Psychonomic Bulletin & Review (2026) 33:38
https://doi.org/10.3758/s13423-025-02807-1

BRIEF REPORT

How syntax promotes stereotypes: Assessing the role of pragmatic inference

Kevin J. Holmes[1] · Sarah H. Wu[1] · Nan Elpers[1,2] · Evan M. Doherty[2] · Stephen J. Flusberg[3]

Received: 10 June 2024 / Accepted: 28 August 2025

Abstract

Subtle linguistic differences can shape beliefs about the social world. For example, the statement "Girls are just as good as boys at math" leads some people to endorse the stereotype that boys have more natural math skill compared with a statement with the positions of the groups reversed. Traditional accounts of linguistic framing characterize such effects as an irrational consequence of biased cognitive and emotional processes. In contrast, we hypothesized that framing effects of this sort depend on the ability to pick up on the pragmatic implications of subject-complement syntax, where the group framed as the complement ("boys") is the implied standard or reference point. We investigated this possibility in two preregistered experiments ($N = 1{,}593$). Overall, participants who were better at *inferring* implicatures from subject-complement syntax were more likely to exhibit a framing effect by *endorsing* the implicature after reading subject-complement statements about math ability. This relationship held even when the statements referenced non-stereotyped groups and when controlling for other social-cognitive abilities associated with pragmatic competence. Framing effects were reduced for participants who explicitly recognized the statements as influencing their evaluations, but only when they invoked a stereotype to be discounted. These results suggest that pragmatic inference plays a crucial role in subject-complement framing but that people do not necessarily accede to what they infer. Our findings add to the growing body of evidence that many framing effects—far from being irrational—are a natural product of human communication.

Keywords Framing · Syntax · Pragmatic inference · Stereotypes · Individual differences

Introduction

Suppose someone tells you, "Girls are just as good as boys at math." You *could* take this statement at face value, as a plain affirmation of gender equality. But if you read between the lines, the syntactic structure of the statement belies the speaker's good intentions: the sentence positions girls as the subject and boys as the complement, implying, as the stereotype goes, that boys set the standard for math ability (Chestnut & Markman, 2016; Gleitman et al., 1996). Recent research shows that statements with this subject-complement framing can inadvertently lead people to endorse their harmful implications (i.e., that boys are more math-inclined by nature; Chestnut & Markman, 2018; Holmes et al., 2022). What mechanisms underlie the framing effects of subject-complement syntax? One key mechanism may be the ability to discern the implicit message behind the syntax—a form of *pragmatic inference* (Flusberg, Thibodeau et al., 2022a; Flusberg et al., 2024).[1] In this paper, we provide the first direct test of the role of pragmatic inference in subject-complement framing.

✉ Kevin J. Holmes
kjholmes@reed.edu

[1] Department of Psychology, Reed College, 3203 Woodstock Blvd, Portland, OR 97202, USA

[2] Department of Psychology, Colorado College, 14 Cache La Poudre St, Colorado Springs, CO 80903, USA

[3] Department of Cognitive Science, Vassar College, 124 Raymond Ave, Poughkeepsie, NY 12604, USA

[1] Following previous work on linguistic framing, we contrast "syntax" (the grammatical structures that give rise to framing effects) with "pragmatic inference" (a candidate mechanism underlying these effects). In using "syntax," we do not assume or endorse theories of language that posit minimalist representations divorced from semantics and pragmatics (e.g., Hauser et al., 2002), but instead use the term in a broad, descriptive sense compatible with constructionist accounts (e.g., Goldberg, 2003).

Published online: 05 January 2026

Framing and pragmatic inference

Traditional accounts of framing neglect pragmatic inference entirely. A common explanation for why ground beef is evaluated more favorably when described as "75% lean" than "25% fat," for example, is that "lean" evokes more positive associations (Levin et al., 1998). This valence-based judgment is seen as flawed or biased because the two linguistic frames are *logically equivalent*, differing in surface form but not literal meaning (De Martino et al., 2006; Tversky & Kahneman, 1981). An alternative explanation is that "75% lean" and "25% fat" carry different *implicit* meanings (i.e., that ground beef is typically less lean or less fatty, respectively) and listeners infer this subtle pragmatic information, adaptively incorporating it into their evaluations (Sher & McKenzie, 2006). Although evidence suggests that sensitivity to implicit meaning can account for several classic and real-world framing effects (e.g., Leong et al., 2017, 2023; Mandel, 2014; Tannenbaum et al., 2013), the role of pragmatic inference remains overlooked in popular characterizations of human decision-making as fundamentally irrational (Ariely, 2008; Kahneman, 2011). Even within the framing literature, research often fails to consider how inferences about implicit meaning might explain heterogeneity in the strength and reliability of a variety of framing effects (Flusberg et al., 2024).

In the case of subject-complement statements like "Girls are just as good as boys at math," initial evidence suggested that their implicit meaning is readily inferred and unquestioningly accepted. In a study by Chestnut and Markman (2018), participants read a report describing scientific evidence for gender equality in math achievement that contained several such statements, framing either girls or boys as the complement. Participants were more likely to affirm the stereotype—judging boys as naturally more skilled at math—when boys were framed as the complement than as the subject, despite the logical equivalence of these statements. When explicitly asked, however, participants rated the statements as relatively egalitarian, suggesting that they were unaware of the implicature they had inferred. Similar framing effects have been observed for subject-complement statements about other gender-stereotyped abilities (e.g., verbal ability; Chestnut & Markman, 2018), stereotyped behaviors (e.g., "Christians are just as likely as Muslims to commit terrorist acts"; Holmes et al., 2022), and even novel or non-stereotyped characteristics (e.g., "Girls are as good as boys at [yuzzing/snapping]"; Chestnut et al., 2021; Qian et al., 2025). These findings converge with evidence that many subtle aspects of grammar can shape reasoning, with important real-world consequences (e.g., Fausey & Boroditsky, 2010; Matlock et al., 2012).

Recent evidence suggests, however, that subject-complement statements do not inevitably give rise to framing effects, either because their implicit meaning goes unnoticed or because it is consciously discounted. In an extension of Chestnut and Markman (2018), Holmes et al. (2022) asked participants to indicate which part of the report on girls' and boys' math achievement had most influenced their judgments of which gender was more skilled. Across four experiments, the 70% of participants who cited the subject-complement statements in the report ("citers") consistently exhibited *weaker* framing effects than the other 30% ("non-citers"). Citers were often no more likely to attribute better math ability to boys, for example, when boys were the complement group than the subject group.

As people vary widely in their ability to extract pragmatic information from language (Bohn et al., 2023; Floyd et al., 2023), citers may have failed to detect the implicature communicated by the syntax. In these experiments, citing subject-complement statements as influential suggests that participants recognized them as communicating the *explicit* message that girls and boys performed equally well. However, it does not necessarily follow that they also inferred the *implicit* message framing the complement group as superior (Chestnut & Markman, 2016). A more direct test of pragmatic sensitivity might reveal citers to be relatively less attuned to the implicature than non-citers. Alternatively, citers may have been equally, if not *more*, attuned to the implicature than non-citers but then consciously rejected it. Several additional findings from Holmes et al. (2022) are consistent with this possibility. First, citers were disproportionately female and politically liberal—characteristics associated with concern for egalitarianism (Graham et al., 2011). Second, even though these statements are generally regarded as unbiased in explicit judgments (Chestnut & Markman, 2018), citers judged them significantly *more* biased than non-citers. Finally, citers sometimes exhibited *reverse* framing effects (e.g., judging the complement group as *less* math-inclined), as if their preexisting counter-stereotypical beliefs motivated them to discount the implicature. A similar pattern has been observed in other framing studies (Flusberg, van der Vord et al., 2022b; Thibodeau & Boroditsky, 2011).

The present research

Why does subject-complement syntax sometimes perpetuate stereotypes and sometimes have no impact or even the opposite effect? Previous studies do not provide a complete answer because they did not directly assess sensitivity to the pragmatic implications of the syntax, instead presuming it from overall framing effects (Chestnut & Markman, 2018; Chestnut et al., 2021) or from whether such effects were moderated by explicitly recognizing the impact of the framing language (Holmes et al., 2022). In the present work, we

created a measure of pragmatic sensitivity to subject-complement framing, modeled after studies assessing individual differences in pragmatic ability (e.g., Mayn & Demberg, 2022) and their role in other framing effects (e.g., Leong et al., 2017). Across two experiments, we assessed whether pragmatic sensitivity underlies framing effects for subject-complement statements about math ability. If so, participants who interpret the statements pragmatically rather than literally should generally exhibit stronger framing effects, attributing better ability to the complement group.

Across experiments, we also varied whether the groups referenced in the subject-complement statements were stereotyped for math ability (e.g., girls and boys) or not (e.g., Montanans and Wyomingites). This enabled us to gain further insight into why citers—those who explicitly recognize the statements as influencing their evaluations—exhibited weaker framing effects than non-citers in prior work (Holmes et al., 2022). If preexisting counter-stereotypical beliefs lead citers to recognize but ultimately reject the implicit message of the statements, they should exhibit weaker effects only if the statements invoke a stereotype. In the absence of a stereotype, citers should have little reason to discount the message and their framing effects should be similar to those of non-citers.

Experiment 1: Stereotyped groups

In Experiment 1, we tested whether pragmatic inference, among several other social-cognitive abilities, predicts the framing effects of subject-complement statements about math ability, and whether such effects are moderated by explicit recognition of the impact of the statements.

Method

We preregistered our methods and analysis plans for both experiments on AsPredicted. Our preregistrations, materials, and data are available on the Open Science Framework (https://osf.io/xqfkb/). Methods were approved by the Institutional Review Board at Reed College. All participants provided informed consent.

Participants

Using CloudResearch, we recruited 658 U.S. adults from Amazon Mechanical Turk with ≥95% approval on ≥100 studies. Participants who failed attention checks ($n = 27$) or did not complete all measures ($n = 34$) were excluded. We aimed for 300 participants per condition, more than in similar framing studies (~100; Flusberg et al., 2024). Our post-exclusion sample size ($N = 597$) provided >99% power to detect a 0.15 odds ratio for the subject-complement framing × citer-status interaction from Holmes et al.'s (2022) pooled analysis of binary math-ability judgments. Experiment 1 also had >80% power to detect weaker interactions (OR ≤ 0.58). Participants from Holmes et al. (2022) were ineligible for the present experiments. Each participant received $1.25 upon completion. Table 1 shows demographic data for both experiments.

Table 1 Demographic data

Variable	Experiment 1	Experiment 2	
		Experiment 2a	*Experiment 2b*
N (sampled / analyzed)	658 / 597	674 / 601	476 / 395
Gender (female / male)	52% / 47%	60% / 39%	53% / 46%
Mean age in years (*SD*)	42 (13)	41 (14)	40 (12)
Race/ethnicity	75%	73%	74%
White	11%	8%	8%
Black	6%	9%	6%
Asian	4%	3%	5%
Latinx	3%	4%	4%
Multiracial			
Political affiliation (Democrat / Republican)	44% / 27%	46% / 23%	44% / 26%

Materials and procedure

The experiment was administered online via Qualtrics.

Framing manipulation and outcome measures First, participants were presented with a report of scientific evidence for gender equality in math achievement (Chestnut & Markman, 2018). Participants were randomly assigned to read one of two versions (see Table 2). In the *Girls = Boys* condition ($n = 299$), girls were framed as the subject and boys as the complement in all three subject-complement statements in the report. In the *Boys = Girls* condition ($n = 298$), the positions of the two groups were reversed.

After 15 s, participants were able to advance to the next screen. There they were asked to choose which gender (girls or boys) they believed has better math ability (*"Based on these findings, who do you think are naturally more skilled at math?"*) and rate how confident they were in their choice (0, *not at all confident*; 100, *extremely confident*).[2] On the

[2] The confidence ratings mitigate demand characteristics that might be invited by the forced-choice judgments. If participants were reluctant to select either gender, they could express little or no confidence in their choice.

Table 2 Framing stimuli in Experiment 1

Recent Study: [Girls/Boys] Equal [Boys/Girls] at Math
A recent study has shown that [girls/boys] do just as well as [boys/girls] at math. At the University of Wisconsin, a team of researchers analyzed scores from standardized tests taken in 2005, 2006, and 2007 by approximately seven million students in ten different states. Overall, they found that [girls/boys] perform as well as [boys/girls] in grades two through eleven. A troubling finding from the study, however, is that many tough math questions seem to have been removed from state tests. The researchers worry that teachers, as a result, may start dropping harder math problems from their curriculums.

Note. Within each set of brackets, the first word was presented in the *Girls = Boys* condition and the second word was presented in the *Boys = Girls* condition

next screen, the report was presented again. Participants were asked to indicate which part was most influential in their judgment by copying and pasting it into a text box, as well as any other information that contributed to their judgment. Whether any of the subject-complement statements were cited in these responses was our measure of explicitly recognizing their impact (Holmes et al., 2022; Rook & Holmes, 2023).

Social-cognitive predictors On subsequent screens, participants completed our pragmatic sensitivity measure (see below) as well as established self-report and performance-based measures of *reflective thinking* (Need for Cognition–NFC: Cacioppo et al., 1984; Cognitive Reflection Test–CRT-2: Thomson & Oppenheimer, 2016), *theory of mind* (Autism Spectrum Quotient–AQ-10: Allison et al., 2012; Reading the Mind in the Eyes–RMET: Olderbak et al., 2015), and self-reported *social desirability bias* (Marlowe-Crowne Scale–SDS: Reynolds, 1982). For additional details, see Supplementary Material. These measures were presented in a fixed order (CRT-2, RMET, pragmatic sensitivity, SDS, NFC, AQ-10) and followed by demographic questions.

Pragmatic sensitivity Subject-complement statements about ability imply that the complement is more skilled than the subject despite literally expressing equivalence. Therefore, one should infer that a person who *believes Xs* are superior to *Ys* is more likely to *say* "*Ys* are as good as *Xs*" than "*Xs* are as good as *Ys*" (*infer frame-choice*). Likewise, one should infer that a person who *says* "*Ys* are as good as *Xs*" is more likely to *believe* that *Xs* are superior than that *Ys* are (*infer belief*). Our measure of pragmatic sensitivity assessed the ability to draw both types of inferences, which have been implicated in other equivalency framing effects (Leong et al., 2017; McKenzie & Nelson, 2003).

The measure began with a cover story about a series of fictional studies comparing two novel groups (Balurians and Arigans; cf. Vasilyeva & Lombrozo, 2020) on tests of four abilities (e.g., running, drawing). The groups were said to have "performed equally well" on all tests despite the existence of stereotypes favoring one group over the other for each ability. The passage explained that a journalist sought to describe these findings accurately, but that he personally endorsed the stereotypes even as the study findings opposed them. After 10 s, participants were able to advance to the next screen, where they answered four *infer-frame-choice* questions. For each, participants read about a stereotype (e.g., "Arigans are stereotyped to be better than Balurians at running") and selected which of two minimally-different subject-complement statements the journalist would be more likely to choose given his beliefs (e.g., "The study showed that [Balurians/Arigans] are just as good as [Arigans/Balurians] at running"). The statement in which the complement was the stereotypically-better group was scored as the pragmatically-correct answer, indicating sensitivity to the implicature.

On the next screen, another passage described additional studies showing that the two groups had performed equally well on tests of four other abilities, again contrary to stereotypes. A second journalist was described as wanting to cover the study findings accurately despite personally endorsing the stereotypes. After 10 s, participants were able to advance to the next screen, where they answered four *infer-belief* questions. For each, participants read a subject-complement statement attributed to the journalist (e.g., "The study showed that Balurians are just as good as Arigans at cooking"). Then they selected which group (Balurians or Arigans) the journalist seemed to believe was more skilled at the referenced ability, given her choice of statement. Selecting the complement group (Arigans in the above example) was scored as pragmatically correct.

Within each question block, the four abilities were presented in a randomized order. For each participant, the order of the two choices was counterbalanced across questions. The measure had good reliability ($\alpha = .83$), suggesting that the *infer-frame-choice* and *infer-belief* questions indexed the same construct. We computed each participant's proportion of pragmatically correct answers across both blocks.

Results

Framing effects

To assess the subject-complement framing effect and whether it was moderated by explicit recognition of the impact of the framing statements, we used a preregistered logistic regression model predicting participants' binary judgments (girls or boys, coded as −1 and 1) from three variables: frame (*Girls* = *Boys*, 1; *Boys* = *Girls*, 0), whether the participant was a "citer" (i.e., cited at least one subject-complement statement from the report, or a paraphrase, as most influential; coded as 1) or "non-citer" (0), and the interaction of these factors. In this analysis, the interaction was significant, $OR = 0.40$, 95% CI [0.19, 0.88], $p = .02$.

To unpack the interaction, we examined the simple effects of frame by contrast-coding the citer-status variable. As shown in Fig. 1, non-citers (29% of participants) exhibited a large framing effect, attributing better math ability to boys far more often in the *Girls* = *Boys* condition (77%, $SE = 4\%$) than the *Boys* = *Girls* condition (35%, $SE = 6\%$), $OR = 6.31$ [3.24, 12.30], $p < .001$. For citers, the framing effect was reduced but still significant (*Girls* = *Boys*: 51%, $SE = 4\%$; *Boys* = *Girls*: 29%, $SE = 3\%$), $OR = 2.55$ [1.71, 3.81], $p < .001$. Notably, judgments differed significantly between citers and non-citers in the *Girls* = *Boys* condition (where the framing aligned with the preexisting stereotype; $p < .001$), but not in the *Boys* = *Girls* condition (where it did not; $p = .32$; see Fig. 1). The same patterns were observed in analogous linear regression analyses of judgments weighted by participants' confidence ratings (see Supplementary Material for results from both experiments).

The role of pragmatic sensitivity

As a proxy for whether individual responses were affected by framing, we coded which group each participant chose as having better math ability across conditions: the complement group (1) or the subject group (−1). This choice was predicted by performance on the pragmatic sensitivity measure: the proportion of pragmatically-correct answers (overall $M = .76$, $SD = .29$) exceeded chance (.5; i.e., literal interpretation) for participants who chose the complement group ($M = .80$, $SD = .27$, $n = 385$) and for participants who chose the subject group ($M = .70$, $SD = .31$, $n = 212$; $ps < .001$), but was significantly higher for those who chose the complement group, $t(595) = 4.06$, $p < .001$, $d = 0.35$. As shown in Table 3, this relationship remained significant in a preregistered logistic regression analysis that included all social-cognitive measures as predictors. Pragmatic sensitivity, though correlated with all but one other measure ($|rs| = .11–.28$; $ps < .01$), was the strongest predictor of choosing

Table 3 Predictors of choosing the complement group

Predictor	Experiment 1		Experiment 2	
	β	*p*	β	*p*
Pragmatic sensitivity	**0.32**	**< .001**	**0.26**	**< .001**
NFC	0.02	.82	−0.03	.65
CRT-2	−0.14	.14	0.08	.27
AQ-10	0.11	.23	< 0.01	.99
RMET	**0.23**	**.02**	0.12	.11
SDS	−0.09	.33	< 0.01	.96

Note. Significant predictors ($p < .05$) are bolded. NFC = Need for Cognition; CRT-2 = Cognitive Reflection Test-2; AQ-10 = Autism Quotient-10; RMET = Reading the Mind in the Eyes Test; SDS = Social Desirability Scale

Fig. 1 Results of Experiment 1, indicating the group to which participants attributed better math ability, by frame and whether or not participants cited subject-complement statements as influential in their judgment. Error bars represent ±1 *SE*

the complement group (see Supplementary Material for analyses of confidence-weighted choices, which yielded similar results in both experiments).[3]

In an exploratory analysis (see Supplementary Material), we added citer status to the regression model as an additional predictor. Citer status and pragmatic sensitivity each explained significant unique variance ($|\beta|$s > 0.09, ps < .03), with non-citers and participants with better pragmatic sensitivity choosing the complement group more often. Pragmatic sensitivity did not differ significantly between citers (M = .77, SD = .29) and non-citers (M = .75, SD = .27), $t(595) = -0.81$, $p = .42$.

Discussion

These results provide initial evidence that pragmatic inference underlies the framing effects of subject-complement syntax. Participants who were generally more sensitive to the implicature—inferring, for instance, that a journalist who used a subject-complement statement believed the complement group was more skilled—exhibited stronger framing effects for seemingly unrelated judgments of math ability, over and above the contribution of other social-cognitive skills. Whereas the pragmatic sensitivity measure assessed explicit reasoning about the link between people's stereotypical beliefs and their use of subject-complement statements, the critical framing task probed whether participants' *own* beliefs were swayed by the preceding syntactic frame. Moreover, the pragmatic sensitivity measure required participants not only to infer someone's belief from their choice of statement (arguably analogous to the main task), but also to infer someone's choice of statement from their belief. Exploratory analyses revealed that these two types of inferences separately predicted framing effects in the main task (see Supplementary Material), suggesting that the role of pragmatic sensitivity in subject-complement framing cannot be reduced to similarities in the measures assessing these constructs.

Additionally, we found that participants who explicitly cited the subject-complement statements as influencing their evaluations exhibited weaker framing effects than those who did not, replicating previous findings (Holmes et al., 2022). Notably, the two factors that moderated framing effects in this experiment—pragmatic sensitivity and explicit recognition—were independent of each other and worked in opposite directions. This suggests that sensitivity to the implicit message of subject-complement statements is not the same as explicitly recognizing their impact: rather than blindly acceding to the message, citers appeared to discount it. In other words, they may have cited the statements to signal awareness of their biasing potential, justifying their choice of the subject group rather than the complement group as more math-inclined. Consistent with this interpretation, citers' and non-citers' judgments only differed in the *Girls = Boys* condition, in which the implicature aligned with the preexisting stereotype, providing sufficient motivation to resist it.

In Experiment 2, we extended these findings by assessing framing effects for statements about groups not specifically stereotyped for math ability. Of interest was whether such effects would still be moderated by pragmatic sensitivity, but no longer by explicit recognition, in the absence of a stereotype motivating people to resist the implicature.

Experiment 2: Unexpected stereotype and no stereotype

The subject-complement statements in Experiment 2 referenced children from two Canadian provinces (Experiment 2a) or two U.S. states (Experiment 2b). As Experiment 2a may have inadvertently invoked ability-relevant stereotypes, the groups in Experiment 2b were pretested to be non-stereotyped.

Method

Participants

Following the criteria of Experiment 1, we recruited 658 participants in Experiment 2a and 476 participants in Experiment 2b. Participants who failed attention checks (n = 67), did not complete all measures (n = 72), or were from the states referenced in the stimuli (Experiment 2b: n = 15) were excluded. We aimed for 200 participants per condition, comparable to Holmes et al. (2022). Our post-exclusion sample sizes (Experiment 2a: n = 601; Experiment 2b: n = 395) provided >92% power to detect the key effect sizes from Experiment 1 (frame × citer-status interaction: OR = 0.40; difference in pragmatic sensitivity between participants who chose the complement vs. subject group: d = 0.35). See Table 1 for participant demographics.

Materials and procedure

Except as indicated below, the methods were identical to Experiment 1. In Experiment 2a, participants were randomly assigned to the *Manitobans = Albertans* (n = 198), *Albertans = Manitobans* (n = 200), or *Baseline* (n = 203) condition. In the first two conditions, the math report described

[3] We also preregistered analyses predicting citer status as an index of whether individual responses were affected by framing (see Supplementary Material), but this is a poor index because even citers chose the complement group more often than not (see Fig. 1).

a fictional study comparing "children from Alberta" and "children from Manitoba" (see Supplementary Material). The *Baseline* condition assessed baseline intuitions about these groups by omitting the math report. In this condition, 73% of participants attributed better math ability to children from Alberta (binomial test: $p < .001$). This unexpected pattern may reflect biases favoring urban over rural students (Bryant, 2010). As a result, Experiment 2a appears to have inadvertently invoked a stereotype, whether preexisting or generated ad-hoc by participants upon comparing the two Canadian provinces.

We addressed this issue in Experiment 2b by first identifying groups judged to have similar baseline math ability. A separate sample of 100 participants were asked to choose, for each of nine pairs of groups (e.g., Swedish- vs. Finnish-speaking children), which group is naturally more skilled at math. The pair yielding closest to an even split were children from Wyoming and children from Montana (57% chose the latter; binomial test: $p = .19$; see Supplementary Material). Accordingly, participants in Experiment 2b were randomly assigned to the *Montanans = Wyomingites* ($n = 198$) or *Wyomingites = Montanans* ($n = 197$) condition and read a corresponding version of the math report (see Supplementary Material).

Results and discussion

Framing effects

Following Experiment 1, we analyzed the binary judgments using a preregistered logistic regression model. In Experiment 2a, the interaction between frame (*Manitobans = Albertans*, 1; *Albertans = Manitobans*, 0) and citer status was not significant, $OR = 0.40$ [0.14, 1.14], $p = .09$, though the framing effect was nominally smaller for citers (81% of participants; see Fig. 2a). This reflects citers' significantly reduced likelihood of choosing the complement group in the *Manitobans = Albertans* condition (where the framing aligned with the unexpected stereotype identified in the *Baseline* condition; $p = .01$), but not in the *Albertans =*

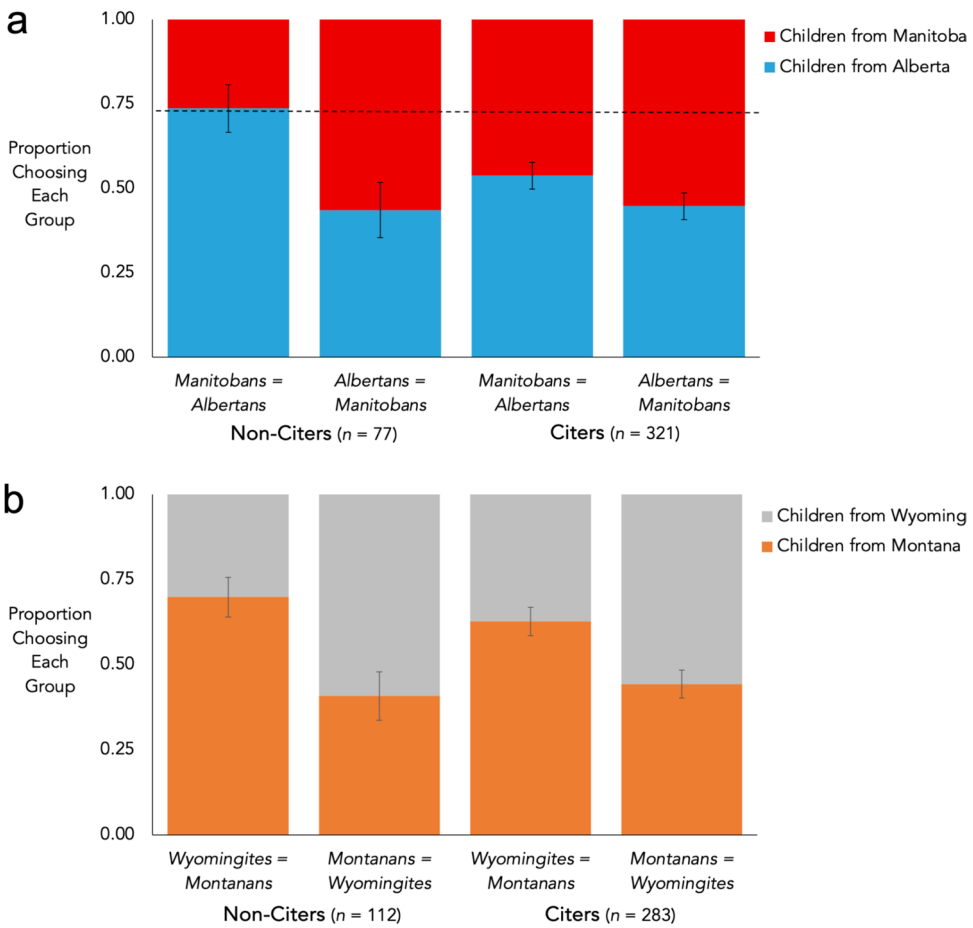

Fig. 2 Results of (**a**) Experiment 2a and (**b**) Experiment 2b, by frame and whether or not participants cited subject-complement statements as influential in their judgment. Error bars represent ±1 *SE*. In (a), the dashed line represents the proportion of *Baseline* participants who chose children from Alberta as superior at math

Manitobans condition (where it did not; $p = .89$). Overall, participants attributed better math ability to children from Alberta more often in the *Manitobans = Albertans* condition (58%, $SE = 4\%$) than the *Albertans = Manitobans* condition (45%, $SE = 4\%$), $OR = 1.69$ [1.14, 2.52], $p = .009$.

In Experiment 2b, the corresponding interaction (*Wyomingites = Montanans*, 1; *Montanans = Wyomingites*, 0) was not significant, $OR = 0.63$ [0.25, 1.57], $p = .32$. Participants attributed better math ability to children from Montana more often in the *Wyomingites = Montanans* condition (65%, $SE = 7\%$) than the *Montanans = Wyomingites* condition (43%, $SE = 7\%$), $OR = 2.42$ [1.61–3.63], $p < .001$. As shown in Fig. 2b, citers (71% of participants) and non-citers exhibited similar framing effects, with no significant difference in their judgments in either condition ($ps > .15$).

To summarize, citers exhibited somewhat weaker framing effects than non-citers in Experiment 2a for unexpectedly-stereotyped groups, driven by the condition in which the implicature aligned with the stereotype. These results mirror the pattern in Experiment 1 for unambiguously-stereotyped groups. In contrast, citers and non-citers exhibited similar framing effects in Experiment 2b for non-stereotyped groups. This suggests that subject-complement framing may have a broader impact when there is no stereotype to discount: in this case, even people who cite the framing statements as influential accede to the implicature.

The role of pragmatic sensitivity

As in Experiment 1, the proportion of pragmatically-correct answers (overall $M = .79$, $SD = .28$) was higher for participants who chose the complement group as better at math ($M = .83$, $SD = .25$, $n = 465$) than for those who chose the subject group ($M = .75$, $SD = .30$, $n = 328$), $t(791) = 4.06$, $p < .001$, $d = 0.29$, although both subsamples performed above chance ($ps < .001$). When controlling for other social-cognitive abilities, pragmatic sensitivity was the only significant predictor of choosing the complement group (see Table 3 and Supplementary Material).

Exploratory regression analyses showed that the framing effects in Experiment 2a were independently moderated by pragmatic sensitivity and citer status. Each of these factors explained significant unique variance in choosing the complement group ($|\beta|s > 0.25$, $ps < .03$), replicating Experiment 1. Absent a stereotype in Experiment 2b, however, citer status was not a significant predictor ($\beta = -0.16$, $p = .15$; see Supplementary Material). Moreover, in Experiments 2a and 2b, citers exhibited better pragmatic sensitivity than non-citers (overall: citers, $M = .82$, $SD = .26$; non-citers, $M = .72$, $SD = .30$), Experiment 2a: $t(396) = -3.00$, $p = .003$, $d = 0.38$; Experiment 2b: $t(393) = -2.94$, $p = .003$, $d = 0.33$, yet their framing effects were weaker in Experiment 2a. These results suggest that subject-complement framing effects generally reflect people's ability to pick up on the pragmatic implications of the framing language, but that those who are *especially* pragmatically savvy recognize when this language aligns with a stereotype and avoid affirming it.

General discussion

Statements like "Girls are just as good as boys at math," though often well-intentioned, risk promoting the stereotype that boys are better. Our findings show that this risk is generally greater for people who are more sensitive to the pragmatic implications of such statements. Across experiments, participants who were better at *inferring* implicatures from subject-complement syntax in our novel pragmatic sensitivity measure were more likely to *endorse* the implicature for statements about math ability in the critical framing task. This relationship held even when the statements referenced non-stereotyped groups and when controlling for social-cognitive abilities like reflective thinking and theory of mind that have been linked to pragmatic competence (Fairchild & Papafragou, 2021; Mayn & Demberg, 2022). These results suggest that pragmatic inference is a crucial mechanism underlying subject-complement framing effects.

At the same time, our findings show that when the framing language invokes a stereotype, many people who infer the implicature do *not* endorse it. Statements about stereotyped groups (e.g., girls and boys) elicited weaker framing effects for participants who explicitly cited them as influencing their evaluations, replicating previous work (Holmes et al., 2022), but statements about non-stereotyped groups did not. Moreover, for stereotyped groups, citers were less likely to endorse the implicature only when the framing aligned with the stereotype (e.g., when boys were the complement group, but not when girls were), and they performed no worse than non-citers, and sometimes significantly better, on our pragmatic sensitivity measure. Together, these results suggest that when stereotypes are at stake, subject-complement statements prompt many people—perhaps those especially adept at discerning their biasing potential—to push back against the implicit message.

Exploratory analyses revealed that pragmatic sensitivity and explicitly recognizing the impact of the statements often counteracted each other, independently moderating the framing effects in Experiments 1 and 2a in opposite directions. We interpret these results as evidence that inferring the implicature is necessary for subject-complement framing effects to occur, but that once inferred, the implicature is routinely discounted when the implied meaning motivates resistance. We acknowledge, however, that our evidence for discounting is indirect, relying on cross-experiment comparisons of participants' post-hoc rationales for their judgments. Notably, this evidence converges with

several previous findings: that, relative to non-citers, citers explicitly rate subject-complement statements as biased, have demographic characteristics associated with resisting anti-egalitarian stereotypes, and sometimes exhibit reverse framing effects (Holmes et al., 2022). Nevertheless, future work would benefit from more directly manipulating people's motivation to push back against the implicit message. For example, subject-complement statements in the critical framing task could be attributed to an ill-informed or ideologically divergent writer. Participants should be highly motivated to discount the implicature in such contexts, and should therefore exhibit relatively weak framing effects (Flusberg et al., 2024; Moty & Rhodes, 2021).

Our findings also have implications for theories of linguistic framing more generally. Subject-complement statements exemplify two types of framing: *equivalency framing*, where logically equivalent messages are presented in different surface forms, and *grammatical framing*, where the grammatical form or syntactic structure of a message is manipulated (Flusberg et al., 2024). Both types of framing effects seem striking—the former because they are traditionally seen as irrational byproducts of biased cognitive and emotional processes (De Martino et al., 2006; Levin et al., 1998), the latter because they suggest that socially significant judgments are swayed by incidental linguistic details (Fausey & Boroditsky, 2010; Matlock et al., 2012). In showing that pragmatic inference underlies subject-complement framing effects, our findings suggest that these effects are not so irrational and that the grammatical details of the message are not so incidental: evaluating social groups differently based on their syntactic positions is a sensible response to the pragmatic information communicated by the syntax (Flusberg et al., 2024). From this pragmatics-informed perspective, what is striking about these framing effects is not that they defy logic or are outsized relative to the linguistic cues that give rise to them. It is that those who are better at discerning the stereotype-affirming message are more susceptible to endorsing it. That said, our findings suggest that endorsement is far from inevitable: pragmatic sensitivity equips people not only to discern the message, but also to interrogate and resist it (Petty & Cacioppo, 1979). This is one reason why framing often backfires when people hold strong prior beliefs about the content of the message (Flusberg et al., 2024).

In future work, it will be important to determine whether our findings generalize beyond attributions of math ability. Statements like "Christians are just as likely as Muslims to commit terrorist acts" may be a useful test case because they seem to elicit more resistance than statements about math ability, perhaps because they invoke a stronger or more morally salient stereotype (Holmes et al., 2022). Future research might also explore whether pragmatic sensitivity for subject-complement statements is associated with sensitivity to implicit meaning in language more generally, and whether a more general measure of pragmatic ability would predict these and other framing effects.

In sum, our results suggest that subject-complement syntax gains the power to promote stereotypes when people use pragmatic inference to discern its implicit message—and then fail to question it. Our findings add to the growing body of evidence that many framing effects are a natural product of human communication.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.3758/s13423-025-02807-1.

Author note We thank Enriqueta Canseco-Gonzalez, Greg Jensen, Liantao Shan, Paul Thibodeau, and members of the Thought Lab at Reed College for helpful discussion.

Funding This work was supported by a James McKeen Cattell Sabbatical Award to K.J.H.

Availability of data and materials Our preregistrations, materials, and data are available on the Open Science Framework (https://osf.io/xqfkb/).

Code availability We used JASP 0.18.3 for all analyses. Analysis files are available upon request.

Declarations

Competing interests The authors have no relevant financial or non-financial interests to disclose.

Ethics approval The experiments were performed in line with the principles of the Declaration of Helsinki. Methods were approved by the Institutional Review Board at Reed College.

Consent to participate All participants provided informed consent.

Consent for publication Not applicable. No identifying information about participants is included in this manuscript.

References

Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief "red flags" for autism screening: The short autism spectrum quotient and the short quantitative checklist in 1,000 cases and 3,000 controls. *Journal of the American Academy of Child & Adolescent Psychiatry, 51*(2), 202-212.e7. https://doi.org/10.1016/j.jaac.2011.11.003

Ariely, D. (2008). *Predictably irrational*. HarperCollins.

Bohn, M., Tessler, M. H., Kordt, C., Hausmann, T., & Frank, M. C. (2023). An individual differences perspective on pragmatic abilities in the preschool years. *Developmental Science, 26*(6), Article e13401. https://doi.org/10.1111/desc.13401

Bryant, J. A., Jr. (2010). Dismantling rural stereotypes. *Educational Leadership, 68*(3), 54–58.

Cacioppo, J. T., Petty, R. E., & Kao, C. F. (1984). The efficient assessment of need for cognition. *Journal of Personality Assessment, 48*(3), 306–307. https://doi.org/10.1207/s15327752jpa4803_13

Chestnut, E. K., & Markman, E. M. (2016). Are horses like zebras, or vice versa? Children's sensitivity to the asymmetries of directional comparisons. *Child Development, 87*(2), 568–582. https://doi.org/10.1111/cdev.12476

Chestnut, E. K., & Markman, E. M. (2018). "Girls are as good as boys at math" implies that boys are probably better: A study of expressions of gender equality. *Cognitive Science, 42*, 2229–2249. https://doi.org/10.1111/cogs.12637

Chestnut, E. K., Zhang, M. Y., & Markman, E. M. (2021). "Just as good": Learning gender stereotypes from attempts to counteract them. *Developmental Psychology, 57*(1), 114–125. https://doi.org/10.1037/dev0001143

De Martino, B., Kumaran, D., Seymour, B., & Dolan, R. J. (2006). Frames, biases, and rational decision-making in the human brain. *Science, 313*(5787), 684–687. https://doi.org/10.1126/science.1128356

Fairchild, S., & Papafragou, A. (2021). The role of executive function and theory of mind in pragmatic computations. *Cognitive Science, 45*(2), Article e12938. https://doi.org/10.1111/cogs.12938

Fausey, C. M., & Boroditsky, L. (2010). Subtle linguistic cues influence perceived blame and financial liability. *Psychonomic Bulletin & Review, 17*(5), 644–650. https://doi.org/10.3758/PBR.17.5.644

Floyd, S., Jouravlev, O., Mineroff, Z., Gibson, E., & Federenko, E. (2023). A tripartite structure of pragmatic language abilities: Comprehension of social conventions, intonation processing, and causal reasoning. *PsyArXiv*. https://doi.org/10.31234/osf.io/e2xta

Flusberg, S. J., Thibodeau, P. H., & Holmes, K. J. (2022). Even simple framing effects are rational: Commentary on Bermúdez. *Behavioral and Brain Sciences, 45*, Article e228. https://doi.org/10.1017/S0140525X22000942

Flusberg, S. J., van der Vord, J., Husney, S. Q., & Holmes, K. J. (2022). Who's the "real" victim? How victim framing shapes attitudes toward sexual assault. *Psychological Science, 33*(4), 524–537. https://doi.org/10.1177/09567976211045935

Flusberg, S. J., Holmes, K. J., Thibodeau, P. H., Nabi, R. L., & Matlock, T. (2024). The psychology of framing: How everyday language shapes the way we think, feel, and act. *Psychological Science in the Public Interest, 25*(3), 105–161. https://doi.org/10.1177/15291006241246966

Gleitman, L. R., Gleitman, H., Miller, C., & Ostrin, R. (1996). Similar, and similar concepts. *Cognition, 58*(3), 321–376. https://doi.org/10.1016/0010-0277(95)00686-9

Goldberg, A. E. (2003). Constructions: A new theoretical approach to language. *Trends in Cognitive Sciences, 7*(5), 219–224. https://doi.org/10.1016/S1364-6613(03)00080-9

Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology, 101*(2), 366–385. https://doi.org/10.1037/a0021847

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it, who has it, and how did it evolve? *Science, 298*(5598), 1569–1579. https://doi.org/10.1126/science.298.5598.1569

Holmes, K. J., Doherty, E. M., & Flusberg, S. J. (2022). How and when does syntax perpetuate stereotypes? Probing the framing effects of subject-complement statements of equality. *Thinking & Reasoning, 28*(2), 226–260. https://doi.org/10.1080/13546783.2021.1963841

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan.

Leong, L. M., McKenzie, C. R., Sher, S., & Müller-Trede, J. (2017). The role of inference in attribute framing effects. *Journal of Behavioral Decision Making, 30*(5), 1147–1156. https://doi.org/10.1002/bdm.2030

Leong, L. M., Müller-Trede, J., & McKenzie, C. R. M. (2023). Is it a judgment of representativeness? Re-examining the birth sequence problem. *Psychonomic Bulletin & Review, 30*, 731–738. https://doi.org/10.3758/s13423-022-02188-9

Levin, I. P., Schneider, S. L., & Gaeth, G. J. (1998). All frames are not created equal: A typology and critical analysis of framing effects. *Organizational Behavior and Human Decision Processes, 76*(2), 149–188. https://doi.org/10.1006/obhd.1998.2804

Mandel, D. R. (2014). Do framing effects reveal irrational choice? *Journal of Experimental Psychology: General, 143*(3), 1185–1198. https://doi.org/10.1037/a0034207

Matlock, T., Sparks, D., Matthews, J. L., Hunter, J., & Huette, S. (2012). Smashing new results on aspectual framing: How people describe car accidents. *Studies in Language, 36*(3), 700–721. https://doi.org/10.1075/sl.36.3.09mat

Mayn, A., & Demberg, V., et al. (2022). Individual differences in a pragmatic reference game. In J. Culbertson et al. (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society* (pp. 3016–3022). Cognitive Science Society. https://escholarship.org/uc/item/0hb5h0hk

McKenzie, C. R. M., & Nelson, J. (2003). What a speaker's choice of frame reveals: Reference points, frame selection, and framing effects. *Psychonomic Bulletin & Review, 10*(3), 596–602. https://doi.org/10.3758/BF03196520

Moty, K., & Rhodes, M. (2021). The unintended consequences of the things we say: What generic statements communicate to children about unmentioned categories. *Psychological Science, 32*(2), 189–203. https://doi.org/10.1177/0956797620953132

Olderbak, S., Wilhelm, O., Olaru, G., Geiger, M., Brenneman, M. W., & Roberts, R. D. (2015). A psychometric analysis of the reading the mind in the eyes test: Toward a brief form for research and applied settings. *Frontiers in Psychology, 6*,1503. https://doi.org/10.3389/fpsyg.2015.01503

Petty, R. E., & Cacioppo, J. T. (1979). Issue involvement can increase or decrease persuasion by enhancing message-relevant cognitive responses. *Journal of Personality and Social Psychology, 37*(10), 1915–1926. https://doi.org/10.1037/0022-3514.37.10.1915

Qian, Y., Goldin-Meadow, S., & Bian, L. (2025). Gesture counteracts gender stereotypes conveyed through subtle linguistic cues. *Proceedings of the National Academy of Sciences of the United States of America, 122*(4), Article e2415671122. https://doi.org/10.1073/pnas.2415671122

Reynolds, W. M. (1982). Development of reliable and valid short forms of the Marlowe-Crowne Social Desirability Scale. *Journal of Clinical Psychology, 38*(1), 119–125. https://doi.org/10.1002/1097-4679(198201)38:13.0.CO;2-I

Rook, E. D., & Holmes, K. J. (2023). How language shapes anti-fat bias: Comparing the effects of disease and fat-rights framing. *Frontiers in Communication, 8*, 1284074. https://doi.org/10.3389/fcomm.2023.1284074

Sher, S., & McKenzie, C. (2006). Information leakage from logically equivalent frames. *Cognition, 101*(3), 467–494. https://doi.org/10.1016/j.cognition.2005.11.001

Tannenbaum, D., Valasek, C. J., Knowles, E. D., & Ditto, P. H. (2013). Incentivizing wellness in the workplace: Sticks (not carrots) send stigmatizing signals. *Psychological Science, 24*(8), 1512–1522. https://doi.org/10.1177/0956797612474471

Thibodeau, P. H., & Boroditsky, L. (2011). Metaphors we think with: The role of metaphor in reasoning. *PLoS One, 6*(2), Article e16782. https://doi.org/10.1371/journal.pone.0016782

Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making, 11*(1), 99–113. https://doi.org/10.1017/S1930297500007622

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453–458. https://doi.org/10.1126/science.7455683

Vasilyeva, N., & Lombrozo, T. (2020). Structural thinking about social categories: Evidence from formal explanations, generics, and generalization. *Cognition, 204*, Article 104383. https://doi.org/10.1016/j.cognition.2020.104383

Open Practices Statement We preregistered our methods and analysis plans on AsPredicted. Our preregistrations, materials, and data are available on the Open Science Framework (https://osf.io/xqfkb/https://osf.io/xqfkb/).