Question Bias

Cory Bill and Todor Koev

Word count (excluding abstract and references) = 6,936

Abstract

This paper spells out the different flavors of question bias (original, contextual, projected) and additionally characterizes the bias inference in terms of its direction, strength, and optionality. With an eye on these distinctions, we review three major approaches to question bias and suggest that this phenomenon is primarily a matter of answer salience, where different kinds of salience correspond to different flavors of bias.

1 Introduction

The conventional role of positive polar questions is to elicit information from the addressee, keeping an equal balance between the two polar answers (Hamblin 1973; Groenendijk and Stokhof 1984; Ciardelli et al. 2019; a.m.o.). For example, in the absence of further contextual cues, the question in (1) would convey no preference for a given answer. On the other end of the spectrum are, for example, polar questions with minimizers, which effectively lack any interrogative force and serve as rhetorical devices (Borkin 1971; Heim 1984; Krifka 1995; Han 2002; Abels 2003; van Rooy 2003; Guerzoni 2004; Caponigro and Sprouse 2007). To illustrate, the question in (2) contains the minimizer *lift a finger* and conveys the message that John did absolutely nothing to help Mary.

(1) Is it raining outside? (positive polar question)

(2) Did John lift a finger to help Mary? (polar question with a minimizer)

Between these two extremes of purely information-seeking questions vs. questions with an obligatory rhetorical effect, there are a large number of non-canonical polar questions that seek to elicit information but that simultaneously lean towards one of the polar answers, thus giving rise to the implication of **question bias**. Some question forms that systematically give rise to this effect are: polar questions with high ('light', 'preposed') negation, accented low negation, a verum accent, or conversational *really*, as in (3) (Romero and Han 2004; a.o.); various tag questions, including nuclear and postnuclear reverse-polarity tag questions, as in (4) (e.g., Bill and Koev 2023a); and rising declaratives, as in (5) (e.g., Malamud and Stephenson 2015).[1] For example, the high negation question in (3a) conveys the speaker's prior belief that Jill eats meat, where this

[1] In linguistic examples, capitalization indicates prosodic prominence, a comma indicates a prosodic break, a question mark indicates a rising boundary tone, and a period indicates a falling boundary tone.

inference may clash with some bit of contextual evidence suggesting—to the contrary—that Jill does not eat meat.

(3) a. Doesn't Jill eat meat? (high negation question)
b. Does Jill NOT eat meat? (accented low negation question)
c. IS Oliver from Australia? (verum question)
d. Is Oliver really from Australia? (question with conversational *really*)

(4) a. It's snowing, ISN'T it. (nuclear reverse-polarity tag question)
b. It's snowing isn't it? (postnuclear reverse-polarity tag question)

(5) You are the new manager? (rising declarative)

Notably, there are also question forms that do not seem able to convey bias at all; for example, alternative questions like (6) (Bolinger 1978; Biezma and Rawlins 2012).

(6) Did ALFONSO or JOANNA give you a ride. (alternative question)

There is a sizeable chunk of literature on the phenomenon of question bias (for seminal ideas and representative accounts, see Ladd 1981; Büring and Gunlogson 2000; van Rooy and Šafářová 2003; Guerzoni 2004; Romero and Han 2004; Reese 2007; Krifka 2015; Malamud and Stephenson 2015; Farkas and Roelofsen 2017; AnderBois 2019; Bill and Koev 2023b; a.o.). In spite of this richness, there is often a lack of clarity regarding the specific flavor of bias that is being captured. When a distinction has been made, it is between original bias and contextual bias. **Original bias** is said to convey the prior belief of the speaker regarding the truth of a given answer (Ladd 1981; Romero and Han 2004), while **contextual bias** is about the evidential support for a given answer provided by the context (Büring and Gunlogson 2000; Romero and Han 2004). A third flavor of bias, which has not been clearly articulated but falls out from some of the theoretical construals, is what we will call **projected bias**. It concerns the speaker's expectation as to which answer will be chosen by the addressee, thus steering the resolution of the issue raised by the question in a particular direction (cf. Krifka 2015; Malamud and Stephenson 2015). Given these three flavors of bias (original, contextual, projected), this paper will provide analytical definitions, offer empirical diagnostics, and determine the degree to which these flavors overlap or contrast with each other.

Moreover, in much of the previous literature, even when a specific flavor of bias has being singled out, the main focus has been on capturing the **direction** of bias, i.e., whether the bias is positive (*for* the overtly expressed proposition, the question 'prejacent') or negative (*against* the overtly expressed proposition). Two additional features that have largely been ignored are **strength** and **optionality**. That is, the bias conventionally associated with a given question form can be weak (expressing a mere preference for one of the answers) or strong (implying a significant amount of certainty about one of the answers). Also, such bias may arise optionally (only in specific contexts) or obligatorily (across the board). Focusing on original bias, we will offer empirical diagnostics that tap into the relevant contrasts in feature settings.

The general goal of this paper is to put the different construals of question bias on an equal footing, thus facilitating their comparison and providing a solid foundation for future work. In doing so, we will attempt to bring some uniformity into this very diverse class of phenomenon by suggesting that question bias is primarily about **answer salience**, where different kinds of answer

salience correspond to different bias flavors. Specifically, it has been pointed out previously that the overt proposition expressed by a polar question is selected based on the speaker's conversational goals and must be compatible with the contextual evidence (Büring and Gunlogson 2000; van Rooy and Šafářová 2003; Romero and Han 2004). In addition to this 'surface' kind of salience, the speaker may choose to enrich a question form with various elements—polarity focus (Bill and Koev 2023b), scalar items like *even* (Guerzoni 2004), etc.—to make 'semantically' salient a given answer, thus conveying an original bias that pulls against the contextual evidence. Given the contrasting nature of original and contextual biases, any bias that is projected to enter the Common Ground will ultimately align with one or the other (cf. Krifka 2015; Malamud and Stephenson 2015). We discuss below the possibilities for how this alignment is ultimately determined.

The discussion to come is empirically restricted in two important ways. First, it is limited to bias in *polar* questions. This is in spite of Guerzoni (2003, ch.3)'s observation that constituent (as well as polar) questions in which *even* associates with a minimal element are biased as well, see (7). Second, we focus on bias in *root* questions, even though—as noted in Ladusaw (1980, ch.7)—bias may arise in embedded questions as well, see (8).

(7) Who can solve even the EASIEST problem?
 ⇝ *The speaker expects that no one can solve the easiest problem.*

(8) Mia wonders if Daniel really is the boss.
 ⇝ *Mia doubts that Daniel is the boss.*

We put these very interesting cases aside because they have barely been touched on in the literature, and also because they bring up issues (the availability of the negative answer in constituent questions, types of question-embedding predicates, etc.) that are largely orthogonal to the topic of question bias.

The paper will proceed as follows. Section 2 spells out the different flavors of bias (original, contextual, projected) and Section 3 discusses the different features of (original) bias, i.e., direction, strength and optionality. Section 4 examines three prior approaches to question bias, with an eye on their predictions regarding bias flavor and bias features. Finally, Section 5 concludes the discussion and proposes answer salience as a possible unifying mechanism for the different flavors of question bias.

2 Flavors of question bias

The phenomenon of question bias typically involves the following kind of scenario. The speaker comes to the conversational Table with a prior belief in mind, which is contradicted by some bit of evidence that has just manifested itself in the context. The speaker then asks a question to sort things out, while simultaneously projecting that one of the answers is more likely to be added to the Common Ground. While not all of these three components (prior belief, current evidence, projected answer) need to always be present, this picture gives a good first sense of the different flavors of bias.[2] The rest of this section elaborates.

[2]For example, rising declaratives do not seem to require a prior belief on the part of the speaker (see e.g. Malamud and Stephenson 2015's 'blushing'/'innuendo' scenario). Also, high negation questions do not always involve contradicting contextual evidence (see Ladd 1981's 'suggestion' scenario).

Perhaps the most prominent flavor of question bias is that of original bias, or what the speaker previously believed to be the true answer. This notion is backward-looking in that it targets the prior epistemic state of the speaker (although this state may still hold at utterance time). It is defined in (9).

(9) *Original Bias* (cf. Ladd 1981; Romero and Han 2004)
A question Q uttered in a context c conveys *original bias* for a possible answer p iff Q indicates that the prior epistemic state of the speaker in c supports p over $\neg p$.

In contrast, contextual bias is about the shared evidence regarding the possible answers, where this evidence has current relevance in the context. In the typical case, this would be evidence that has been implied by the addressee's prior utterance. This notion is defined in (10).

(10) *Contextual Bias* (cf. Büring and Gunlogson 2000; Romero and Han 2004)
A question Q uttered in a context c conveys *contextual bias* for a possible answer p iff Q indicates that some piece of evidence that has just become available in c supports p over $\neg p$.

A standard way of diagnosing both original and contextual bias is to set up the utterance context such that the assumed biases are entailed. Importantly, it has been pointed out that these two bias flavors are typically in disagreement about the question prejacent (Roelofsen et al. 2012; Domaneschi et al. 2017). The opposing directions of original and contextual bias are illustrated in (11) for the case of high negation questions, which typically express a combination of positive original bias (here, that we are going to a Vietnamese restaurant, due to a prior conversation) and negative contextual bias (here, that we are going to a Korean restaurant, due to the immediately preceding utterance).

(11) *A and B bump into each other at a linguistics conference. They agree to go for dinner to a Vietnamese restaurant. The last talk is finally over, and A and B are about to head out.*
A: Let's go to Sun Nong Dan, the best Korean restaurant in town.
B: Aren't we going to a Vietnamese restaurant?

Beyond these two flavors, several theoretical construals make predictions about bias that has to do with which of the answers is projected to enter the Common Ground (Krifka 2015; Malamud and Stephenson 2015; see also Abels 2003; Guerzoni 2004; AnderBois 2019). This notion is forward-looking as it is about the answer that is expected to be chosen by the addressee. It is tentatively defined in (12).

(12) *Projected Bias*
A question Q uttered in a context c conveys *projected bias* for a possible answer p iff Q indicates that the speaker in c expects answer p rather than $\neg p$.

Since the notion of projected bias has not been singled out, to our knowledge there are no set diagnostics that tap into it. But given the definition in (12), one would expect this flavor of bias to give rise to contrasts between plain responses (i.e., responses that merely select a given partition cell, like *yes* and *no*), where the partition cell representing the expected answer is more readily

accessible. The contrast in (13) illustrates this kind of diagnostic, which we will call the **plain response test**.[3] According to it, high negation questions convey (by default) a negative projected bias.

(13) A: Isn't Deniz going to Berlin?
B: No. / ?Yes.

Notice that with more elaborate responses the contrast is lost. For example, (13) could naturally be answered with *Well, yes, he actually is*. This is to be expected, as this longer response explicitly flags the selected (positive) answer as being the less likely choice.

Once these three flavors of bias are teased apart, the question arises as to how they are aligned with respect to each other. As already mentioned, original and contextual bias usually pull into opposite directions, one of them supporting the positive answer and the other supporting the negative answer.[4] Projected bias will then have no choice but to be aligned with the one and misaligned with the other in each particular context. Is there any systematicity here? One very simple strategy would be that projected bias is consistently aligned with one of the other two bias flavors. If this is original bias, it could be because the speaker prefers not to revise their beliefs (AnderBois 2019); and if this is contextual bias, it could be because the context is taken to provide objective evidence that generally trumps the speaker's subjective beliefs (cf. Koev 2019). One could also imagine a mixed strategy, where projected bias follows whichever of these two biases is perceived to be stronger.

An initial survey seems to indicate that projected bias is aligned with contextual bias. This is already apparent in (13), where the preferred negative answer echoes the negative contextual bias that may arise in high negation questions. Considering the set of biased questions in (3)–(5) leads to similar conclusions. That is, all these questions seem to display some preference for the plain response that can be assumed to agree with the contextual evidence. Specifically, accented low negation questions display a preference for the negative answer, while verum questions and questions with conversational *really* display a preference for the positive answer, as illustrated in (14).[5] This in line with the assumed contextual biases in such questions (Romero and Han 2004).

(14) a. Does Jill NOT eat meat? (No. / ??Yes.)
b. IS Oliver from Australia? (Yes. / ?No.)
c. Is Oliver really from Australia? (Yes. / ?No.)

Moreover, rising declaratives prefer the positive plain response, as shown in (15). And indeed, they are typically used when the prejacent is supported by the contextual evidence (Gunlogson 2008; Malamud and Stephenson 2015).

(15) You are the new manager? (Yes. / ?No.)

However, reverse-polarity tag questions seem to paint a different picture. That is, both nuclear and postnuclear variants show some preference for the positive plain response, see (16).

[3]The plain response test is suggested in Ladusaw (1980, ch.8) and Guerzoni (2004). In these works it is applied to different data, i.e., polar questions with minimizers like (2), which receive a rhetorical interpretation.

[4]This makes sense pragmatically, because if these two bias flavors were aligned, the issue expressed by the prejacent may count as settled and there may be no reason to ask the question in the first place.

[5]The dispreferred response seems worse in (14a) than in (14b)–(14c). This may be due to the fact that polar responses to low negation questions are ambiguous between targeting the negative or the positive proposition (Kramer and Rawlins 2011; Holmberg 2013; Krifka 2013; Roelofsen and Farkas 2015; Goodhue and Wagner 2018).

(16) a. It's sunny outside, ISN'T it. (Yes. / ?No.)
b. It's sunny outside isn't it? (Yes. / (?)No.)

This data is our first indication that contextual bias and projected bias may not always track together. That is, reverse-polarity tag questions have been widely assumed to convey an original bias for the anchor proposition (Rando 1980; Ladd 1981; Romero and Han 2004; Reese 2007; Krifka 2015; Farkas and Roelofsen 2017; Bill and Koev 2023a), and—crucially—a contextual bias against said proposition (Romero and Han 2004).

Another case where projected bias appears to align with original bias is that of high negation questions in 'suggestion' scenarios, in which the contextual bias is neutral. As shown in (17), in such scenarios there is a preference for the positive plain response, which aligns with the positive original bias of high negation questions.

(17) *After a long day at a linguistics conference.*
A: Hey, let's go eat!
B: Sure. Isn't there a vegetarian restaurant around here?
A: Yes. / ?No.

The last two data points (i.e. (16) and (17)) argue against the collapsing of contextual and projected bias into the same phenomenon. This leads us to propose that these two biases are truly different—though often aligned in direction, presumably due to some general constraint. Specifically, it seems plausible to us that a projected bias aligns with whichever of original and contextual bias is deemed to be stronger. In general, contextual bias may be deemed stronger, due to it being perceived as more objective or due to it being the most recent piece of evidence relevant to the prejacent. Then, when it comes to tag questions (i.e. (16)), the original bias may trump the contextual bias due to it being in some sense conveyed by both the anchor and the tag (via an elliptical high negation question). As for high negation questions in 'suggestion' scenarios (i.e. (17)), the lack of any countervailing evidence in the context would trivially result in the original bias being the stronger of the two, and thus the only option for projected bias to align with.[6]

We close this section by commenting on other potential bias flavors. Notice that all three flavors of bias distinguished above (original, contextual, projected) are 'epistemic' in some broad sense. That is, they convey information about beliefs, evidence, or commitments. Huddleston and Pullum (2002, p.879) suggest that non-epistemic flavors of bias exist as well, including 'deontic' bias (e.g., *Aren't you ashamed of yourselves?*) and 'bouletic' bias (e.g., *Can I have some more ice cream?*). Needless to say, non-epistemic biases could potentially open up new and exciting avenues for future research. However, at this point it is unclear how systematic such biases are or whether they are associated with specific question forms, so we put these aside.

3 Bias features

We have argued that question bias comes in different (epistemic) flavors, i.e., as original, contextual, or projected, where the latter two tend to align in direction. Whatever the particular bias flavor,

[6]Note that the example in (17) also indicates that the negative projected bias in high negation questions, demonstrated in (13), is optional in the sense that it can be overridden by contextual factors.

question biases have several features, such as direction, strength, and optionality. This section discusses these features, as exhibited by original bias.

3.1 Direction

Perhaps the most salient feature of original bias is its direction, which can be positive (agreeing with the question prejacent) or negative (disagreeing with the question prejacent). The direction of original bias is usually explicated by the utterance context, as in (11) above. However, intuitions about this feature may get blurry in the presence of opposing contextual bias. This is due to the fact that these two bias flavors diverge in timing, where original bias is about the speaker's *prior* belief, while contextual bias has *current* relevance. It thus may be unclear whether what is being judged is the speaker's prior or current epistemic state, where the latter would be inconsistent if the contravening contextual evidence had been factored in.

In order to sharpen the intuitions about bias direction, we introduce what we call the **evidential follow-up test**. This diagnostic involves a follow-up utterance which states the evidence on which the speaker's prior belief is based. We use specifically *reportative* evidence, as such evidence is compatible with any degree of certainty (cf. AnderBois 2014) and does not intervene with the bias strength feature, to be discussed in the following subsection. To illustrate, high negation questions are felt to convey a positive original bias, while questions with conversational *really* are felt to convey a negative such bias. The evidential follow-up test confirms these intuitions, as shown in (18)–(19).[7]

(18) Isn't Zelda coming to the party? I mean, she said she would / #wouldn't.

(19) Is Zelda really coming to the party? I mean, she said she wouldn't / #would.

3.2 Strength

It has been recognized that the original biases conventionally associated with different question forms may display contrasts in strength. For example, nuclear tag questions are generally felt to convey a stronger bias than do postnuclear tag questions, a difference that arguably goes back to their specific contours and associated discourse moves (Ladd 1981; Reese and Asher 2010). Despite such initial stabs at the data and proposed explanations, the literature offers no systematic investigation of bias strength. Here we will demonstrate that the intuitions about degrees of bias strength are on the right track, motivating a minimal weak/strong bias distinction.

Consider again the high negation question and the *really*-question in (18)–(19). Through application of the evidential follow-up test, we have established that the original biases conveyed by these two forms differ in direction: positive in the former case, negative in the latter case. As it turns out, these two biases differ in strength as well. That is, in (18) the speaker's prior belief that Zelda is coming to the party can be a mere suspicion, while in (19) the speaker's prior belief that Zelda is not coming to the party must be quite strong.[8] This intuition is further corroborated if we try to match the implied bias strength with an epistemic expression of a comparable strength, such

[7] We intend (18) and (19) to be pronounced with a 'default' intonation, whereby the nuclear pitch accent goes on *party* in the former question and on *really* in the latter question.

[8] The latter is a potential consequence of the intensifier semantics of *really* (Bill and Koev 2022).

as *I thought* (weak bias) vs. *I was sure* (strong bias). The preferences in (20)–(21) are sharp enough to illustrate the viability of this empirical diagnostic, which we dub the **epistemic follow-up test**.

(20) Isn't Zelda coming to the party? Because I thought / ?I was sure she was.

(21) Is Zelda really coming to the party? Because I was sure / ?I thought she wasn't.

3.3 Optionality

Original bias, like any other bias flavor, is just an inference with certain properties. This raises the issue of whether this inference arises obligatorily with any given biased questions or not. Previous literature has pointed out that original bias is sometimes optional. For example, the original bias associated with verum questions is known to disappear in certain contexts (Gutzmann et al. 2020; Bill and Koev 2021; Goodhue 2022). At the same time, the original bias conveyed by many other question forms—incl. high negation or accented low negation questions, questions with conversational *really*, and various tag questions—turns out to be obligatory (Roelofsen et al. 2012; Domaneschi et al. 2017; Bill and Koev 2023a). This subsection outlines two diagnostics for distinguishing between optional and obligatory bias.

On the face of it, the verum question in (22) and the *really*-question in (23) seem to contribute the same original bias, both signaling the speaker's prior belief that Mary did not join the team.

(22) DID Mary join the team?

(23) Did Mary really join the team?

However, when placed in a context which provides evidence both for and against the question prejacent, (22) fits in but loses its original bias, whereas (23) becomes degraded. This is shown in (24)–(25), which illustrate what may be called the **conflicted context test**. It suggests that verum questions are optionally biased while *really*-questions are obligatorily biased.

(24) DID Mary join the team? Because I had no idea what would happen, but some say she did, others say she didn't.

(25) ?Did Mary really join the team? Because I had no idea what would happen, but some say she did, others say she didn't.

In addition, Sadock (1971) noticed that markers like *by any chance* select for neutral questions, thus excluding biased questions or plain assertions. The effect of Sadock's ***by any chance* test** is illustrated in (26) and lends additional support for the contrast in bias optionality between verum questions and *really*-questions already observed in (24)–(25).

(26) *Dan is having a conversation about tofu with some of his friends.*
Dan: By any chance, DO you like tofu?
Dan: #By any chance, do you really like tofu?

3.4 Theoretical implications

The observed contrasts in direction, strength and optionality of the original bias inference raise several theoretical issues. The most basic one concerns the source of these contrasts. For example, why is original bias optional/obligatory in certain questions only? One simple idea is that optional

biases are due to the context, while obligatory biases are hard-coded in the linguistic form. However, this idea fails to draw a distinction between e.g. positive polar questions, which are typically (though not always) interpreted as unbiased, and verum questions, which are typically (though not always) interpreted as biased. Another issue concerns the granularity of bias strength, i.e., is bias strength a binary notion or does it fall on a spectrum? Finally, our discussion has focused on *original* bias. This begs the question of how the different bias features interact with other flavors of bias. For example, it is conceivable that some flavors of bias vary in strength while others do not. All these issues need to be addressed in future work.

4 Approaches to question bias

There are several approaches to question bias, depending on whether the key factor is taken to be the choice of question prejacent, the effects of a specialized epistemic operator, a discourse-relevant conventional implication, etc. Importantly, these different approaches cover specific data or bias flavor and are not necessarily incompatible. This section critically evaluates three main approaches, with an eye on their predictions regarding bias flavor, direction, strength, and optionality.[9]

4.1 Surface form approach

One approach is based on the idea that, in asking a question, the speaker faces a choice between different surface forms that result in the same semantic partition. Which particular surface form is selected will depend on the speaker's beliefs and desires, which determine the utility (usefulness) of the answers in the sense of decision theory (Savage 1954; Jeffrey 1965; a.o.). In the case of polar questions, the speaker always selects the prejacent with the higher 'utility value'. Following this line, van Rooy and Šafářová (2003) propose that polar questions are subject to the felicity constraint in (27) (see also AnderBois 2019 and Goodhue 2019).

(27) *Felicity Constraint on Polar Questions*

A polar question is felicitous only if the utility value of the pronounced cell (positive or negative) exceeds the utility value of the unpronounced cell. If both cells are pronounced (as in the case of alternative questions), the utility values of the positive and the negative cells are the same.

One major motivation for this constraint comes from the observation that negative and positive polar questions have a different distribution. For example, on a children's hospital flier that advertises a given nutritional product, one would read *Is your child not eating properly?* rather than *Is your child eating properly?*.

van Rooy and Šafářová (2003) discuss three types of polar questions: positive, negative, and alternative. Although each of these questions are assumed to produce a regular partition of the form $\{p, \neg p\}$, (27) imposes different restrictions on the utility values of the two answers. That is, in positive polar questions the positive cell must be ranked higher, in negative polar questions the

[9] We will ignore the approach of Abels (2003) and Guerzoni (2004) to questions like *Can Sue even solve the easiest problem?*, which exploits the lexical presuppositions of *even* to bring the set of answers to a singleton. The reason is that this approach is primarily targeted at rhetorical interpretations, which are not our focus here.

negative cell must ranked higher, and in alternative questions the two cells must be on a par. This is schematized in (28), where UV stands for the utility value function.

(28) a. Positive polar questions: $?p$, where $UV(p) > UV(\neg p)$
b. Negative polar questions: $?\neg p$, where $UV(p) < UV(\neg p)$
c. Alternative questions: $?(p \vee \neg p)$, where $UV(p) = UV(\neg p)$

The utility value of a proposition is a formal construct that amounts to different things, depending on the discourse strategy being pursued by the speaker. One common strategy is trying to find out what the world is like. On this fact-finding strategy, the utility value of a proposition reduces to its informativity (or 'surprisal') that results from learning that proposition. Essentially, informativity is a matter of likelihood: less likely propositions are more informative and more likely propositions are less informative. The conditions in (28) then boil down to the general requirement that the pronounced cell be less likely, or—if both cells are pronounced—that they be equally likely. This is stated in (29), with P standing for a probability function.

(29) a. Positive polar questions: $?p$, where $P(p) < P(\neg p)$
b. Negative polar questions: $?\neg p$, where $P(p) > P(\neg p)$
c. Alternative questions: $?(p \vee \neg p)$, where $P(p) = P(\neg p)$

van Rooy and Šafářová (2003) do not distinguish between different flavors of bias, although one could interpret the different strategies being pursued by the speaker as generating different flavors. For example, their fact-finding strategy predicts the direction of original bias, at least in some cases. That is, according to (29b) negative polar questions prioritize the positive answer and are correctly predicted to convey a positive original bias. Moreover, the prediction of (29c) that alternative questions are unbiased is in line with the literature (Bolinger 1978; Biezma and Rawlins 2012). However, the prediction of (29a) that positive polar questions are negatively biased is clearly incorrect. While such questions may sometimes be associated with a negative original bias, they are also compatible with a positive original bias or no bias at all (Roelofsen et al. 2012; Domaneschi et al. 2017).

As for bias strength, if taken at face value, the account only ever requires a mild preference for one of the two answers. In order to distinguish between weak vs. strong bias, we would need to explicitly state how far apart the likelihoods of the two answers are (cf. Farkas and Roelofsen 2017). One problem is that fine-tuning bias strength like that would amount to a long list of stipulations, as we would have to impose specific pragmatic conditions on individual linguistic forms. At the same time, such an account can capture any degree of bias strength, no matter how fine-grained.

Finally, the surface form approach faces difficulties with the optionality feature. That is, despite the finding that bias is obligatory in several question forms, the account predicts that the bias inference should always be cancelable. This is due to the open-ended nature of the proposed explanation, which allows that the speaker may be pursuing different discourse strategies. In other words, even if the felicity constraint in (27) is assumed to be non-violable, there are always going to be several ways of satisfying it and presumably only some of the pursued strategies will generate bias.

4.2 Epistemic approach

Another approach to question bias is rooted in the idea that certain question partitions reveal an epistemic attitude towards one of the answers (Romero and Han 2004; Romero 2006; Repp 2013; Frana and Rawlins 2019; Goodhue 2019; Silk 2019). Romero and Han (2004) apply this idea to polar questions with a verum accent, high or accented low negation, or conversational *really*. Their main claim is that all these question forms share the meta-conversational operator VERUM (cf. Höhle 1992), which signals certainty that the question prejacent should be added to the Common Ground. In addition to introducing epistemic content, the use of VERUM is constrained by an economy principle, which requires an epistemic conflict between the speaker's prior belief and the contextual evidence. The semantic entry for VERUM (anchored to an agent x) and a slightly simplified version of said economy principle are stated in (30)–(31).

(30) $[\![\text{VERUM}_x]\!] = ForSureCG_x \quad (= \lambda p \lambda w\,.\,\forall w' \in Epi_{x,w}\,\forall w'' \in Conv_{x,w'}\,[p \in CG_{x,w''}]\,)$

(31) *Principle of Economy*
Do not use a meta-conversational move unless necessary to resolve an epistemic conflict.

(32) and (33) demonstrate what this account predicts for polar questions with a verum accent and high negation, where the pronounced cell of the question partition is underlined.[10]

(32) IS Jane coming? (original bias: negative, contextual bias: positive)
a. [$_{\text{CP}}$ Q VERUM$_x$ [$_{\text{TP}}$ Jane coming]]
b. $\{\underline{ForSureCG_x(\lambda w.coming_w(jane))}\,,\,\neg ForSureCG_x(\lambda w.coming_w(jane))\}$

(33) Isn't Jane coming (too)? (original bias: positive, contextual bias: negative)
a. [$_{\text{CP}}$ Q not VERUM$_x$ [$_{\text{TP}}$ Jane coming]]
b. $\{ForSureCG_x(\lambda w.coming_w(jane))\,,\,\underline{\neg ForSureCG_x(\lambda w.coming_w(jane))}\}$

The two denotations amount to the same partition, presenting a choice between being certain and not being certain that the prejacent should be added to the Common Ground. The only difference lies in which cell is being pronounced—a concept that Romero and Han (2004) call the 'intent' of the question—i.e., the positive cell (32) or the negative cell (33). This correctly predicts the opposing directions of original vs. contextual bias, due to the assumption that the speaker always requires evidence *for* the pronounced cell. That is, in (32) the pronounced cell is the positive one, corresponding to being certain that *Jane is coming* should be added to the Common Ground. Since the speaker requires evidence for that cell, given the epistemic conflict mandated by the economy principle in (31), the speaker must be doubting the positive cell (hence the negative original bias) and the contextual evidence must be in favor of it (hence the positive contextual bias). In contrast, in (33) the pronounced cell is the negative one, corresponding to not being certain that *Jane is coming* should be added to the Common Ground. In this latter case, the speaker requires evidence for that negative cell, i.e., she requires evidence against the positive cell. Given the epistemic conflict mandated by the economy principle in (31), the speaker must be believing the positive

[10] (33) illustrates the so-called 'outer' negation reading for high negation questions, which double-checks the positive question alternative (Ladd 1981; Romero and Han 2004; Sailor 2013; Romero et al. 2017; Goodhue 2019; Jeong 2021).

cell (hence the positive original bias) and the contextual evidence must point against it (hence the negative contextual bias).

The epistemic account does not do so well with deriving bias strength, though. For example, in (32) the speaker is asking for evidence for the pronounced positive cell, so she must be doubting the prejacent to some degree. This is compatible with the speaker's having either a weak or a strong bias against the prejacent, although it is unclear how to make the correct choice in a given context. Even worse, in (33) the speaker is requiring evidence for the pronounced negative cell, i.e., that it is not certain that *Jane is coming* should be added to the Common Ground. That is, since the speaker is requiring evidence against adding *Jane is coming* to the Common Ground, she must be strongly biased for that proposition. However, (20) above demonstrated that in fact high negation questions convey a weak original bias.

As for optionality, the initial expectation is that bias should be optional in all of the targeted question forms, due to the reliance on pragmatic reasoning about evidence. This contradicts the finding that most of the targeted question forms convey original bias obligatorily (Bill and Koev 2023b). Romero and Han (2004, ft.1) thus stipulate that the conversational principles involved are non-violable. Although this makes the right predictions in most cases, it still fails to predict attested variability, given that e.g. verum questions are only optionally biased (Gutzmann et al. 2020; Bill and Koev 2021; Goodhue 2022).

4.3 Conventional implication approach

A third approach codes the bias implication directly into the conventional meaning of certain non-canonical questions in the form of an indirect assertion (Reese 2007; Reese and Asher 2010), a projected commitment of a discourse participant (Krifka 2015; Malamud and Stephenson 2015), an evidence-backed belief (Farkas and Roelofsen 2017), a secondary discourse topic (AnderBois 2019), or a contrasting focus alternative (Bill and Koev 2023a,b). Here we illustrate this approach on the account put forward in Bill and Koev (2023a,b). This account covers essentially the same data as that of Romero and Han (2004), which includes polar questions with a verum accent, high or accented low negation, or conversational *really*, as well as reverse-polarity tag questions. However, instead of attributing the biases of all these questions to a specialized epistemic operator, these biases are fundamentally derived from the well-established linguistic category of 'polarity focus', i.e., regular focus marking on some polar operator. The basic idea is that, by marking a question with polarity focus, the speaker points to a contrasting antecedent that entails one of the question partition cells, thus making this cell semantically salient and signaling original bias for it. This is codified in (34).

(34) *Salient Cell Principle*

Do not make semantically salient one of the question partition cells unless your prior belief state supports that cell over the alternative.

Notice that this principle says nothing about the specific mechanism through which such cell salience is achieved. That is, polarity focus is expected to be just one way, among possibly many, of raising the salience of a given partition cell and generating a bias.[11]

[11] For example, a given partition cell may also be made semantically salient by excluding alternatives that clash with the lexical presuppositions of *even* (Abels 2003; Guerzoni 2004).

As an illustration, consider the question in (35), where the accent on negation is a reflex of standard F-marking (Rooth 1985, 1992).

(35) Does Susan NOT do weightlifting? (original bias: positive, weak, obligatory)

a. $[_{\text{CP}}\ \text{Q}\ [_{\text{PolP}}\ [\text{not}]_F\ [_{\text{TP}}\ \text{Susan do weightlifting}]]]_\phi \sim C]$

b. $C = \lambda w\,.\,do_w(susan, weightlifting)$

Through the diagnostics presented in Section 3, it is easily established that the original bias conveyed by such questions is positive, weak, and obligatory. This specific bias profile is derived as follows. First, the bias is positive because the polarity focus on negation points to the contrasting positive alternative, corresponding to *Susan does weightlifting*. Since this alternative is equivalent to the positive partition cell, this cell has been made salient and the principle in (34) mandates that the speaker is biased for it. Moreover, the principle in (34) calls for a mere preference for the salient cell, and so—in the absence of further contextual cues—the original positive bias is predicted to be weak. Finally, due to the presence of focus marking, the process described above will always apply and the bias inference will arise obligatorily.

More generally, the salience-through-focus account of Bill and Koev (2023a,b) makes several good predictions about original bias. For one, because of the contrastive focus interpretation, it captures the fact that the polarity of such bias is always opposite to that of the question prejacent: negative in positive questions, positive in negative questions. The account also correctly predicts that original bias in questions marked with polarity focus is weak by default, given that the principle in (34) establishes a mere preference. Cases in which such bias is consistently strong, as in questions with conversational *really*, are explained by the lexical semantics of the polar operator in focus (here through the intensifier semantics of *really*). The account also captures the observation that, once polarity focus is present, the bias is obligatory (verum-accented questions are claimed to be optionally biased as they may, but need not, be focus marked). In spite of all that, the account is limited in scope to original bias.[12] It makes no predictions as to how the notion of salience may play into deriving the properties of contextual and projected bias.

5 Conclusion and outlook

We distinguished between three different flavors of question bias, i.e., original, contextual, and projected. Zeroing in on original bias, we employed various empirical diagnostics to additionally characterize a given bias inference as positive or negative, weak or strong, optional or obligatory. Finally, we reviewed three major approaches to question bias, spelling out their predictions regarding flavor and the additional features.

The main issue raised by our discussion is whether question bias is a uniform phenomenon or, alternatively, whether it is composed of a number of semantically or pragmatically unrelated effects. While it would be premature to stake out a definite position at this point, we will conclude this paper by emphasizing certain general tendencies that point towards linguistic uniformity. What we suggest is that question bias is primarily a matter of answer salience, where different kinds of salience correspond to different bias flavors. Specifically, Büring and Gunlogson (2000) note that the overtly expressed proposition in polar questions cannot contradict the contextual evidence and

[12] Other accounts within the conventional implication approach are primarily focused on a single bias flavor as well.

must be compatible with it. The choice of question prejacent then constitutes a kind of 'surface' salience that conveys contextual bias. Additionally, the speaker may choose to raise the salience of a given answer by triggering a certain implication through polarity focus marking, presupposition triggers, etc. (Abels 2003; Guerzoni 2004; Bill and Koev 2023a,b). This kind of 'semantic' salience will amount to original bias and diverge in direction with the contextual bias expressed by the question prejacent (Romero and Han 2004; Roelofsen et al. 2012; Domaneschi et al. 2017). Now, given these two contrasting kinds of answer salience and corresponding biases, projected bias will have to follow more closely one of them, where the default choice is alignment with contextual bias due to some general constraint (see Section 2). We leave the substantiation of these speculative remarks to future work.

References

Abels, K. (2003). Who gives a damn about minimizers in questions? In *Proceedings of Semantics and Linguistic Theory 13 (SALT 13)*, pp. 1–18.

AnderBois, S. (2014). The semantics of sluicing: beyond truth-conditions. *Language 90*(4), 887–926.

AnderBois, S. (2019). Negation, alternatives, and negative polar questions in American English. In K. von Heusinger, V. E. O. Gaspar, and M. Zimmermann (Eds.), *Questions in Discourse, Volume 1: Semantics*, pp. 118–171. Brill.

Biezma, M. and K. Rawlins (2012). Responding to alternative and polar questions. *Linguistics and Philosophy 35*, 361–406.

Bill, C. and T. Koev (2021). Verum accent is verum, but not always focus. In *Proceedings of Linguistic Society of America 6(1)*, pp. 188–202.

Bill, C. and T. Koev (2022). *Really*: Ambiguity and question bias. In *Proceedings of Sinn und Bedeutung 26*.

Bill, C. and T. Koev (2023a, March). Bias in tag questions. To appear in *Perspectives on Biased Questions*. Available at https://lingbuzz.net/lingbuzz/007302.

Bill, C. and T. Koev (2023b, April). Question bias from polarity focus. Under review. Available at https://ling.auf.net/lingbuzz/007301.

Bolinger, D. (1978). Yes–no questions are not alternative questions. In H. Hiż (Ed.), *Questions*, pp. 87–105. Dordrecht: Springer.

Borkin, A. (1971). Polarity items in questions. In *Proceedings of Chicago Linguistic Society 7 (CLS 7)*, pp. 53–62.

Büring, D. and C. Gunlogson (2000). Aren't positive and negative polar questions the same? Technical report, University of California, Santa Cruz.

Caponigro, I. and J. Sprouse (2007). Rhetorical questions as questions. In *Proceedings of Sinn und Bedeutung 11 (SuB 11)*, pp. 121–133.

Ciardelli, I., J. Groenendijk, and F. Roelofsen (2019). *Inquisitive Semantics*. Oxford: Oxford University Press.

Domaneschi, F., M. Romero, and B. Braun (2017). Bias in polar questions: Evidence from english and german production experiments. *Glossa 23*, 1–28.

Farkas, D. F. and F. Roelofsen (2017). Division of labor in the interpretation of declaratives and interrogatives. *Journal of Semantics 34*, 237–289.

Frana, I. and K. Rawlins (2019). Attitudes in discourse: Italian polar questions and the particle *mica*. *Semantics & Pragmatics 12*(16), 1–48.

Goodhue, D. (2019). High negation questions and epistemic bias. In *Proceedings of Sinn und Bedeutung 23(1)*, pp. 469–485.

Goodhue, D. (2022). All focus is contrastive: On polarity (verum) focus, answer focus, contrastive focus and givenness. *Journal of Semantics 39*, 117–158.

Goodhue, D. and M. Wagner (2018). Intonation, *yes* and *no*. *Glossa: A Journal of General Linguistics 3*(1), 5.

Groenendijk, J. and M. Stokhof (1984). On the semantics of questions and the pragmatics of answers. In F. Landman and F. Veltman (Eds.), *Varieties of Formal Semantics*, pp. 143–170. Dordrecht: Foris.

Guerzoni, E. (2003). *Why even ask?: On the pragmatics of questions and the semantics of answers*. Ph. D. thesis, Massachusetts Institute of Technology.

Guerzoni, E. (2004). *Even*-NPIs in yes/no questions. *Natural Language Semantics 12*, 319–343.

Gunlogson, C. (2008). A question of commitment. *Belgian Journal of Linguistics 22*(1), 101–136.

Gutzmann, D., K. Hartmann, and L. Matthewson (2020). Verum focus is verum, not focus: Cross-linguistic evidence. *Glossa: A Journal of General Linguistics 5*(1), 51.

Hamblin, C. L. (1973). Questions in Montague English. *Foundations of Language 10*, 41–53.

Han, C.-h. (2002). Interpreting interrogatives as rhetorical questions. *Lingua 112*(3), 201–229.

Heim, I. (1984). A note on negative polarity and de-ness. In C. Jones and P. Sells (Eds.), *Proceedings of North East Linguistics Society 14 (NELS 14)*, pp. 98–107.

Höhle, T. N. (1992). Über Verum-Fokus im Deutschen. In J. Jacobs (Ed.), *Informationsstruktur und Grammatik*, pp. 112–141. Opladen: Westdeutscher Verlag.

Holmberg, A. (2013). The syntax of answers to polar questions in english and swedish. *Lingua 128*, 31–50. SI: Polarity emphasis: distribution and locus of licensing.

Huddleston, R. and G. Pullum (2002). *The Cambridge Grammar of the English Language*. Cambridge, UK: Cambridge University Press.

Jeffrey, R. C. (1965). Ethics and the logic of decision. *The Journal of Philosophy 62*(19), 528–539.

Jeong, S. (2021, 10). Deriving Dual Dimensions of Bias: Preposed Negation Questions with EVEN. *Journal of Semantics 38*(1), 49–94.

Koev, T. (2019). Strong beliefs, weak commitments. In *Proceedings of Sinn und Bedeutung 23, vol. 2*, pp. 1–18.

Kramer, R. and K. Rawlins (2011). Polarity particles: An ellipsis account. In *Proceedings of the North East Linguistics Society 39*, pp. 479–492.

Krifka, M. (1995). The semantics and pragmatics of polarity items. *Linguistic Analysis 25*, 1–49.

Krifka, M. (2013). Response particles as propositional anaphors. In T. Snider (Ed.), *Proceedings of Semantics and Linguistic Theory (SALT) 23*, pp. 1–18.

Krifka, M. (2015). Bias in commitment space semantics: Declarative questions, negated questions, and question tags. In *Proceedings of Semantics and Linguistic Theory 25*, pp. 328–345.

Ladd, D. R. (1981). A first look at the semantics and pragmatics of negative questions and tag questions. In *Proceedings of Chicago Linguistic Society 17*, pp. 164–171.

Ladusaw, W. A. (1980). *Polarity Sensitivity as Inherent Scope Relations*. Ph.D. dissertation, University of Texas, Austin.

Malamud, S. A. and T. Stephenson (2015). Three ways to avoid commitments: Declarative force modifiers in the conversational scoreboard. *Journal of Semantics 32*(2), 275–311.

Rando, E. (1980). Intonation in discourse. In *The melody of language*, pp. 243–278. Baltimore: University Park Press.

Reese, B. and N. Asher (2010). Biased questions, intonation and discourse. In M. Zimmermann and C. Féry (Eds.), *Information Structure*, pp. 139–173. Oxford University Press.

Reese, B. J. (2007). *Bias in Questions*. Ph.D. dissertation, University of Texas at Austin.

Repp, S. (2013). Common ground management: Modal particles, illocutionary negation and verum. In D. Gutzmann and H.-M. Gärtner (Eds.), *Beyond Expressives: Explorations in Use-conditional Meaning*, pp. 231–274. Leiden: Brill.

Roelofsen, F. and D. F. Farkas (2015). Polarity particle responses as a window onto the interpretation of questions and assertions. *Language 91*(2), 359–414.

Roelofsen, F., N. Venhuizen, and G. W. Sassoon (2012). Positive and negative questions in discourse. In *Proceedings of Sinn und Bedeutung 17*, pp. 455–472.

Romero, M. (2006). Biased yes/no questions: The role of verum. *Sprache und Datenverarbeitung 30*(1), 9–24.

Romero, M., A. Arnhold, B. Braun, and F. Domaneschi (2017). Negative polar question types in english. In *Proceedings of the North East Linguistics Society 47*, Volume 3, pp. 35–48.

Romero, M. and C.-H. Han (2004). On negative *yes/no* questions. *Linguistics and Philosophy 27*, 609–658.

Rooth, M. (1992). A theory of focus interpretation. *Linguistics and Philosophy 1*, 75–116.

Rooth, M. E. (1985). *Association with Focus*. Ph.D. dissertation, University of Massachusetts at Amherst.

Sadock, J. (1971). Queclaratives. In *Proceedings of Chicago Linguistic Society 7*, pp. 223–231.

Sailor, C. (2013, January 3). Questionable negation. Technical report, LSA Annual Meeting.

Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley.

Silk, A. (2019). Expectation biases and context management with negative polar questions. *Journal of Philosophical Logic*.

van Rooy, R. (2003). Questioning to resolve decision problems. *Linguistics and Philosophy 26*(6), 727–763.

van Rooy, R. and M. Šafářová (2003). On polar questions. In *Proceedings of Semantics and Linguistic Theory 13*, pp. 292–309.