IJPLH
ISSN: 3082-415X (Online)

International Journal of Philosophy, Linguistics, and Humanities

Journal Homepage: https://ijmshe.com/index.php/ijplh



Research Article

Undirected Need Not Be Uninformative: Symmetric Conflicting Evidence and the Measurement of QUD Informativity

Muteb Alqarni [1]*

[1] Professor, King Khalid University, Abha, Saudi Arabia

https://doi.org/

Article history:

Received: 15 October 2026
Revised: 28 October 2026
Accepted: 06 November 2026

© The Author(s) 2026. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/

Abstract

Feinmann (2026) argues that linguistic relevance cannot be reduced to informativity relative to the Question Under Discussion (QUD), because a response can be uninformative to the QUD yet be perceived as relevant. His notion of informativity is that of Probabilistic QUD Relevance: a non-trivial directional effect on the QUD-cell probabilities. This article accepts his judgments and conclusion under that gloss, and asks what the cases show about accounts making informativity necessary for relevance. A directional criterion returns the same verdict for two different situations: one in which evidence on the question enters on both sides and offsets, and one in which none enters. An exploratory study (N = 27; six polar-question scenarios; three counterbalanced lists) compared explicit conjunctive cancellation, one-sided evidence from the same sources, and a topically matched Inert control, eliciting probability and sureness judgments before and after each response plus relevance and informativeness ratings. Conflict and Inert responses both produced very little directional probability change, yet were sharply distinguished on a seven-point self-report measure of perceived informativeness (M = 3.15 vs. 1.04; dz = 1.59). Feinmann and Ramotowska's (2026) relevance pattern is reproduced. The result motivates, but does not establish, a richer conception of QUD-directed epistemic impact.

Keywords: relevance; Question Under Discussion; informativity; conflicting evidence; perceived informativeness; experimental pragmatics

International Journal of Philosophy, Linguistics, and Humanities
E-ISSN: 3082-415X

1. Introduction

In Gricean and Neo-Gricean pragmatics the Maxim of Relation (Grice, 1975) is standardly cashed out in terms of the Question Under Discussion. A contribution counts as relevant to a QUD just in case it is informative relative to that QUD: in non-probabilistic implementations by eliminating at least one cell of the partition (Groenendijk & Stokhof, 1984; Roberts, 2012), and in probabilistic implementations by shifting the probabilities assigned to those cells (Büring, 2003; van Rooy, 2003). Feinmann (2026) labels the shared core Informativity-based QUD Relevance and argues that it fails. The QUD tradition more broadly treats discourse as organized around explicit or implicit questions and an evolving conversational context (Ginzburg, 2012; Stalnaker, 2002; van Kuppevelt, 1995).

The argument centers on a class of responses, illustrated by the following example (Feinmann, 2026, p. 6):

(1) A: Is Mary coming to the party tonight?
B: Mary's parents gave me conflicting reports.

Suppose the parents are equally reliable and their reports conditionally independent given whether Mary attends. The testimonies then carry equal and opposite evidential force, and conditioning on B's assertion leaves each QUD cell exactly where it was. Yet B's response cannot be challenged with "And what does that have to do with anything?", whereas an off-topic response can. Feinmann concludes that relevance cannot be identified with QUD informativity.

The judgments seem to me correct and I do not dispute them; nor do I dispute that the conclusion follows under Feinmann's own gloss, on which being uninformative to the QUD is having no detectable effect on the QUD cells, a gloss he states in the sentence that states the conclusion and alongside which he says he has no quarrel with Info-QUD-R as such (Feinmann, 2026, pp. 6–7). What I question is narrower: what the cases show about the wider family of accounts that make informativity necessary for relevance.

They establish relevance without directional change in the QUD-cell probabilities. Whether they establish relevance without any plausible form of QUD-directed informativity turns on a distinction a directional criterion cannot draw. Two quite different situations leave P(q) where it was: in one, evidence bearing on the question enters the common ground on both sides and offsets; in the other, none enters at all.

Call these cancellation and independence. Feinmann's targets are cases of the first, his off-topic controls of the second, and the distinction is roughly the

one Keynes (1921) drew between the balance of the evidence and its weight (see Joyce, 2005).

This is a claim about the reach of the cases, not a claim that Feinmann's conclusion is false: if the two situations were alike for language users, nothing would follow from the distinction. Whether they are alike carries a differential prediction. If conflicting evidence is treated as it would be if nothing bearing on the question had entered, it should pattern with inert information on every index of perceived informativeness; if it is merely undirected, it should pattern with inert information on directional probability change while separating from it on a graded measure.

The gap this study addresses is a measurement gap. What previous work establishes is that relevance judgments come apart from directional QUD-informativity: responses producing little or no change in the probability of an answer are nonetheless judged relevant (Feinmann, 2026; Feinmann & Ramotowska, 2026; Warstadt & Agha, 2022), and representations that add a higher-order component to the listener's state capture part of what a first-order measure misses (Warstadt, Agha, & Franke, 2026). What remains unresolved is whether a response coded uninformative on a directional measure is also perceived as informationally on a par with a response that bears on nothing. Feinmann and Ramotowska's informativeness measure is a three-way direction report, and their target replies were compared with off-topic rather than topically matched controls, so their data cannot answer that question, which is not one they set out to answer. The present study adds a graded rating of perceived informativeness and a topically matched Inert control. Its contribution is to that measurement question; the judgments themselves are not in dispute.

This article reports a small exploratory study designed to test that prediction. Three response conditions were compared within participants across six everyday scenarios with polar questions: Conflict (two equally well-placed sources reporting in opposite directions), Directional (one of the same sources reporting in one direction), and Inert (a topically matched response about the same people and situation that carries no evidence bearing on the question).

The study's status should be clear at the outset. N = 27, the sample was recruited by convenience, each participant contributed two trials per condition, and several features of the data would be unacceptable in a confirmatory design. It is a proof of concept for a theoretical point, not experimental validation of it.

The closest experimental work is Feinmann and Ramotowska (2026), who tested these cases with one hundred participants and twelve items and report that disagreement and uncertainty targets were judged uninformative at rates not

International Journal of Philosophy, Linguistics, and Humanities
E-ISSN: 3082-415X

reliably different from off-topic controls, while being judged relevant on 95% of trials against 3%. Their relevance pattern is reproduced here; their informativeness result is a null, which a study of this size cannot be said to replicate. Section 3.2.2 sets the comparison out.

The most robust finding is the distinction between Conflict and Inert responses in perceived informativeness: both produced very little directional probability change, and were nonetheless sharply distinguished on a graded self-report measure. Sections 3.2.1 and 3.2.4 are explicit about how far that falls short of a claim about information, weight, or higher-order belief. Two further results are exploratory only: sureness increased more in Conflict than in Inert, but fragilely; and conflicting evidence moved non-neutral priors toward the midpoint wherever the prior permitted it.

One of Feinmann's findings is reproduced here and should be preserved. Conflicting responses were rated exactly as relevant as one-sided evidence from the same sources, while being rated far less informative. Relevance and perceived informativeness do come apart, and they come apart in the direction he describes.

Section 1.1 states the argument and the question this article raises about it. Section 1.2 draws the distinction the question turns on and derives the predictions. Sections 2 and 3.1 report the study. Section 3.2 discusses what survives, and Section 5 sets out the limitations.

1.1. Theoretical Background

1.1.1. Informativity-Based QUD Relevance

On the standard non-probabilistic account, a declarative sentence φ is relevant to a QUD Q iff φ gives at least a partial answer to Q, that is, iff φ is not a contextual contradiction and its truth conditions fail to intersect at least one cell of Q (Groenendijk & Stokhof, 1984; Roberts, 2012). Feinmann calls this Standard QUD Relevance, and it is well known to be too strong: as Simons et al. (2010, pp. 316–317) observe, responses that raise or lower the probability of a cell without eliminating any are intuitively relevant and wrongly excluded. Related work on information structure likewise models interpretation against a structured common ground and local communicative needs (Krifka, 2008).

The probabilistic repair replaces cell elimination with probability shift. Let P be a probability distribution over the set of possible worlds, Q a partition of that set, and φ a declarative sentence.

Then φ is relevant to Q under P iff $P(\varphi) > 0$ and there is a cell $q \in Q$ such that $P(q) \neq P(q \mid \varphi)$ (Feinmann, 2026, p. 3), a definition Feinmann calls Probabilistic QUD Relevance. The two implementations share a core he labels

International Journal of Philosophy, Linguistics, and Humanities
E-ISSN: 3082-415X

Informativity-based QUD Relevance: a sentence is relevant iff, given the epistemic state, it is informative to the QUD.

1.1.2. The Cancellation Cases

Feinmann's central data are responses compatible with every cell of the QUD that also, under a stated assumption about the prior, leave the cell probabilities exactly as they were. The conjunctive version is his (3)B1 and the non-conjunctive version his (4)B1:

(2) A: Is Mary coming to the party tonight?
B: Mary's father told me she's coming, but Mary's mother said the opposite.
(3) A: Is Mary coming to the party tonight?
B: Mary's parents gave me conflicting reports.

These are contrasted with an off-topic control. The contrast is clear and I take it at face value: (2)B and (3)B are related to the question in a way the control is not, and Feinmann's diagnostic (the felicity of the follow-up "And what does that have to do with anything?") separates them.

The stipulation that generates the theoretical problem is symmetry. Feinmann assumes that there is no reason to think one parent more likely than the other to report truthfully, which he glosses as the assumption that each conjunct shifts the odds in opposite directions with the same strength (Feinmann, 2026, p. 5). Under that assumption the derivation is immediate. Write q for the proposition that Mary comes and let each parent report correctly with probability r. Then $P(\text{father says yes, mother says no} \mid q) = r(1 - r) = P(\text{father says yes, mother says no} \mid \neg q)$, so the likelihood ratio is 1, and conditioning on the conjunction returns the prior unchanged for any prior whatsoever. The same holds for (3)B, which denotes the union of the two disagreement cells. Feinmann's arithmetic is correct.

Feinmann gives the two versions different roles in his own dialectic. He notes that (2)B might be handled by assessing relevance incrementally, clause by clause, since the first conjunct raises the probability of the YES cell and the second lowers it; and he observes that the atomic paraphrase (3)B is not open to that reply, since there is no principled way to decompose (3)B into sub-clausal units that are independently informative relative to the QUD (Feinmann, 2026, pp. 5–6). The atomic version is therefore the load-bearing case. This matters for Section 2, where only the conjunctive form was tested.

1.1.3. The Question This Raises

From this Feinmann concludes that the responses are uninformative to the QUD, and that a sentence may be uninformative to the QUD and still be related

to it, so that relevance in the sense of being related cannot be assimilated to relevance in the sense of being informative.

It matters to represent that claim exactly, because it is more careful than it is sometimes taken to be. Feinmann glosses his own use of the term in the sentence that states the conclusion: a sentence may not be informative to the QUD, which is to say it may have no detectable effect on the QUD cells, and may still be related to the QUD (Feinmann, 2026, p. 6). He adds that he has no qualms with Info-QUD-R as such, calling it a perfectly natural notion with empirical reflexes of its own, his point being that it rules out as irrelevant contributions that do not give rise to a judgment of irrelevance (Feinmann, 2026, p. 7). He therefore makes no claim that these responses carry no information in any unrestricted sense, and none is attributed to him here. His notion of informativity is the one PQR defines, a non-trivial directional effect on the probabilities of the QUD cells, and his argument is that it does not exhaust perceived relatedness.

The question this paper raises is therefore not whether his conclusion follows from his premises. Under his own gloss it does. The question is what the cases show about the space of accounts on which informativity is necessary for relevance. They establish relevance without directional QUD-cell probability change. They do not by themselves establish relevance without every plausible form of QUD-directed informativity, because a criterion that reads only the movement of P(q) returns the same verdict for two situations that differ in whether anything bearing on the question has entered the common ground at all. Whether language users treat those situations alike is an open empirical question, and if they do not, what the cases refute is the directional implementation rather than the family.

Feinmann identifies the historical precedent for his data in exactly the place where this reading is available. He quotes Keynes (1921, p. 72) on the thought that evidence part of which is favorable and part unfavorable should count as relevant even if, taken as a whole, it leaves the probability unchanged, and reads this as Keynes dismissing the probabilistic characterization that Carnap (1950) later revived. The more usual reading is that Keynes was pointing to a second dimension of an evidential state rather than abandoning the first. That dimension is the subject of the next section.

1.2. Cancellation and Independence

1.2.1. Balance and Weight

An epistemic state relative to a binary question is not exhausted by the single number P(q). Keynes (1921) distinguished the balance of the evidence,

International Journal of Philosophy, Linguistics, and Humanities
E-ISSN: 3082-415X

which the probability records, from its weight, the amount of evidence bearing on the question. Joyce (2005) puts the distinction precisely: probabilities record how decisively the available evidence tells for or against a hypothesis, which is a different matter from how much evidence there is, the latter reflected in the stability of the credence under further evidence rather than in its value. Two agents can assign the same probability while standing in very different evidential relations to the proposition: one on a single equivocal report, the other on a large and internally balanced body of testimony.

1.2.2. Two Configurations

Probabilistic QUD Relevance, as formulated in Section 1.1.1, asks one question: does conditioning on φ move P(q)? Two quite different situations answer it in the negative.

(4) a. Cancellation: φ puts evidence bearing on q into the common ground on both sides, and the contributions offset.

b. Independence: φ puts no evidence bearing on q into the common ground at all.

These are descriptive labels for two configurations, not a measure. I attach no numbers to them, and nothing below depends on a formalization; the terms balance and weight are used in this paper in Keynes's and Joyce's informal sense and are not given an implementation here.

Feinmann's symmetry stipulation is a stipulation about the relation between two contributions, namely that they offset. It is sufficient to force the change in P(q) to zero. It is not sufficient to make the case one in which nothing bearing on q has entered, and nothing in the examples requires that stronger claim. The conjunctive version (2)B makes this plain: it explicitly reports two testimonies about whether Mary is coming, one on each side.

Feinmann's off-topic control, in which B says that Jeff Bezos hates Bill Gates, is a case of (4b). Nothing about whether one technology executive dislikes another bears on whether Mary attends a party. The conflicting-reports responses are cases of (4a). A criterion that reads only the movement of P(q) returns the same verdict for both, and that is the whole of the observation.

1.2.3. First-Order and Higher-Order Epistemic Change

The distinction has a recent implementation worth invoking for one limited purpose. Warstadt, Agha, and Franke (2026) represent a listener's state with respect to a polar QUD as a first-order probability of a "yes" answer together with a second-order distribution over it, recovered from ratings of commitment to one's own probability judgment; their motivating example leaves the first-order estimate exactly where it was while reducing higher-order

International Journal of Philosophy, Linguistics, and Humanities
E-ISSN: 3082-415X

uncertainty, and in their experiments measures of higher-order belief change contribute to relevance ratings over and above first-order change. Warstadt and Agha (2022) had already found responses judged relevant that produced little or no prior-to-posterior probability change, which they attribute to the assumption that belief states can be represented as a single distribution over the QUD alternatives.

The purpose is only to establish that the distinction between first-order and higher-order epistemic change is independently motivated and empirically tractable. It is no part of my argument that their measures would register Feinmann's configuration, which their published items do not settle. Since a well-motivated representation exists on which zero first-order change is compatible with non-zero epistemic change, a demonstration of the former is not by itself a demonstration of the latter.

1.2.4. What This Does and Does Not Establish

None of this is a counterexample to Feinmann. Two things would close the gap between what the cases show and what a wider reading of them would assert: an argument that nothing beyond the movement of P(q) counts as informativity relative to a QUD; or a demonstration that language users treat cancellation and independence alike. The second possibility is empirically testable because it generates distinct predictions. The remainder of this paper states that prediction and reports a first attempt to test it.

1.2.5. Predictions

Consider three response types to a single polar QUD, holding the protagonists and the situation constant: Conflict, in which two equally well-placed sources report in opposite directions; Directional, in which one of the same sources reports in one direction; and Inert, a response about the same people and situation carrying no evidence bearing on either answer. These are the experimental realizations of cancellation, ordinary one-sided evidence, and independence. Four predictions follow, of which only the second discriminates between the two accounts.

First, directional probability change should be large in Directional and small in both Conflict and Inert.

This is common ground, and functions here as a manipulation check on the symmetry of the Conflict materials rather than as a test.

Second, and critically, perceived informational content should distinguish Conflict from Inert. If Conflict responses are genuinely uninformative relative to the QUD, they should pattern with Inert responses on an informativeness measure.

International Journal of Philosophy, Linguistics, and Humanities
E-ISSN: 3082-415X

If they are merely undirected, that is, if evidence bearing on the question has entered without net directional effect, they should be rated as substantially more informative than Inert responses, while still less informative than Directional responses. This is a between-condition comparison, with Inert rather than zero supplying the baseline, since any response at all may lift a rating slightly; and it concerns a graded measure, since a judgment of whether the probability of an answer went up, down, or neither does not discriminate the two accounts.

Third, relevance ratings should be high and approximately equal for Conflict and Directional, and low for Inert. This is Feinmann's prediction and also mine, included because his dissociation deserves an independent test and because nothing here claims that relevance tracks informativity proportionally.

Fourth, and treated throughout as secondary, sureness about one's own probability estimate should increase more in Conflict than in Inert, since a listener who has learned what two well-placed sources reported is in a different position from one who has learned nothing bearing on the question. This is the most direct test available here of the claim that something bearing on the question has entered, but it is also the most vulnerable to measurement problems, and in this study it turned out to be exactly that.

2. Materials and Methods

2.1. Research Design

The study used a within-participants design with three response conditions (Conflict, Directional, Inert) crossed with six scenarios over three counterbalanced lists. Each participant saw all six scenarios, two in each condition; each scenario appeared in each condition across the lists. A two-message procedure elicited prior and posterior measures: participants first gave a probability estimate and a sureness rating for each context and question, and only after replying received the responses and rated them again on both scales plus relevance and informativeness. Because each scenario appeared in each condition across the three lists, scenario content was balanced across conditions at the group level; within any one list, however, scenario and condition were confounded, so the list-wise contrasts reported in Section 3.1.3 are the only check on scenario-specific effects the design affords (Section 5).

2.2. Materials

Six everyday scenarios were used, involving a workplace potluck, a delayed flight, a job offer, a late package, an outdoor concert, and whether a friend would bring his dog on a trip.

International Journal of Philosophy, Linguistics, and Humanities
E-ISSN: 3082-415X

Each context named two people and stated that they were equally well placed to know the answer, so that the symmetry assumed in Feinmann's analysis was part of the stimulus rather than imposed by the analyst. The Conflict response then gave one report from each source, in parallel wording and with opposite polarity. Conditional independence of the two reports was assumed, as in Feinmann's analysis, not tested, and no manipulation check of perceived source reliability was collected (Section 2.4 lists the complete instrument). Participants were expected to treat the two sources as equally reliable, but perceived symmetry is therefore an assumption of the design rather than a measured property of it; the movement toward the midpoint reported in Section 3.1.6 bears on it. The context did not state that either person had in fact obtained information, so that the Inert and Directional responses would not implicate that information was being withheld. The potluck scenario illustrates the set:

(5) Context: Dana might come to her office potluck on Friday. Her coworkers Ben and Carla both know her well and are equally likely to know her Friday plans.
A: Is Dana coming to the potluck on Friday?
Conflict: Ben told me she's coming, and Carla told me she isn't.
Directional: I only heard from Ben, and he told me she's coming.
Inert: Ben and Carla both joined Dana's team the same month she did.

Three features matter. The Inert responses are topically connected, concerning the same protagonists and situation, so that the contrast with Conflict isolates evidential bearing rather than topic. The Directional responses reuse one of the same two sources. And within each list one Conflict item stated the yes-supporting report first and one the no-supporting report first, while one Directional item pointed toward "yes" and one toward "no," counterbalancing order and polarity within participants.

One feature of the materials must be stated plainly, because it bounds what the study can show. All six Conflict responses were of the conjunctive form, "X reported p, and Y reported not-p." None used the atomic paraphrase, "Her parents gave me conflicting reports." As noted in Section 1.1.2, Feinmann raises an incremental, clause-by-clause reading for exactly the conjunctive form and observes that the atomic paraphrase is not open to it, which makes the atomic version his load-bearing case. The experiment therefore tests the weaker of his two cases. Section 3.2.4 sets out what follows.

2.3. Respondents and Sampling Procedure

Twenty-seven adult native speakers of American English took part, recruited by convenience through social-media direct messages and assigned to lists in order of reply (nine per list).

International Journal of Philosophy, Linguistics, and Humanities
E-ISSN: 3082-415X

No compensation was offered. Participants were told only that this was a short language study, were never told its purpose, and were never shown condition labels. Responses were recorded without usernames.

Approximately thirty participants were planned; twenty-seven were available and took part. No formal power analysis was conducted; the study is exploratory, and its sample size is treated as a limitation in Section 5. No demographic variables beyond the eligibility characteristics just stated (adult native speakers of American English) were collected.

Research ethics. Participation was voluntary. Recipients of the invitation indicated their willingness to take part by replying to it and completing the research messages; no separate written consent form was obtained. Participants were free to decline before starting, to discontinue, or to have their responses withdrawn after completion. Participants were told that the study was a short language study; its specific purpose was not disclosed to them, and no debriefing took place afterwards. Only the responses were extracted for analysis; participant names and usernames were not included in the analytical data. The study was not submitted to an institutional ethics committee, and no formal exemption determination was obtained; the author's understanding was that the institution did not require prior ethics approval for non-identifying humanities research of this kind.

2.4. Research Instruments and Data Analysis

Probability was elicited as "How likely is it that the answer is YES?" on a 0–100 scale, and sureness as "How sure are you that this number is a good estimate?" on a 1–7 scale anchored at 1 = "not at all sure; my estimate could easily be quite different" and 7 = "completely sure." After the response, participants rated relevance ("How relevant was the new information to the question?") and informativeness ("How informative was it for answering the question?"), both 1–7.

The informativeness item was worded "for answering the question" so as to direct participants to the response's bearing on the question asked rather than to its general interest. It is nevertheless a single self-report rating, and perceived informativeness in this sense may overlap with neighbouring constructs: perceived helpfulness, topical relevance, perceived diagnosticity, and whether the response is about the proposition at issue.

The study does not separate these. The ratings are therefore treated throughout as perceived informativeness, not as a demonstration of information gain or evidential weight; Section 3.2.1 sets out the alternative readings this leaves open, and Section 4 the multiple-indicator measurement a follow-up study would need.

International Journal of Philosophy, Linguistics, and Humanities
E-ISSN: 3082-415X

Analyses use participant means, each participant contributing one value per condition. This is deliberately conservative: as Section 5 documents, most participants gave identical answers to their two items within a condition, so the trials per cell are not independent replicates and trial-level models treating them as such are anticonservative. Participant means address the non-independence of a participant's two trials; they do not address item-level variability or the confounding of item with list described in Section 5, which the design cannot resolve. Equivalence testing uses two one-sided tests (Lakens, 2017). Paired mean differences are calculated from unrounded participant-level values and therefore need not equal the differences between the rounded marginal means in Table 1. No participant and no item was excluded from any analysis below.

3. Results

3.1. Results

The design was fully balanced in the realized data: 162 trials, 54 per condition, each scenario appearing nine times in each condition. One participant returned only two numbers per item in the second message, so relevance and informativeness are missing for that participant's six trials; those measures are based on 26 participants and the others on 27. Table 1 gives the descriptive statistics.

Table 1. Descriptive statistics by condition (trial level; 54 trials per condition; relevance and informativeness 52 per condition).

Measure	Conflict	Directional	Inert
Signed Δprobability, M (SD)	−2.22 (7.87)	0.00 (37.47)	−0.09 (0.68)
Absolute Δprobability, M (SD)	2.78 (7.69)	35.37 (11.36)	0.09 (0.68)
Trials with Δprobability exactly 0	75.9%	3.7%	98.1%
Trials with absolute Δprobability ≤ 5	88.9%	3.7%	100%
Δsureness (1–7), M (SD)	+0.87 (2.16)	+2.44 (1.73)	−0.09 (1.01)
Relevance (1–7), M (SD)	6.96 (0.19)	6.96 (0.19)	1.58 (0.50)
Informativeness (1–7), M (SD)	3.15 (1.51)	6.31 (0.54)	1.04 (0.19)

Note. Trial-level descriptive statistics. The inferential tests reported in the text use participant means (one value per condition per participant); see Section 2.4.

3.1.1. Manipulation Check

Mean absolute probability change was 35.37 points in Directional against 2.78 in Conflict, $t(26) = 14.88$, $p < .001$. Of the 54 Directional trials, 51 moved in the direction of the reported evidence; the three exceptions were two trials from a participant who gave identical answers throughout and one apparent misreading. Sureness rose by 2.44 points in Directional against −0.09 in Inert, $t(26) = 8.72$, $p < .001$, $dz = 1.68$.

Participants were reading the materials and updating on them.

3.1.2. Probability Change in the Conflict Condition

Conflict responses produced very little directional change. Three quarters of Conflict trials (75.9%) showed no change at all, and 88.9% a change of five points or less, against 3.7% in Directional. The mean signed change was −2.22 points, 95% CI [−4.41, −0.04]. Two things should be said about this figure, and said together.

First, the change is not zero: a conventional test against zero is significant, $t(26) = -2.09$, $p = .046$. Second, it is small relative to the manipulation: a post hoc two one-sided test against bounds of ±5 points rejects an effect larger than those bounds, $p = .007$, though the same test against ±3 points does not, $p = .235$. Those bounds were chosen after the fact, so the claim carried forward is the descriptive one, that Conflict produced very little directional change, not that it was equivalent to zero. Two facts qualify even that: 63% of participants gave a prior of exactly 50 on every scenario, the one value at which symmetric conflicting evidence cannot move the estimate directionally, and all eleven Conflict trials with a non-neutral prior moved toward 50 (Section 3.1.6). The small mean is in part a property of where the priors sat. The Inert condition, by contrast, really was flat: 98.1% of trials showed no change and the largest in the condition was five points.

3.1.3. Perceived Informativeness

The discriminating prediction was borne out. Conflict responses were rated substantially more informative than Inert responses: 3.15 against 1.04 on the 1–7 scale, a mean paired difference of 2.12, 95% CI [1.58, 2.66], $t(25) = 8.12$, $p < .001$, $dz = 1.59$. Every participant's difference was non-negative, only one participant placed Conflict at the floor of the scale, and the contrast was stable across the three lists (mean differences of 2.33, 2.00 and 2.00 for lists A, B and C). It is also unaffected by the participant subgroup that drives the fragile sureness result reported below: restricted to the twenty participants who both supplied these ratings and did not begin at the floor of the sureness scale, the difference is 2.20, $p < .001$.

That subgroup is twenty rather than the twenty-one of Table 2 because the participant missing relevance and informativeness data is one of the non-floor participants. This is the most robust finding in the study.

International Journal of Philosophy, Linguistics, and Humanities
E-ISSN: 3082-415X

Three further descriptions of the same data are worth giving, since Section 3.2.2 turns on the comparison. Coded as a direction report rather than a magnitude, the probability data classify 75.9% of Conflict trials as showing no change at all, against 98.1% of Inert and 3.7% of Directional. Note that this coding does not make Conflict and Inert alike: they differ on it as they do on the continuous measure. Binarizing relevance at the scale midpoint classifies 100% of Conflict and Directional trials as relevant against 0% of Inert. Binarizing informativeness the same way classifies 21% of Conflict trials as informative against 0% of Inert; read as a magnitude, 94% of Conflict trials sit above the floor of the scale against 4% of Inert. The graded gap is not carried by the minority who rated Conflict above the midpoint: among the 21 participants whose Conflict mean fell at or below it, a different subgroup from the prior-sureness one in Table 2 that happens to be the same size, Conflict still exceeds Inert by 1.62 points, $t(20) = 8.18$, $p < .001$, $dz = 1.79$.

At the same time, Conflict responses were rated far less informative than Directional responses: 3.15 against 6.31, a mean difference of −3.15, 95% CI [−3.66, −2.64], $t(25) = -12.81$, $p < .001$, $dz = -2.51$. The three conditions therefore order themselves as Directional > Conflict > Inert on perceived informativeness, with Conflict close to the middle of the scale rather than at either end.

3.1.4. Relevance

Relevance ratings showed a different pattern. Conflict and Directional were both at 6.96, with a difference of exactly zero for all 26 participants who supplied these ratings; Inert was at 1.58. Within Conflict the gap between relevance and informativeness was 3.81 points, $t(25) = 14.08$, $p < .001$, against 0.65 in Directional and 0.54 in Inert. A response rated just over three on informativeness is rated at seven on relevance, and exactly as relevant as one rated over six. Relevance is not proportional to perceived informativeness.

3.1.5. Sureness

The secondary prediction was that sureness would increase more in Conflict than in Inert, and taken at face value it was confirmed: mean difference +0.96, 95% CI [0.23, 1.70], $t(26) = 2.68$, $p = .013$, $dz = 0.52$, with a Wilcoxon signed-rank test at $p = .015$ and a sign-flip permutation test at $p = .012$.

International Journal of Philosophy, Linguistics, and Humanities
E-ISSN: 3082-415X

That summary would mislead on its own, and Table 2 reports the tests alongside the checks that undermine them. Three problems are serious enough that the effect should not be treated as established. A sign test is not significant: 13 participants showed a larger increase in Conflict, 7 a smaller one, 7 no difference, $p = .26$; the five largest contributors account for 67% of the summed positive difference, and excluding them the mean falls to +0.23, $p = .30$. Second, and more damaging, the effect is strongly associated with where participants placed themselves on the sureness scale before seeing any information: the correlation between mean prior sureness and the Conflict-minus-Inert difference is $r = -.535$, $p = .004$. The six participants reporting prior sureness of 1 on every item show a mean difference of +3.50; the remaining twenty-one show +0.24, $t(20) = 0.97$, $p = .34$, $dz = 0.21$. Third, the effect does not survive removal of one of the three lists: dropping list A reduces it to +0.36, $p = .25$.

Table 2. The sureness contrast (Conflict minus Inert) and its robustness. Participant means, one value per condition per participant.

Analysis	n	Mean diff.	t	p	dz
All data (primary)	27	+0.96	2.68	.013	0.52
Sign test on the same data	27	13+ / 7− / 7 ties	—	.26	—
Excluding prior-sureness floor (= 1)	21	+0.24	0.97	.34	0.21
Excluding list A	18	+0.36	1.20	.25	0.28
Excluding list B	18	+0.97	1.91	.073	0.45
Excluding list C	18	+1.56	3.49	<.01	0.82
Comparison: Directional minus Inert	27	+2.54	8.72	<.001	1.68

A cautious interpretation is that the sureness findings are consistent with the argument in Section 1.2.2 but do not establish it. The pattern within the floor subgroup is the predicted one, since those participants increased in Conflict and not in Inert, an asymmetry a pure response-style account does not predict. But with a scale whose bottom anchor was the natural resting point for a no-information state, the pilot cannot separate a real effect from an artifact of where participants started. Hence the argument here rests on the informativeness contrast rather than this one.

3.1.6. An Unexpected Observation

One pattern was not predicted and is reported as an observation. Most participants gave a prior of exactly 50, but eleven Conflict trials had a non-neutral prior. In all eleven the posterior moved toward 50, by a mean of 8.64 points, and in none away.

International Journal of Philosophy, Linguistics, and Humanities
E-ISSN: 3082-415X

The cell is small and is a by-product of a design that did not push priors off the midpoint, so no weight should be placed on it. It is noted because such movement is inconsistent with the likelihood-ratio derivation in Section 1.1.2, which preserves any prior exactly. Several explanations are possible and the design does not separate them. Participants may not have perceived the two sources as perfectly symmetrical despite the stated equal placement; they may have treated disagreement itself as a reason to reconsider an initial estimate; the instructions or the act of re-entering a number may have drawn responses toward the midpoint; or the non-neutral priors may simply have been unstable. These are listed as possibilities, not as mechanisms. Separating them requires a larger study with deliberately varied prior probabilities and a check on perceived source reliability.

3.2. Discussion

3.2.1. The Narrow Conclusion

The claim the study supports is narrow, and it is a claim about perceived informativeness. Conflict and Inert responses both produced very little directional probability change, and were nonetheless sharply distinguished on the graded item: 3.15 against 1.04 on a 1–7 scale, with the Inert control essentially at floor. On that item Conflict responses sit well above floor and below the midpoint. The Conflict mean signed change was small but not zero (−2.22 points, 95% CI [−4.41, −0.04]), and the equivalence bounds in Section 3.1.2 were chosen post hoc, so the claim is that directional change was small relative to the Directional condition, not that it was equivalent to zero.

Three readings of that gap must be distinguished, and the present design does not distinguish them. On the first, participants register that a report of two well-placed sources disagreeing puts genuine evidence about the question into the common ground, on both sides. On the second, participants register only that the response is about the proposition under discussion: Conflict responses state the QUD proposition explicitly, once affirmed and once denied, whereas Inert responses never mention it, so a rating that tracks aboutness predicts the result with no appeal to evidence entering the common ground. On the third, the rating tracks how helpful or conversationally appropriate the response is, the dimension Feinmann and Ramotowska (2026) name. Because the experiment cannot separate these, the informativeness rating is not treated here as direct evidence that participants registered an increase in the weight of evidence. What it shows is that a response producing very little directional probability change is not perceived as informationally on a par with an Inert one; which of the three readings carries that gap is left open.

International Journal of Philosophy, Linguistics, and Humanities
E-ISSN: 3082-415X

What follows is the claim set out in Section 1.1.3, now with evidence attached: the two configurations of Section 1.2.2 are indistinguishable to a directional criterion and are not, on this evidence, indistinguishable to language users. That motivates a richer conception of QUD-directed epistemic impact. It does not establish one, and Section 3.2.4 gives four reasons why the present study cannot.

This leaves the description of the phenomenon different from the familiar one. Described as relevance without informativity, the cases look like a problem for the whole family of accounts. Described as high relevance with intermediate perceived informativeness, contrasting with Inert responses that have neither, they look like a problem for the directional implementation, which is a smaller claim and one compatible with some informativity condition remaining necessary for relevance.

3.2.2. Relation to Feinmann and Ramotowska (2026)

Feinmann and Ramotowska (2026) provide the closest experimental test of the relationship at issue, and their design is worth setting out because the comparison turns on it. In a within-subjects study with one hundred participants and twelve items they compared informative on-topic replies, uninformative off-topic replies, and a target class of replies that are intuitively uninformative yet do not seem to incur a relevance violation. Relevance was measured by a forced choice between "And what does that have to do with anything?" and "Alright, thanks." Informativeness was measured by a three-point question asking whether the answer had become less likely, neither more nor less likely, or more likely, the directional responses being recoded as informative. Target replies were judged relevant on 95% of trials against 3%, while the rate of uninformative judgments did not differ reliably across the two conditions. Table 3 sets the two designs side by side.

Table 3. The present study and Feinmann and Ramotowska (2026) compared.

Dimension	Feinmann and Ramotowska (2026)	Present study
Participants	100	27
Items	12	6 scenarios, three counterbalanced lists
Relevance measure	Forced choice between "And what does that have to do with anything?" and "Alright, thanks."	Seven-point rating
Informativeness measure	Three-point directional judgment (less likely / neither / more likely)	Seven-point perceived-informativeness rating, with 0–100 probability before and after the response

International Journal of Philosophy, Linguistics, and Humanities
E-ISSN: 3082-415X

Control	Off-topic reply	Topically matched Inert reply (same protagonists and situation)
Main focus	Relevance–informativity relationship	Cancellation versus independence
Principal result	Relevance without directional change (95% vs. 3%)	Conflict rated more informative than Inert (3.15 vs. 1.04) with very little directional change

The second measure is where the two papers meet. It operationalizes informativeness as directional first-order probability change: whether the reply moved the estimate up, down, or not at all, so that a reply coded uninformative is one for which participants reported no directional shift in the probability of the YES cell. That is a legitimate operational definition, and it is exactly the one Probabilistic QUD Relevance licenses; the objection here is not that they measured the wrong thing but that the measure is incomplete for the inference drawn from it. Being no independent graded measure of informational content, it can code a reply uninformative while leaving open whether that reply is perceived as carrying information. Their finding is that replies producing no reported directional change are nonetheless judged relevant: the antecedent of the inference at issue, established at a scale the present study cannot match, and not its consequent.

On their relevance measure the present data reproduce their pattern: the binarized relevance ratings classify 100% of Conflict and 0% of Inert trials as relevant, against their 95% and 3%. Their informativeness result is a null, and a null is not something a study of this size can be said to replicate. What can be said is this. Coded as a direction report, the present data show the same predominance of little or no directional movement in the Conflict condition, 75.9% of trials at exactly zero against 3.7% of Directional. But they also show a small detectable Conflict–Inert difference, on that coding (75.9% against 98.1%) and on the continuous measure (2.78 against 0.09 points, $p = .014$), which a trichotomous direction report is not designed to register. Where the studies part company is on a measure they did not take: a graded judgment of how informative the reply was, on which Conflict sits three scale points above Inert.

Their own data run in the same direction without reaching significance: in the target condition 52 of 400 judgments were coded informative, against 34 of 400 in the off-topic condition ($\beta = -1.96$, $z = -1.78$, $p = .08$). They report this correctly as no reliable difference, and it is not evidence for one; neither is it evidence of equivalence.

International Journal of Philosophy, Linguistics, and Humanities
E-ISSN: 3082-415X

Underneath the compatibility lies a disagreement about where the residue is located. Feinmann and Ramotowska suggest that target replies convey meta-information about the conversational situation, signalling that the speaker is not in a position to answer, and they hold that such replies have no informational impact on the QUD. The present suggestion places the residue inside the QUD rather than beside it: a report that two well-placed sources disagree puts evidence on both sides of the question asked into the common ground while leaving P(q) where it was. The proposals are close, and their observation that such replies guide the next conversational move is very nearly the point. They differ testably: on theirs the graded informativeness of a disagreement report should be no higher than that of an off-topic reply, and on this one it should be higher, which is the pattern Section 3.1.3 reports, though Section 3.2.4 notes that the present item cannot rule out other readings of it.

Feinmann and Ramotowska anticipate part of this. They observe that their target reply can naturally be followed with "Well, that's not very helpful…" whereas their informative reply cannot, and suggest that informativeness seems to track another dimension of linguistic evaluation, namely perceived helpfulness (Feinmann & Ramotowska, 2026, p. 422). The graded item used here may be quantifying that dimension rather than anything about the participant's beliefs, and the present design cannot tell; Section 3.2.4 returns to it.

One candidate explanation can be set aside: the gap is not the product of a weak control. Their off-topic replies are topically unrelated ("Mary lives in an old house"), whereas the Inert responses here named the same protagonists and situation, which should if anything raise their ratings and compress the contrast. They nonetheless sat at 1.04, with 96% of trials at the absolute floor.

One hundred participants across twelve items is a better basis for a claim about relevance than twenty-seven across six, and nothing here overturns their result. What is at issue is what follows from it, and a measure recording only the direction of first-order change cannot by itself settle whether the responses it codes uninformative are perceived as informationally on a par with responses that bear on nothing.

3.2.3. What Survives of the Original Argument

A good deal does, and these results are no general vindication of Informativity-based QUD Relevance. First, Feinmann's central dissociation is reproduced here. Conflict responses were rated exactly as relevant as Directional responses while being rated at half their informativeness.

International Journal of Philosophy, Linguistics, and Humanities
E-ISSN: 3082-415X

Feinmann and Ramotowska's (2026) data make the point independently, since their informativeness judgments did not predict their relevance judgments. The present ratings are consistent with it, but they are at ceiling for both Conflict and Directional (6.96, SD 0.19), so the exact equality should be read as an upper bound on any difference rather than as an established identity; what they show is that a large difference in perceived informativeness is not accompanied by any detectable difference in rated relevance.

Second, Standard QUD Relevance remains refuted: Conflict responses eliminate no cell of a polar QUD and are nonetheless relevant. Third, the present argument does not show what the epistemic effect participants register consists in. The ratings show that Conflict is not perceived as empty; they do not show what participants take it to contain, and the sureness data were too fragile to say.

Fourth, nothing here bears on Feinmann's discussion of QUD accommodation, strategies of inquiry, ignorance inferences, or subject matter, which address what a positive theory of relevance should look like. This paper does not offer one.

3.2.4. Scope and Alternative Explanations

Four restrictions on the conclusion should be made explicit. First, the study tested one configuration, and the weaker of Feinmann's two. All six Conflict responses were conjunctive (Section 2.2), the form for which he raises an incremental, clause-by-clause reading, observing that the atomic paraphrase is not open to it (Feinmann, 2026, pp. 5–6). Incremental informativeness is therefore a live alternative explanation of the graded rating: participants may have rated these responses as partly informative because each clause is informative taken alone. Nothing here distinguishes that from the reading in Section 1.2.2. His game-show case, and the expressions of unsettledness in the target class of Feinmann and Ramotowska (2026), are likewise outside the scope of anything claimed here.

Second, the 1–7 item asked how informative the response was for answering the question, and what it records is perceived informativeness. The three readings set out in Section 3.2.1 cannot be excluded on this design, and the item was collected immediately after a relevance judgment about the same response, so carryover from that judgment is possible. The experiment therefore does not establish that conflicting responses carry weight in any formal sense, or that any higher-order epistemic change occurred, and Section 1.2 should be read as a proposal about where to look rather than as something these data confirm. What it establishes is narrower: that a response producing very little directional probability change need not be perceived as informationally on a par with an inert one. In the item's defence, the same participants had just given probability estimates that barely moved, so a simple conflation of informativeness with belief change did not produce the contrast.

International Journal of Philosophy, Linguistics, and Humanities
E-ISSN: 3082-415X

Third, and relatedly, the Conflict and Inert responses differ in more than evidential bearing. As Section 3.2.1 sets out, the contrast is confounded with congruence between the response and the subject matter of the QUD, in the sense Feinmann discusses in his own Section 1.2.4, and an account on which perceived informativeness tracks whether a response is about whether p predicts the result with no appeal to evidence entering the common ground. Separating the two would require materials matched for QUD-aboutness, which these are not. I record this as an alternative explanation rather than attempting to defeat it.

Fourth, N = 27, against one hundred in the closest comparable study. A single small study recruited by direct message does not settle a theoretical question. It shows that the question is empirically tractable, and that on a first look the data fall on the side that a directional criterion alone would not have predicted.

4. Conclusion

Feinmann (2026) presents symmetric conflicting evidence as a response that is uninformative to the QUD yet perceived as relevant. Under his own gloss, on which being uninformative to the QUD is having no detectable effect on the QUD cells, the description is accurate and the conclusion follows; nothing here disputes either, and he is explicit that he has no quarrel with Info-QUD-R as such.

What the cases establish is relevance without directional QUD-cell probability change; what they do not by themselves establish is relevance without every plausible form of QUD-directed informativity, since a directional criterion cannot separate a response that puts evidence bearing on the question into the common ground on both sides from one that puts none there at all. A small exploratory study found that language users separate them: explicit conjunctive cancellation and a topically matched Inert control both produced very little directional probability change, yet were sharply distinguished on a graded measure of perceived informativeness.

The conclusion is correspondingly narrow, and it is a motivation rather than a result. Very little directional probability change does not by itself warrant treating a response as perceived informationally equivalent to an inert one, which is reason to look for a richer conception of QUD-directed epistemic impact. What that conception should be, and whether the contrast survives materials matched for QUD-aboutness and an atomic rather than conjunctive cancellation form, are questions these data do not answer.

What the result contributes to probabilistic approaches to QUD relevance is correspondingly limited.

International Journal of Philosophy, Linguistics, and Humanities
E-ISSN: 3082-415X

It indicates that a criterion reading only the direction of first-order change is incomplete as a measure of what language users treat as QUD-directed informativeness, and that perceived informativeness is a dimension such approaches will need to represent; it does not supply a model of that dimension. The most important next steps follow from the restrictions in Section 3.2.4: Conflict responses in the atomic as well as the conjunctive form; control responses matched for explicit reference to the QUD proposition; a manipulation check of perceived source reliability; several indicators of perceived informativeness, separating evidential content, diagnosticity, helpfulness, and subjective uncertainty reduction; deliberately varied prior probabilities; a larger and more diverse sample with a preregistered analysis plan; and more scenarios per condition, assigned to conditions independently of list and analysed with a mixed-effects model treating participants and items as random effects.

5. Limitations of the Study

The study is exploratory and several of its features would not be acceptable in a confirmatory design. They are listed in full, because the informativeness result should be assessed against them.

The sample is small (N = 27) and was recruited by convenience through direct messages rather than a participant platform; several response vectors are near-identical, which a recruitment chain of this kind cannot distinguish from a homogeneous sample. Item and list are perfectly confounded in the critical condition, since each scenario appears in Conflict on exactly one list, and the large participant-level effect size for the informativeness contrast does not reduce the uncertainty that comes from six items so distributed. Roughly 70% of participants gave identical answers to their two items within the Conflict and Inert conditions, so the effective number of independent observations per cell is closer to one than two; this is why participant means are used throughout, and why trial-level mixed models, which give far smaller p values for the sureness contrast, are not primary. Sixty-three percent gave an identical prior response to all six scenarios, which is why Section 3.1.6 rests on so few trials. Relevance ratings for Conflict and Directional are at ceiling (6.96, SD 0.19), so their exact equality may be compression.

Six participants began at the floor of the sureness scale and two at its ceiling, and one gave identical answers throughout and failed the manipulation check; all were retained. All Conflict responses were conjunctive and the Inert responses were not matched for explicit reference to the QUD proposition (Sections 2.2 and 3.2.4), which bears directly on the interpretation of the central result, and the informativeness measure is a single self-report item (Section 2.4).

International Journal of Philosophy, Linguistics, and Humanities
E-ISSN: 3082-415X

Finally, the participants were native speakers of American English and the materials six everyday scenarios with polar questions, so the results cannot be assumed to generalize to other linguistic populations, discourse contexts, or question types.

What these limitations bear on differs by result. The informativeness contrast is large, present in every list and every participant who supplied ratings, and unaffected by the floor subgroup; it is least exposed to the sampling and response-style limitations, though its interpretation is bounded by the stimulus and measurement limitations just noted. The sureness contrast is exposed to all of them.

Supplementary Details

Supplementary Materials & Data Availability Statement:
Non-applicable.

Author Contributions:
The author is the sole contributor to this manuscript and was responsible for all aspects of the research, including conceptualization, methodology, software, validation, formal analysis, investigation, data curation, writing—original draft, and writing—review and editing.

This journal uses the Contributor Roles Taxonomy (CRediT).

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Muteb Alqarni	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			✓	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

Funding:
This research received no external funding

Institutional Review Board Statement:
Non-applicable.

Informed Consent Statement:
Non-applicable.

Conflicts of Interest:
The authors declares no conflict of interest.

Disclosure Statement:
ChatGPT-5 (OpenAI) was used solely for proofreading and language editing to improve the clarity, grammar, and readability of the manuscript. All research content, ideas, analyses, interpretations, and conclusions are the author's own. The author reviewed and verified the AI-assisted edits and takes full responsibility for the accuracy, integrity, and final content of the manuscript.

References

Büring, D. (2003). On D-trees, beans, and B-accents. Linguistics and Philosophy, 26(5), 511–545. https://doi.org/10.1023/A:1025887707652

Carnap, R. (1950). Logical foundations of probability. University of Chicago Press.

Feinmann, D. (2026). A note on relevance. Journal of Semantics, 43(1), 1–15. https://doi.org/10.1093/jos/ffaf013

Feinmann, D., & Ramotowska, S. (2026). It may or may not be relevant: Testing the relationship between relevance and informativeness. In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana, & T. Torrent (Eds.), Proceedings of the 48th Annual Conference of the Cognitive Science Society (pp. 417–423). Cognitive Science Society. https://escholarship.org/uc/item/7gz1k15k

Ginzburg, J. (2012). The interactive stance: Meaning for conversation. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199697922.001.0001

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), Syntax and semantics: Vol. 3. Speech acts (pp. 41–58). Academic Press.

Groenendijk, J., & Stokhof, M. (1984). Studies on the semantics of questions and the pragmatics of answers [Doctoral dissertation, University of Amsterdam].

Joyce, J. M. (2005). How probabilities reflect evidence. Philosophical Perspectives, 19(1), 153–178. https://doi.org/10.1111/j.1520-8583.2005.00058.x

Keynes, J. M. (1921). A treatise on probability. Macmillan.

Krifka, M. (2008). Basic notions of information structure. Acta Linguistica Hungarica, 55(3–4), 243–276. https://doi.org/10.1556/ALing.55.2008.3-4.2

Lakens, D. (2017). Equivalence tests: A practical primer for t tests, correlations, and meta-analyses. Social Psychological and Personality Science, 8(4), 355–362. https://doi.org/10.1177/1948550617697177

Roberts, C. (2012). Information structure in discourse: Towards an integrated formal theory of pragmatics. Semantics and Pragmatics, 5(6), 1–69. https://doi.org/10.3765/sp.5.6 (Original work published 1996)

Simons, M., Tonhauser, J., Beaver, D., & Roberts, C. (2010). What projects and why. Semantics and Linguistic Theory, 20, 309–327. https://doi.org/10.3765/salt.v20i0.2584

Stalnaker, R. (2002). Common ground. Linguistics and Philosophy, 25(5–6), 701–721. https://doi.org/10.1023/A:1020867916902

van Kuppevelt, J. (1995). Discourse structure, topicality and questioning. Journal of Linguistics, 31(1), 109–147. https://doi.org/10.1017/S002222670000058X

van Rooy, R. (2003). Questioning to resolve decision problems. Linguistics and Philosophy, 26(6), 727–763. https://doi.org/10.1023/B:LING.0000004548.98658.8f

Warstadt, A., & Agha, O. (2022). Testing Bayesian measures of relevance in discourse. Proceedings of Sinn und Bedeutung, 26, 872–886. https://doi.org/10.18148/sub/2022.v26i0.1034

Warstadt, A., Agha, O., & Franke, M. (2026). Quantifying epistemic relevance. Open Mind, 10, 502–546. https://doi.org/10.1162/opmi.a.332

Cite as: Alqarni, M. (2026). Undirected Need Not Be Uninformative: Symmetric Conflicting Evidence and the Measurement of QUD Informativity. *International Journal of Philosophy, Linguistics, and Humanities,* 2(2), 77-101. https://doi.org/10.70847/744173