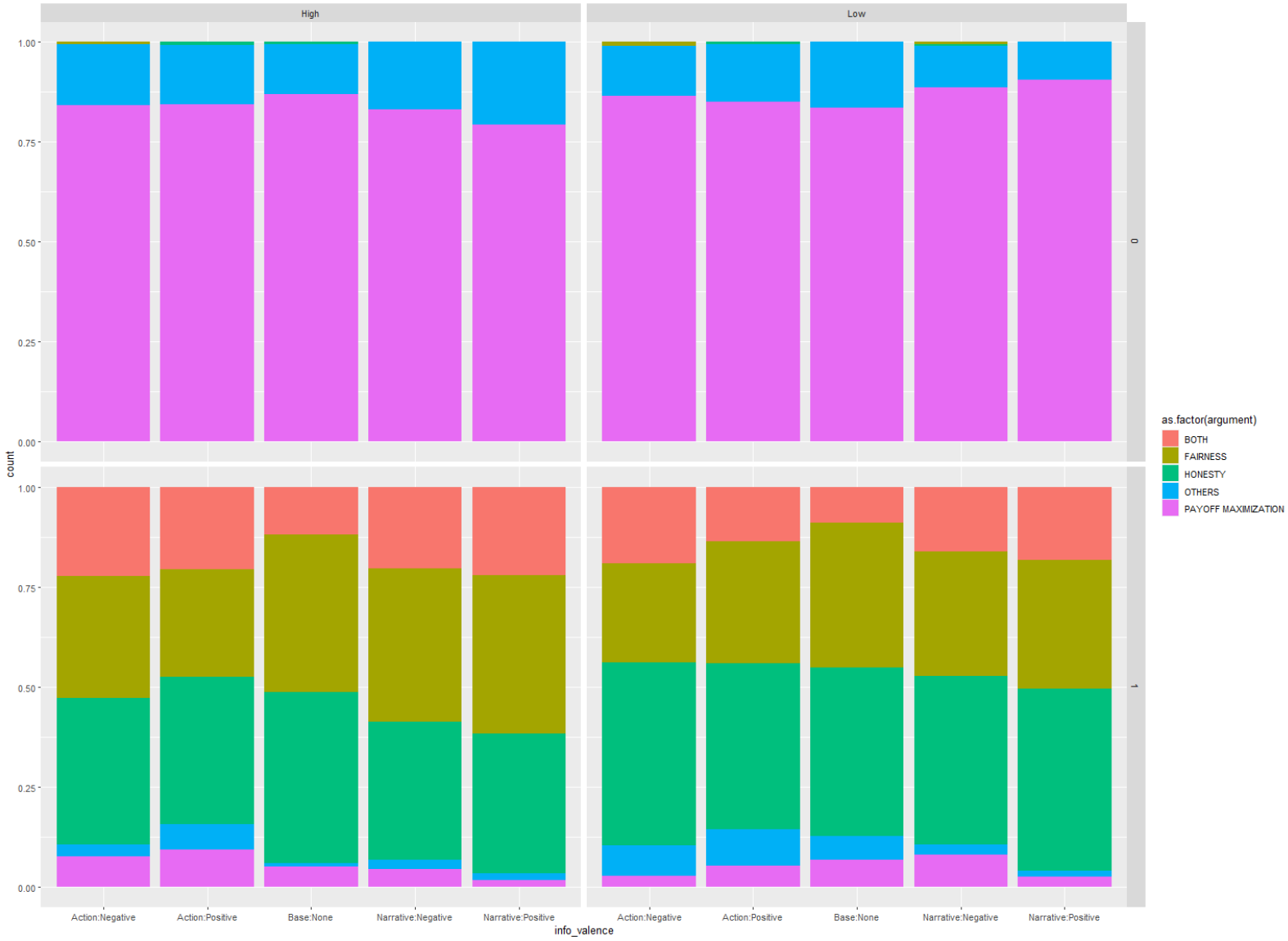
Narrative selection (2)

- Recruited a separate Prolific sample (N=240) to rate their persuasiveness
  - *"We ask you to imagine that each explanation will be presented to another participant in a similar decision situation, and consider the extent to which it will succeed in persuading the other participant to make the same choice, i.e., sending a message {corresponding to the actual outcome of the die roll [positive]/different than the actual outcome of the die roll [negative]}. Please rate how persuasive you think each explanation is from "Not at all persuasive" (1) to "Very persuasive" (7)."*
- Presented with the description of the deception game (HIGH or LOW) and asked to rate 30 out of 45 randomly selected narratives
  - Same narrative (positive or negative) was rated both under HIGH/LOW
  - Raters saw only either positive or negative narratives
- Selected the 4 **most persuasive positive** and the 4 **most persuasive negative** narratives elicited under HIGH and LOW
  - 2 from HIGH and 2 from LOW
  - Average persuasiveness did not differ across HIGH and LOW

NDT

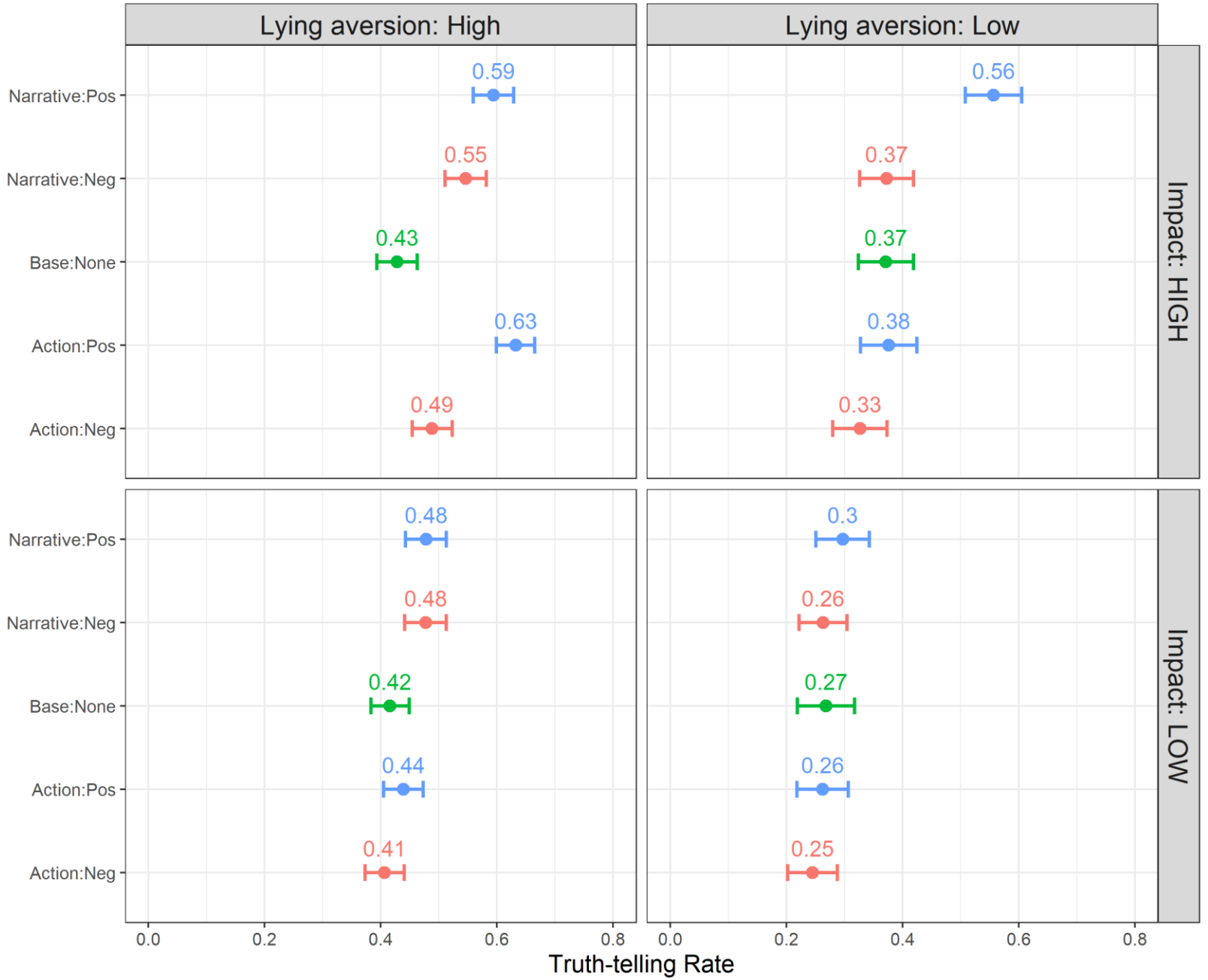

Percentage (%)
0
10
20
30
40
50
60
70
80
32.92%
39.51%
60.42%
72.84%
74.17%
79.84%
Both Socially Inappropriate
Both are Neither
Both Socially Appropriate
HIGH
LOW

Narratives content



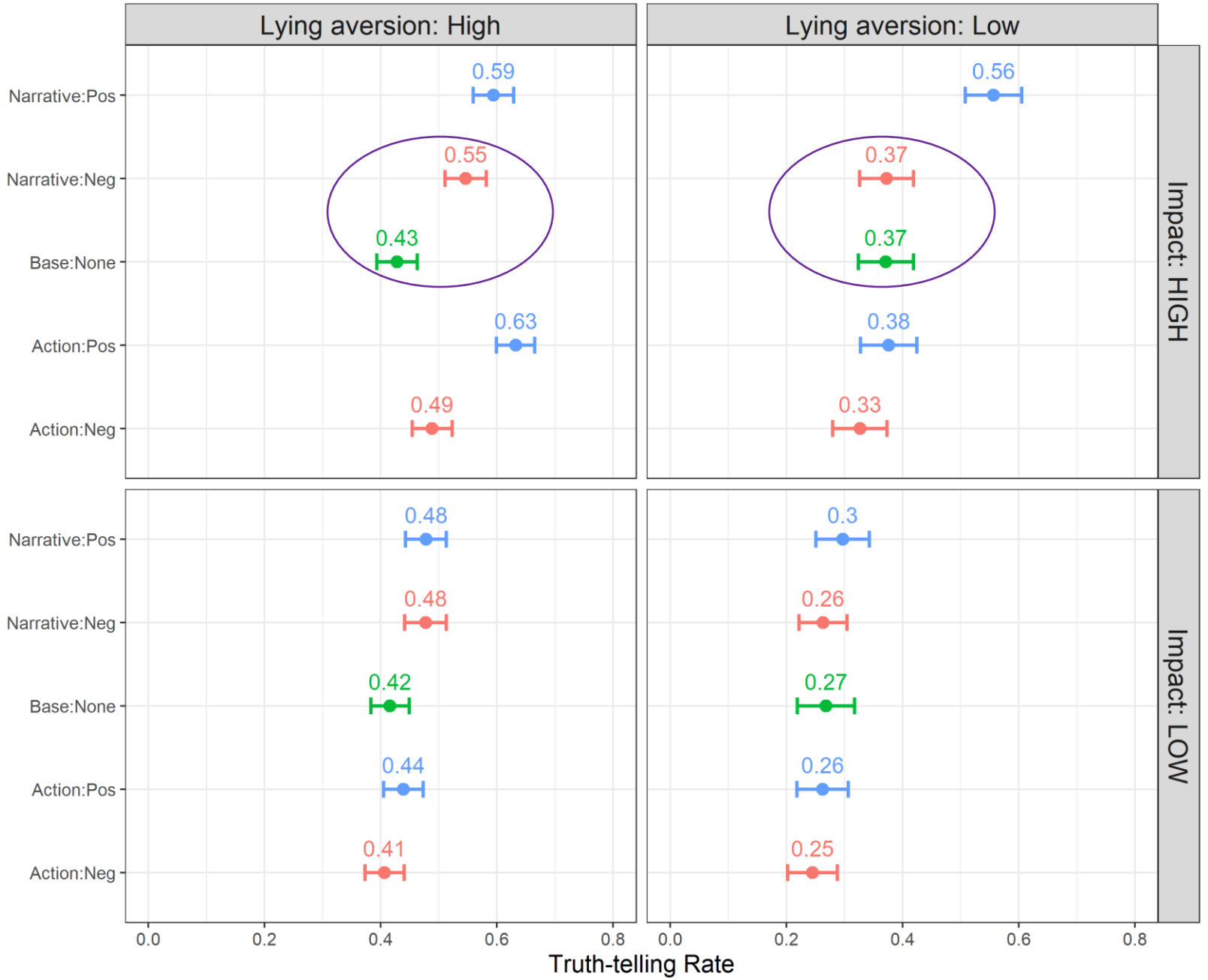
Study 1 by Type

By Lying type



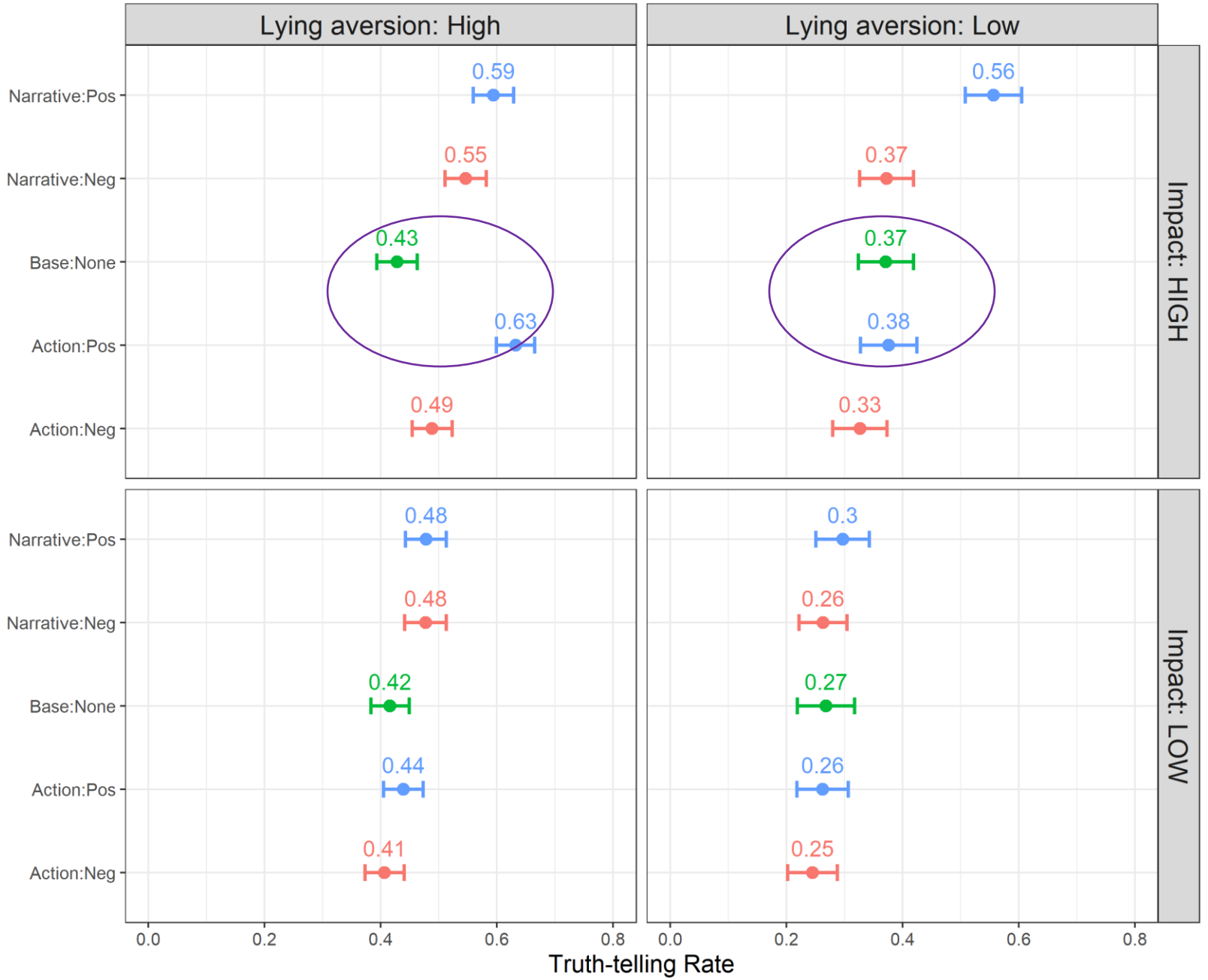
By Lying type

- Positive effect of negative narratives is driven mostly by types with high lying aversion



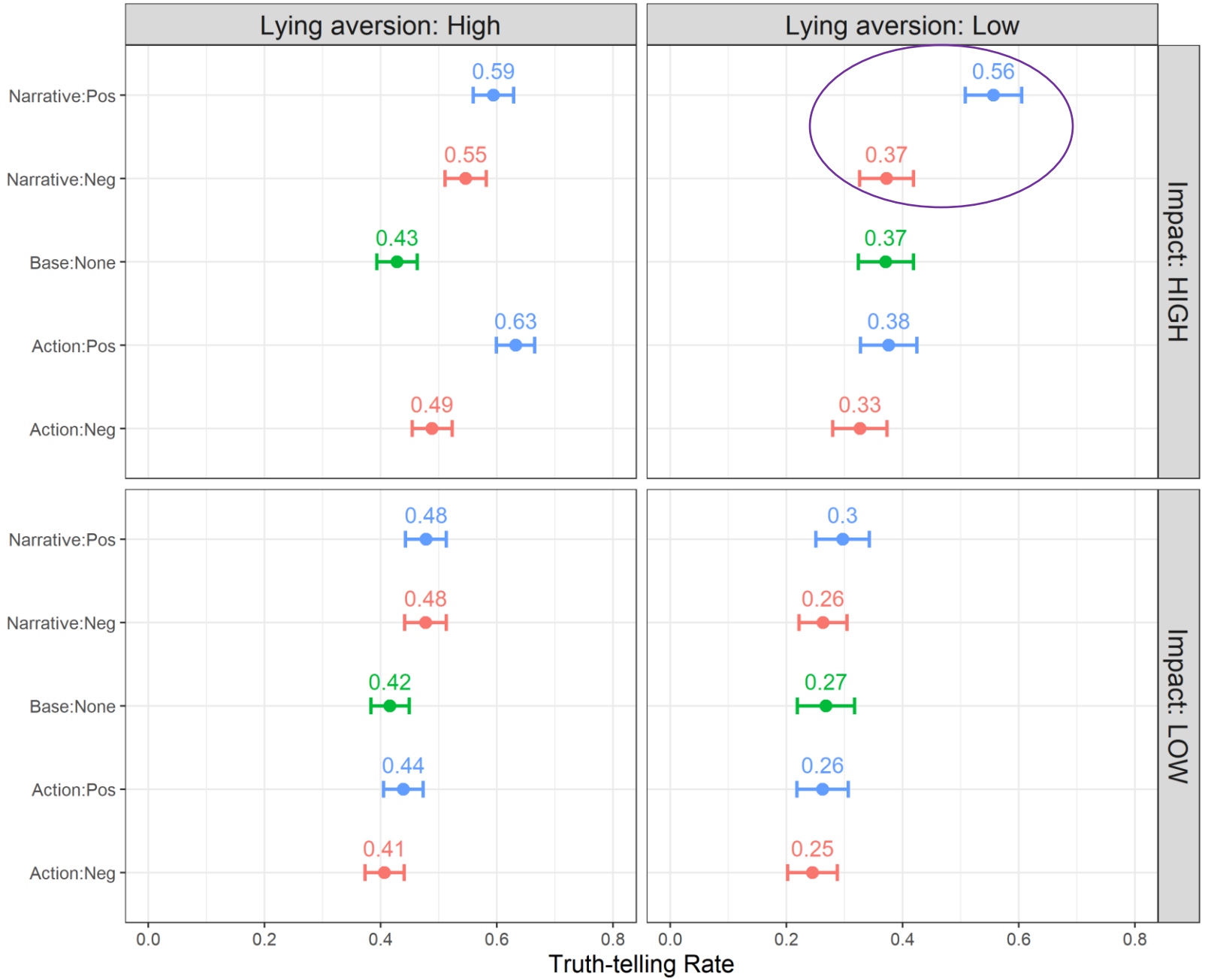
By Lying type

- Positive effect of positive actions is driven mostly by types with high lying aversion



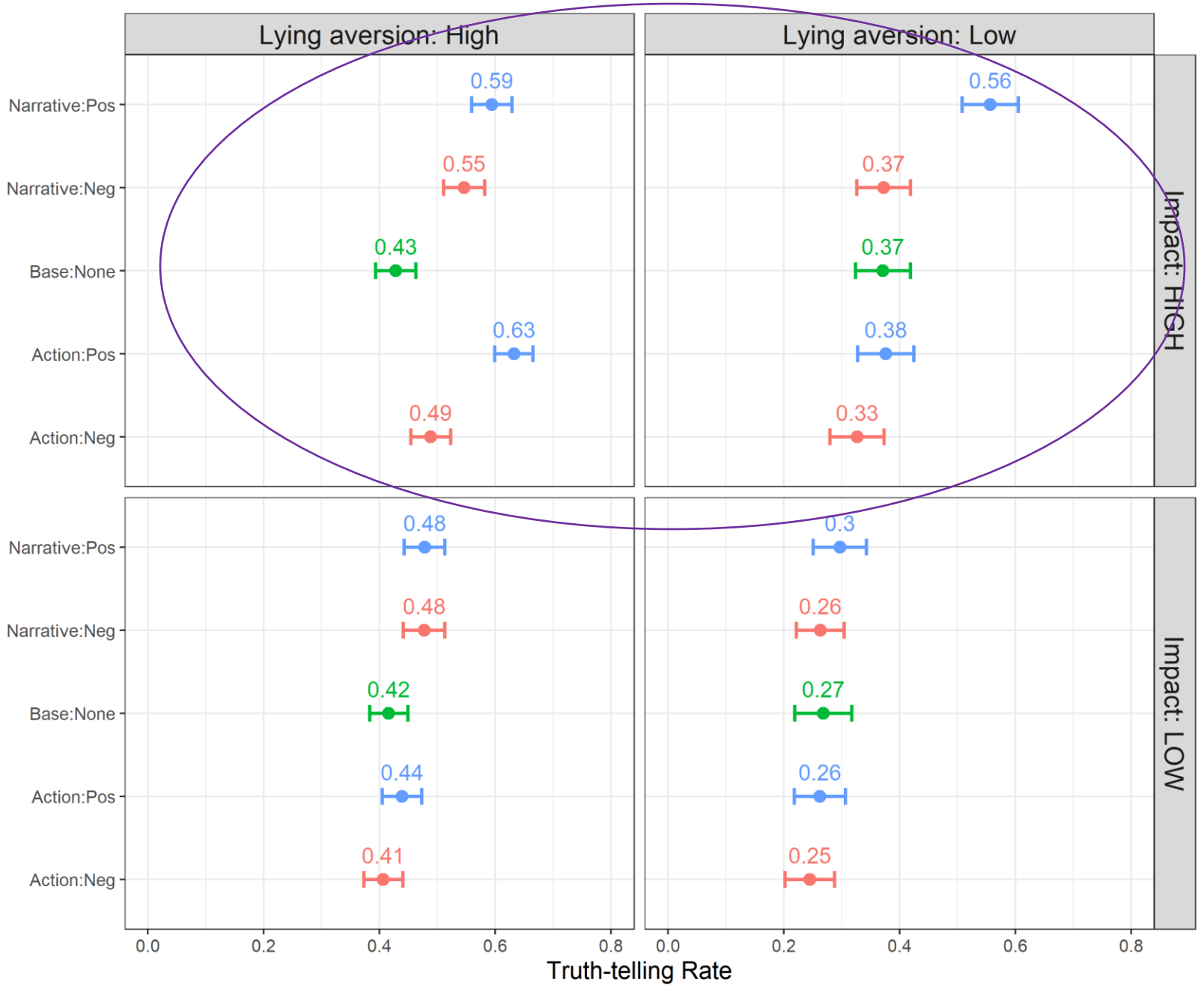
By Lying type

- Types with low lying aversion are positively affected by positive narratives only



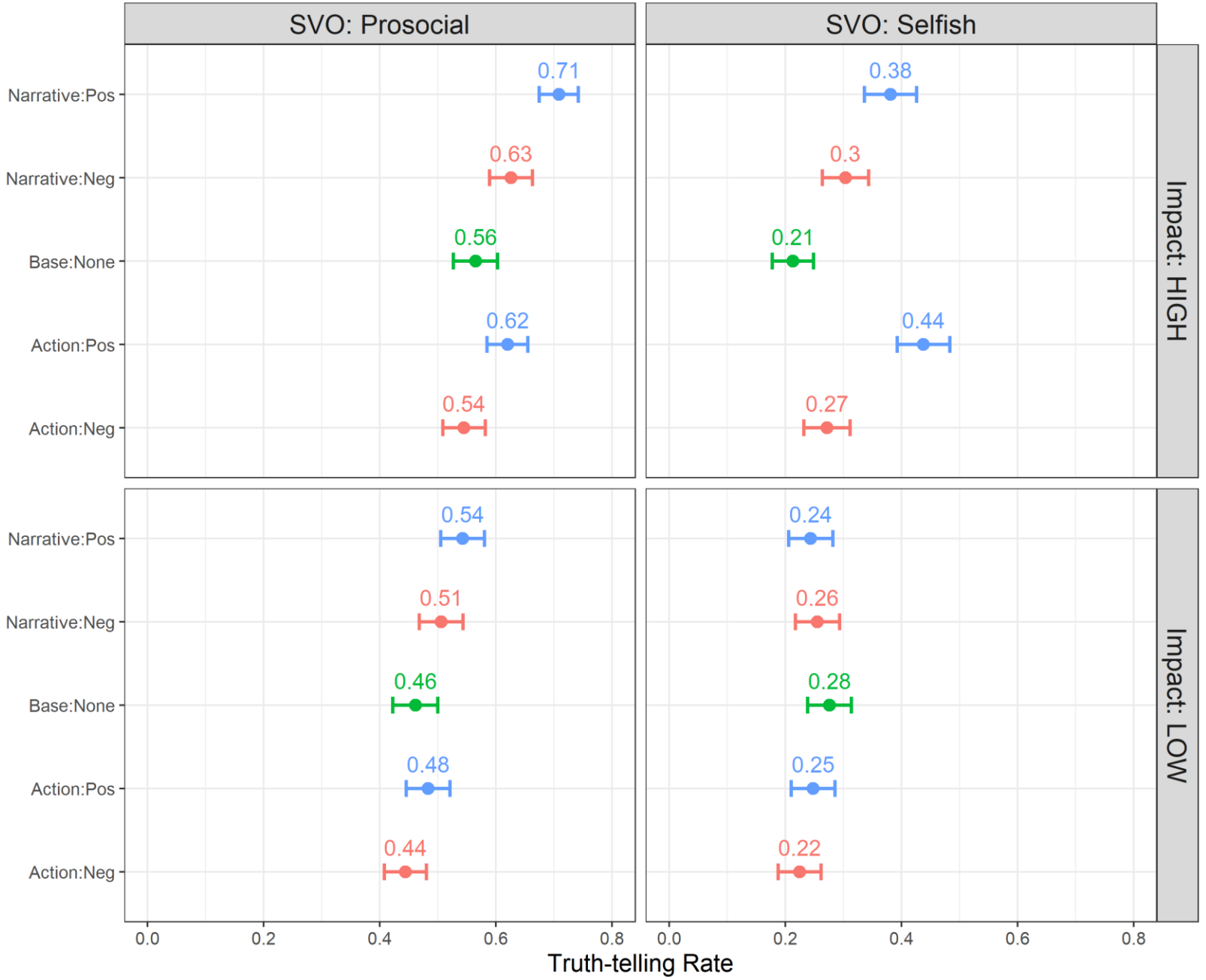
By Lying type

- Both types are affected by (some) social information only when impact is HIGH



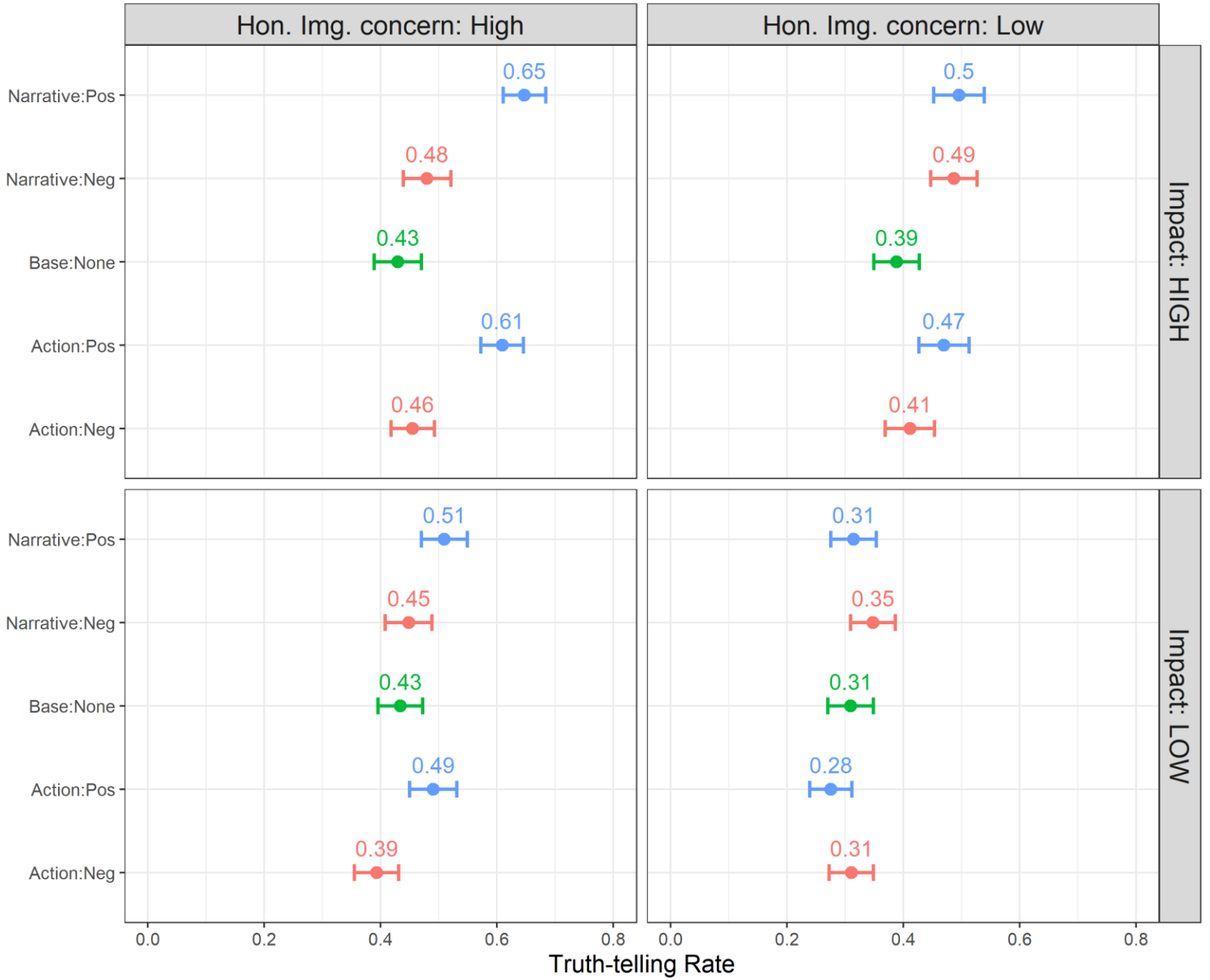
By SVO type

- Prosocial types react more to positive narratives than positive actions
- Selfish types react equally to positive narratives and positive actions
- Only when impact is HIGH

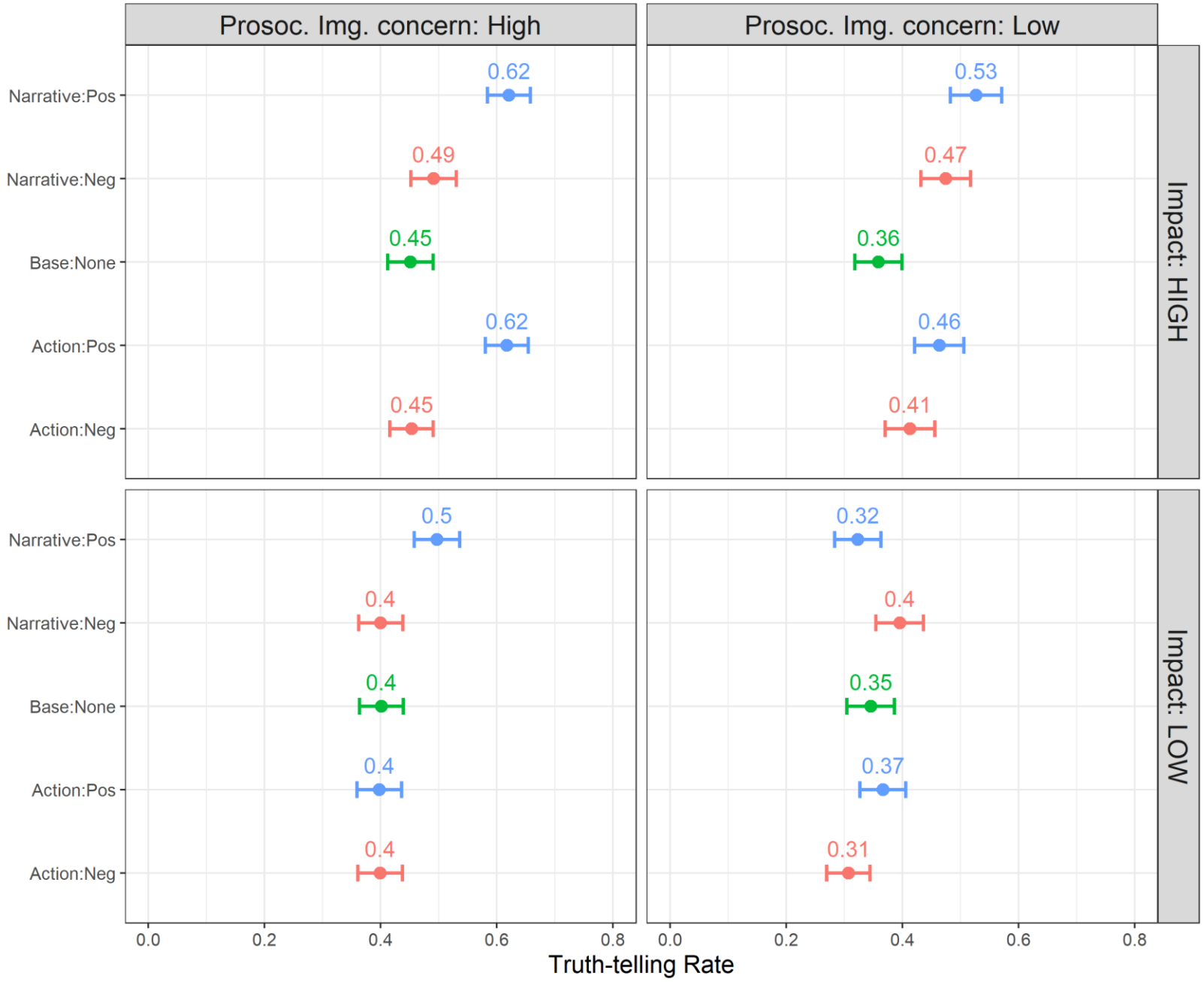


By Honesty Image Type

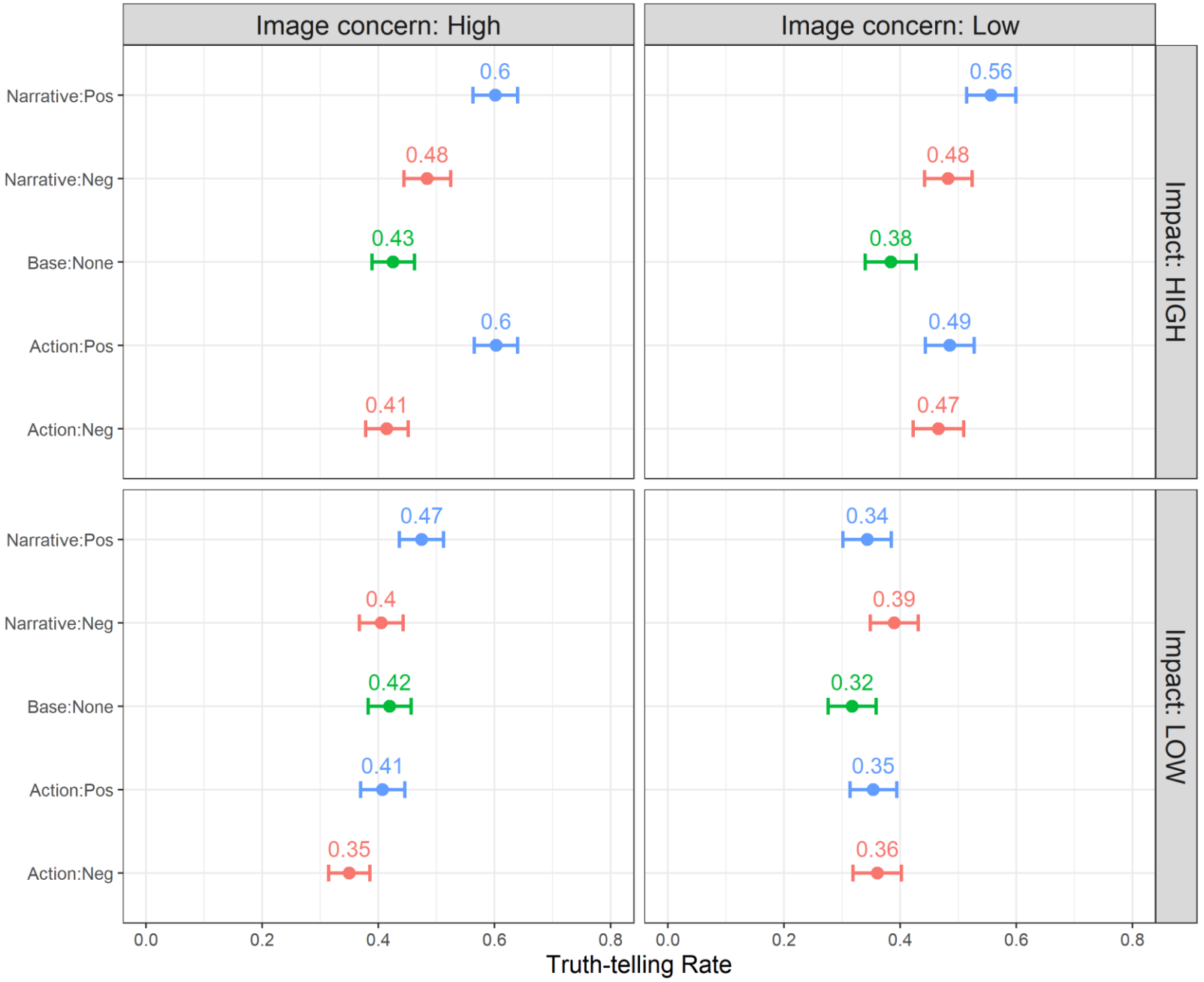
- Types with high image concerns are equally affected by positive narratives and actions (not negative ones)
- Types with low image concerns react more to narratives - both positive and negative



By Prosocial Image Type



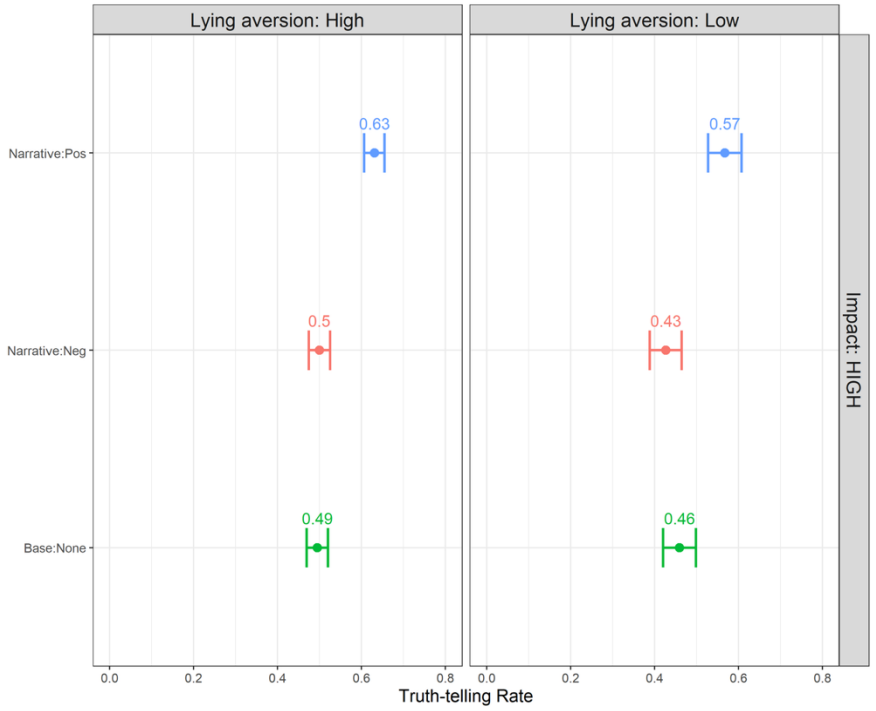
By General Image Type



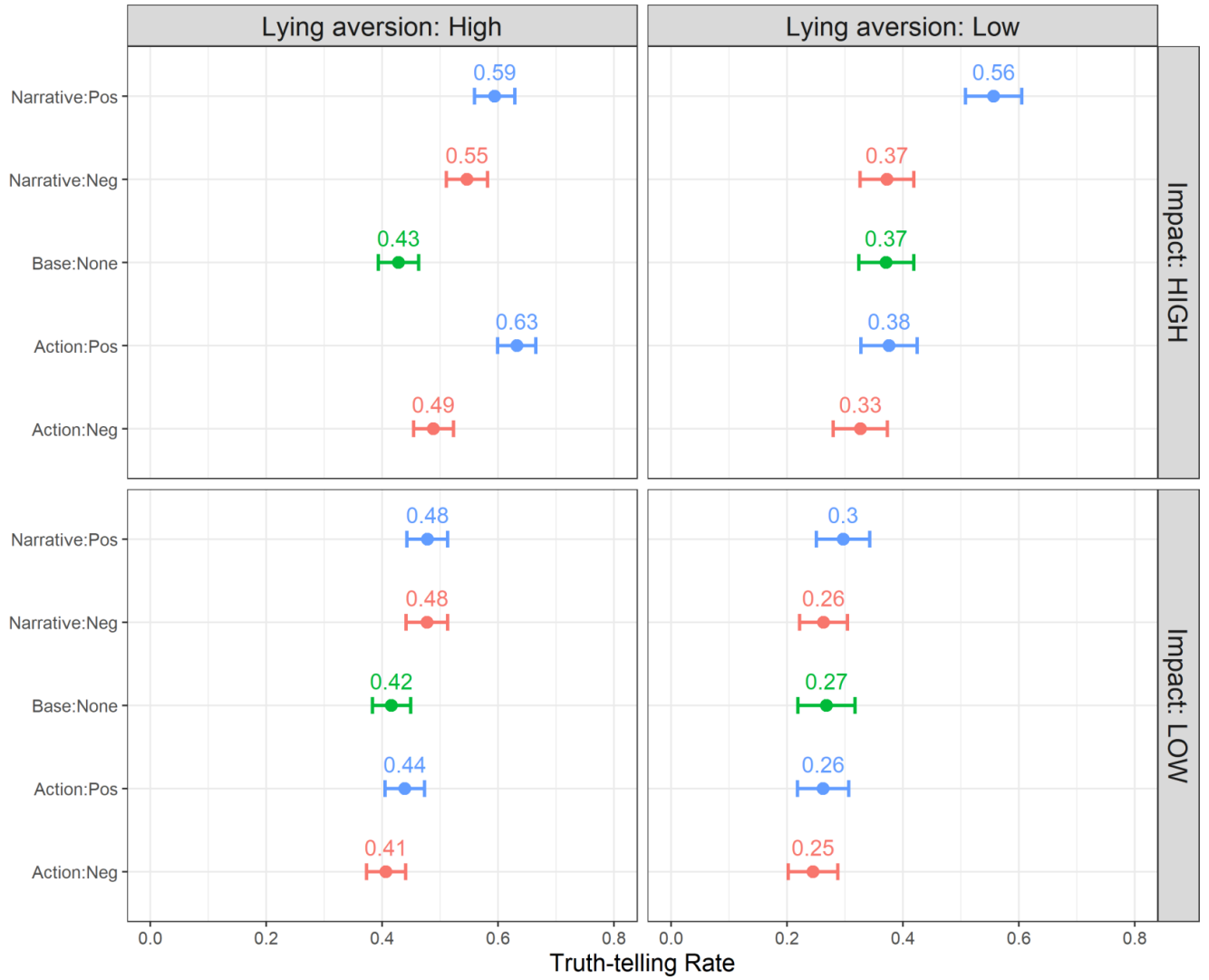
Study 1 and 2 by Type

By Lying type

STUDY 2

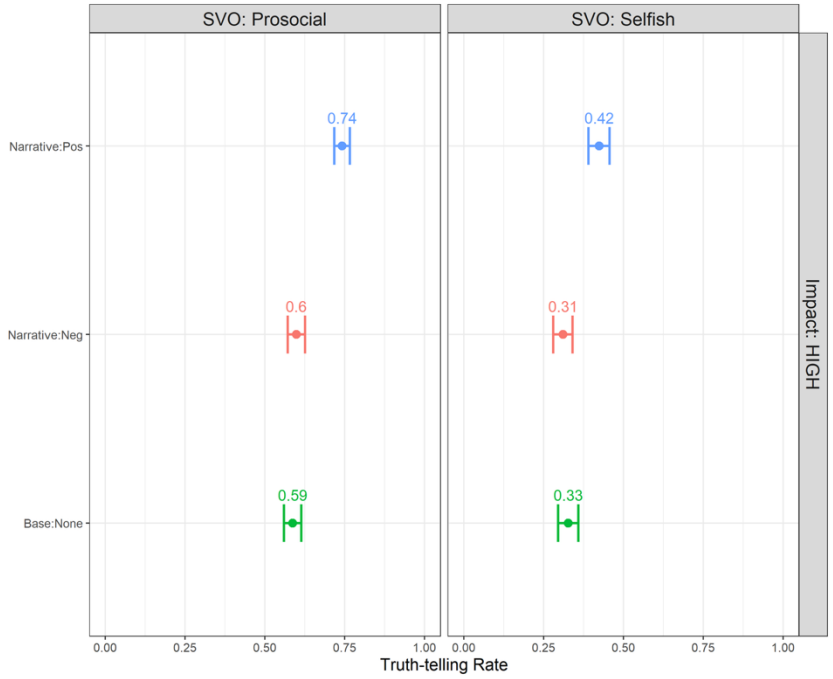


STUDY 1

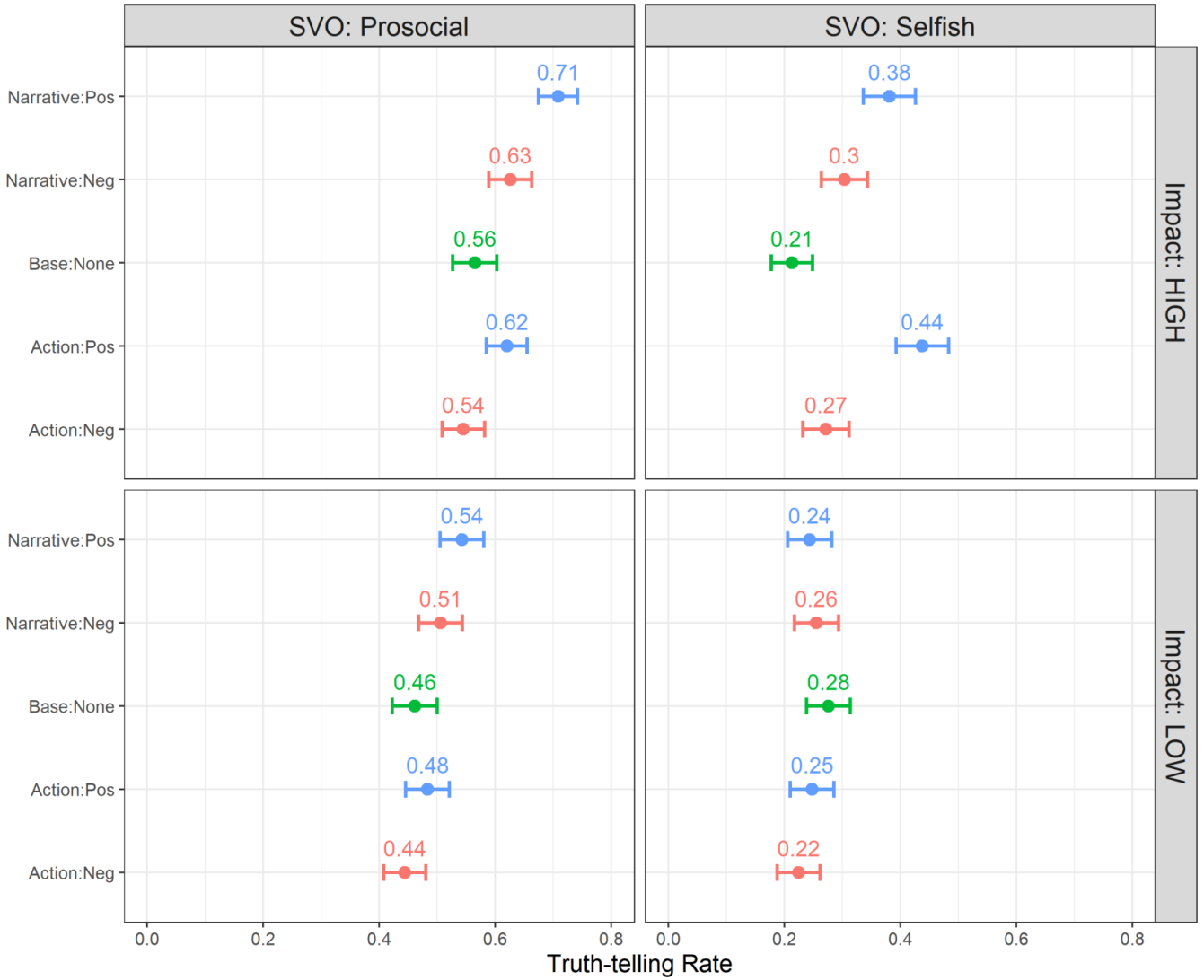


By SVO type

STUDY 2

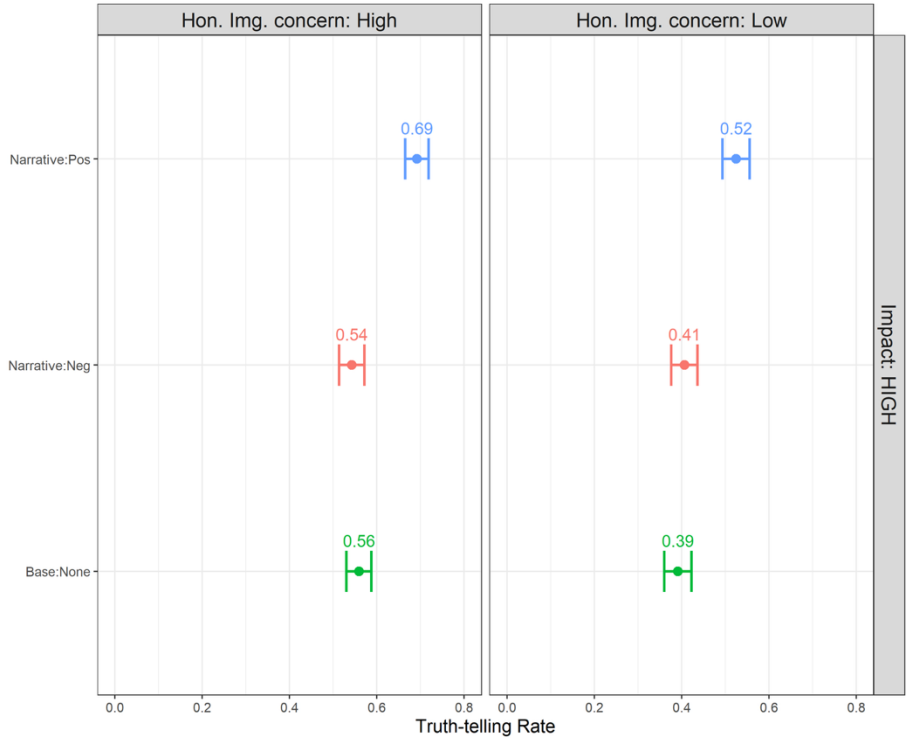


STUDY 1

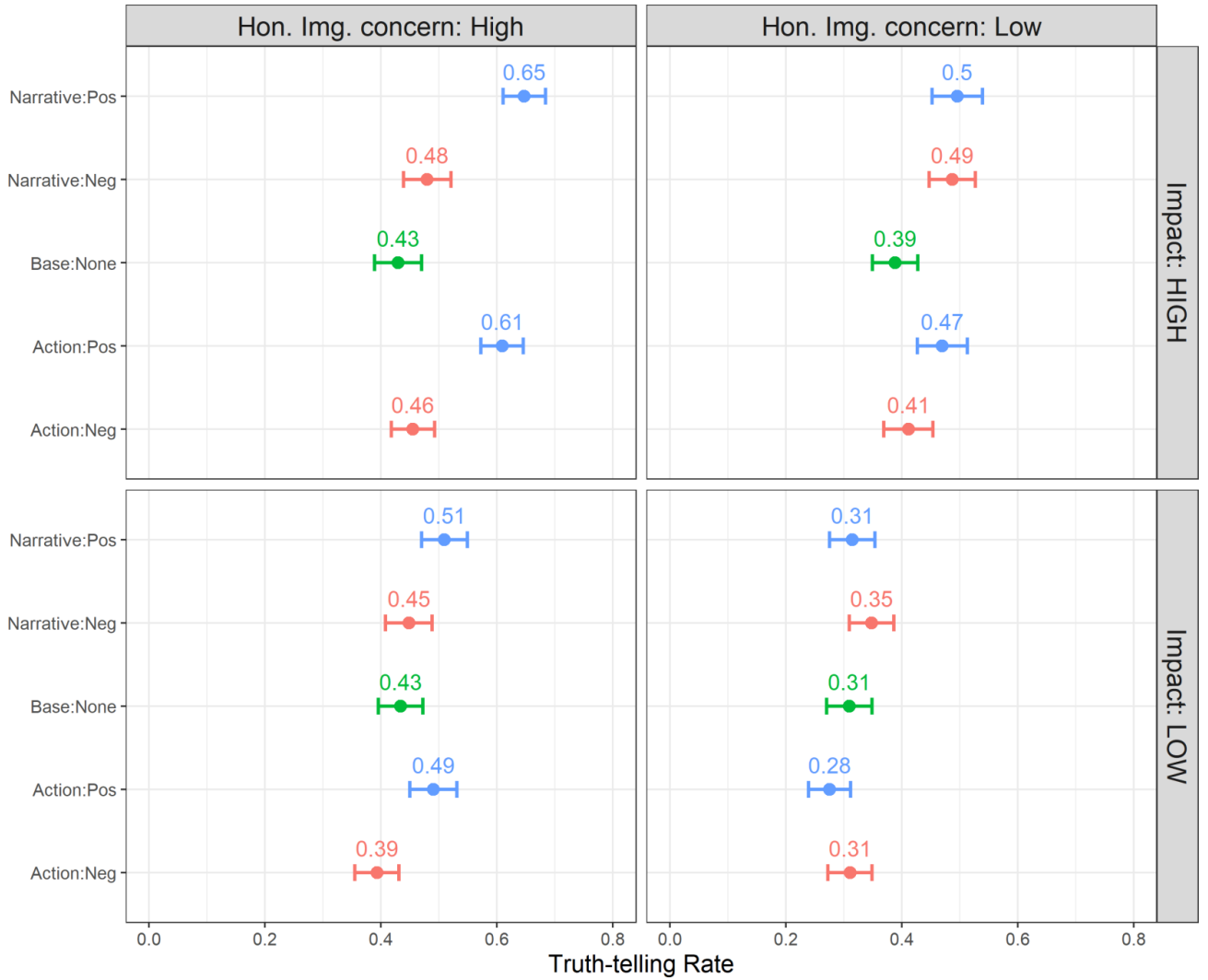


By Honesty Image Type

STUDY 2

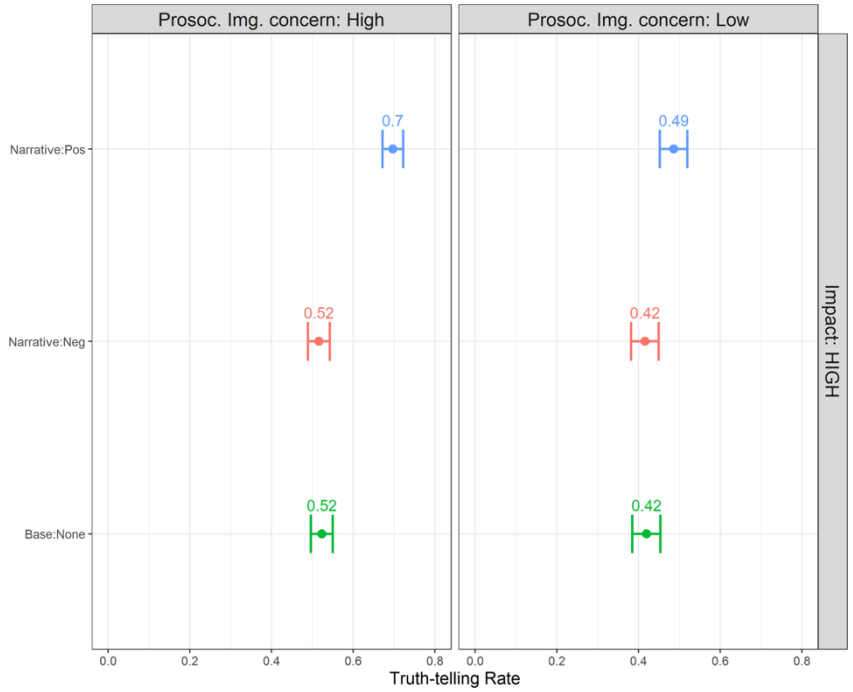


STUDY 1

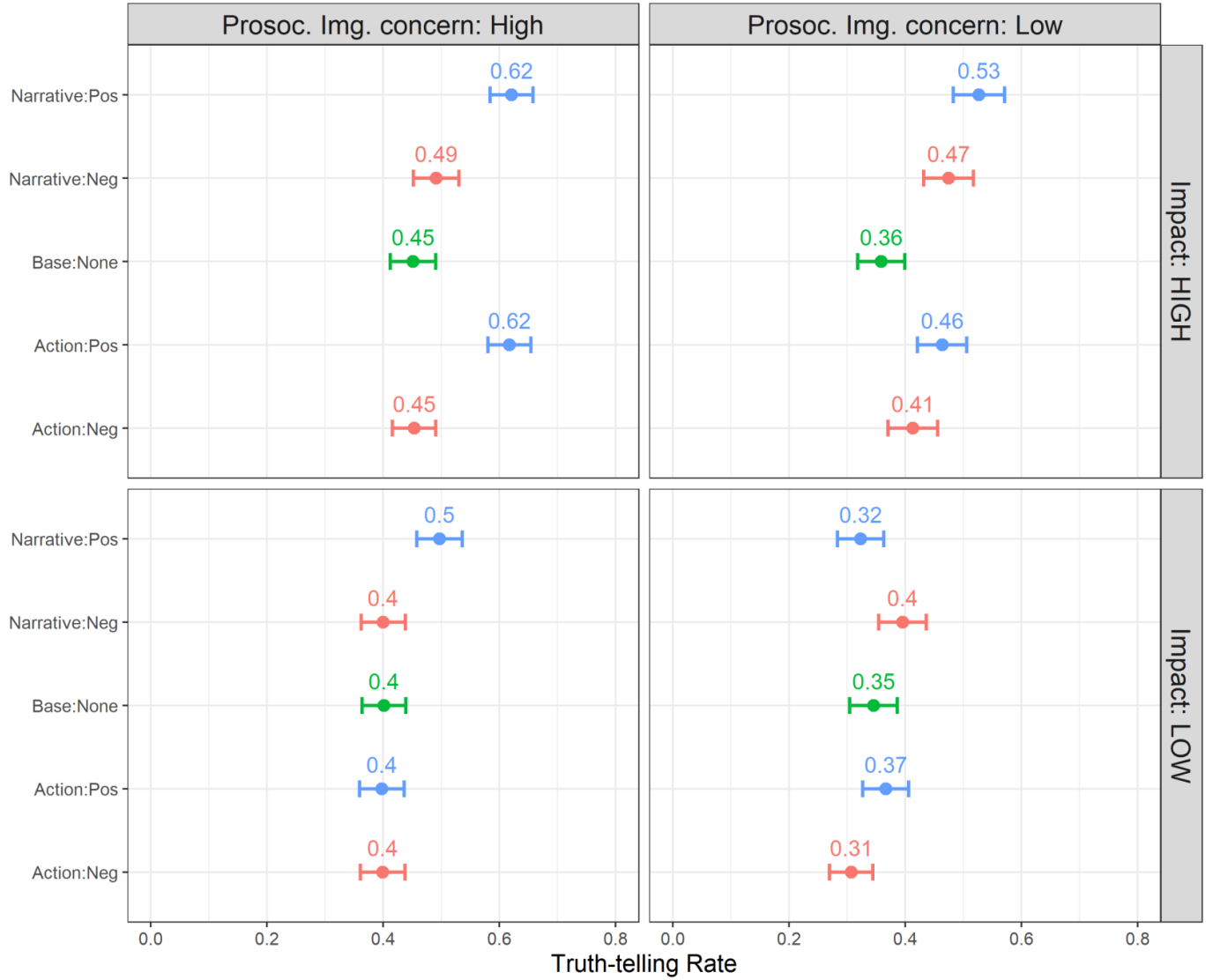


By Prosocial Image Type

STUDY 2

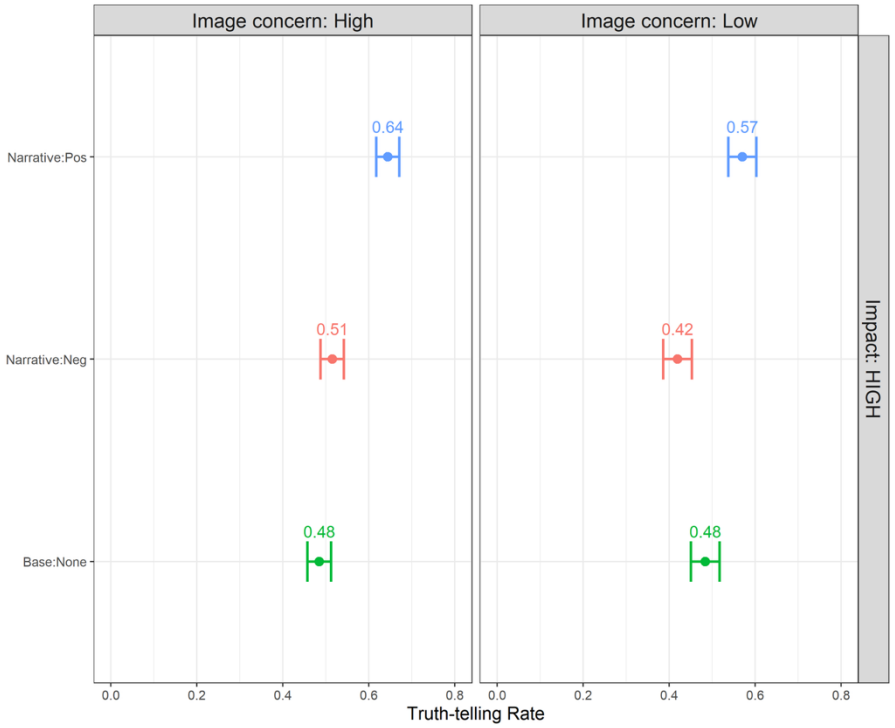


STUDY 1

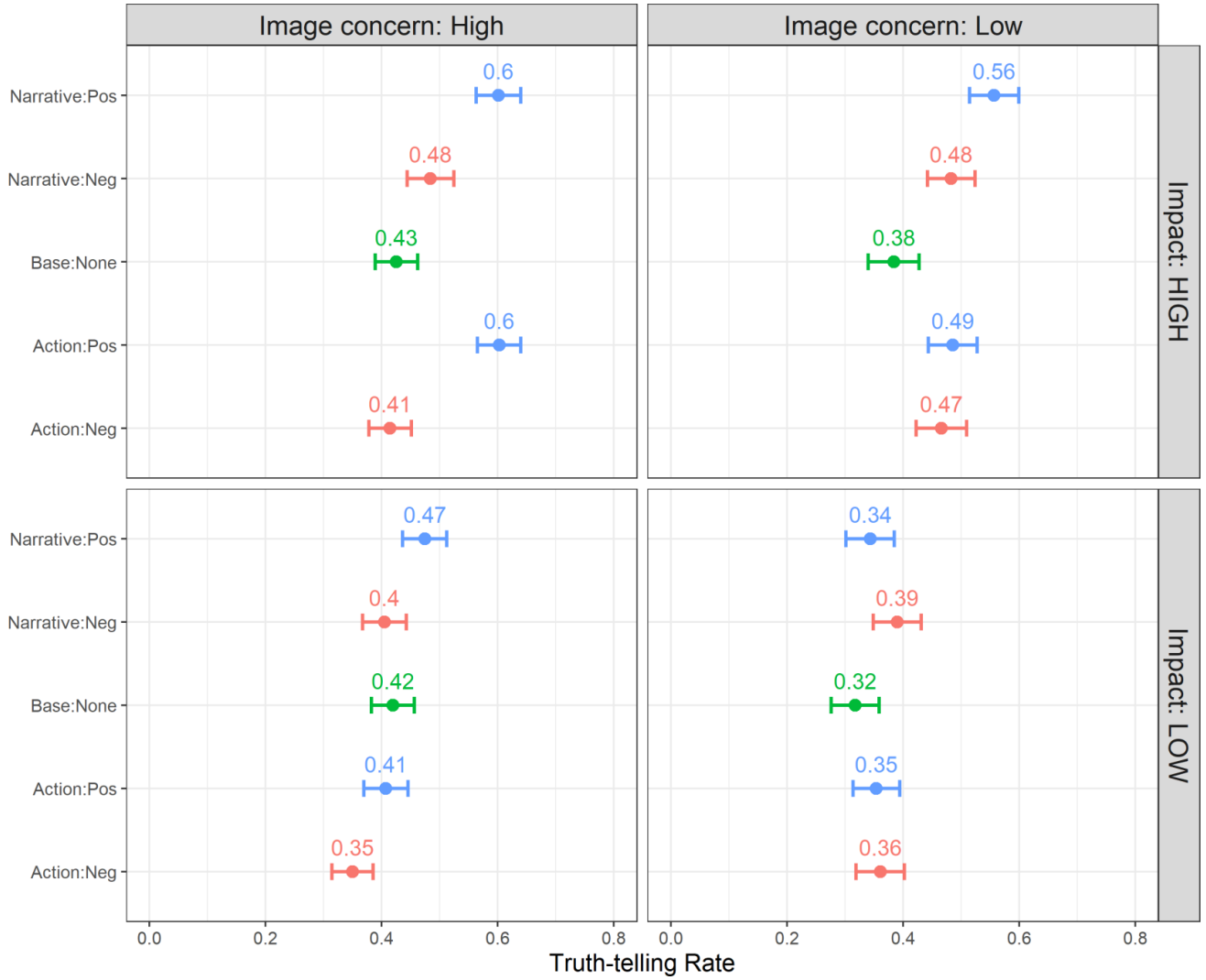


By General Image Type

STUDY 2



STUDY 1



Effect on narratives' content

- Coded by GPT-4
  - Positive/Negative
    - 94% consistency with S action
  - Persuasiveness (1-7)