What is at-issueness? An experimental comparison of diagnostics

At-issueness is a key concept in theoretical semantics/pragmatics, but there is no consensus about how it is defined or diagnosed (e.g., Tonhauser 2012; Tonhauser et al. 2018; Koev 2018). We present experimental data investigating whether four widely used diagnostics for at-issueness yield consistent results. Our findings reveal significant differences across diagnostics, indicating they are not interchangeable. Since the diagnostics target distinct theoretical conceptions of at-issueness, these differences offer insight into their comparability.

At-issueness: Diagnostics and theoretical conceptions. Four commonly used diagnostics for at-issueness are illustrated in (1–4) for sentence-medial non-restrictive relative clauses (NRRCs). As their content is generally taken to be not-at-issue, participants are expected to: Give low naturalness ratings under the QUD diagnostic (1) and the direct dissent diagnostic (3), not interpret the speaker to be asking about the content under the 'asking-whether' diagnostic in (2), will choose one of the *yes*-responses under the 'yes, but' diagnostic in (4). These diagnostics reflect different theoretical conceptions of at-issueness: The QUD diagnostic targets Q-at-issueness (Koev 2018), where at-issue content addresses the QUD; the 'asking whether' diagnostic assumes that at-issue content of interrogatives partitions the context set (Tonhauser et al. 2018); and the direct dissent and 'yes, but' diagnostics reflect P-at-issueness, where at-issue content constitutes a proposal to update the common ground (Koev 2018).

(1) QUD diagnostic (e.g., Tonhauser 2012; Chen 2024)
A: *What did Greg buy?* B: *Greg, who bought a new car, is envied by his neighbor.*
Question to participants: How well does B's response fit A's question?

(2) 'asking whether' diagnostic (e.g., Tonhauser et al. 2018; Solstad and Bott 2024)
Is Greg, who bought a new car, envied by his neighbor?
Question to participants: Is the speaker asking whether Greg bought a new car?

(3) Direct dissent diagnostic (e.g., Tonhauser 2012; Syrett and Koev 2015)
A: *Greg, who bought a new car, is envied by his neighbor.* B: *No, that's not true, he didn't buy a new car.*
Question to participants: How natural is B's rejection of A's utterance?

(4) 'yes, but' diagnostic (e.g., Xue and Onea 2011; Destruel et al. 2015)
A: *Greg, who bought a new car, is envied by his neighbor.*
B: *Yes, but he didn't buy a new car.* / *Yes, and he didn't buy a new car.* / *No, he didn't buy a new car.*
Task for participants: Choose the response that sounds best.

Prior research has identified disagreements, potentially arising from diagnostic differences: While Syrett and Koev 2015 found that sentence-medial NRRCs are more at-issue than sentence-final ones using the direct dissent test, Drozdov 2024 found no difference with the 'asking whether' diagnostic. To investigate how consistent the diagnostics are, we conducted four experiments measuring the at-issueness of the same contents across diagnostics.

Methods. For each of the four experiment, we recruited 80 participants on Prolific, and tested the same seven types of content: the content of sentence-medial and -final NRRCs, and the content of the clausal complements of *discover, know, be right, confirm* and *confess*. Each content type was instantiated by one of seven items (e.g., 'Greg bought a new car') and realized as either an assertion (1), (3), (4) or a polar question (2). Participants responded for the seven target stimuli and two control stimuli, by adjusting a slider for (1–3), or by choosing a response in (4).

Results. Figures 1 and 2 show the mean responses by content for the four diagnostics. Comparing the results across diagnostics reveals some key differences. First, the diagnostics vary in their sensitivity to differences between contents: The by-content means in Experiment 2 ('asking whether' diagnostic) show a larger range (Figure 1b) than in the other three experiments (Figures 1a, 1c and 2). Second, the content manipulation affects the ratings differently across the four diagnostics, sometimes in opposite directions. This results in a different order of predicates by response means between experiments. For instance, *be right* ranks highest under the 'asking whether' diagnostic (Figure 1b), and the 'yes, but' test (Figure 2), but ranks lowest under the QUD-diagnostic (Figure 1a), and shows no clear effect in the direct dissent diagnostic (Figure 1c).

Discussion. The differening results between diagnostics suggest that they are not interchangeable. In particular, the varying relative order of by-content means across diagnostics provide an initial argument that they target distinct properties of the content. Further, while the 'asking whether' diagnostic, for contents embedded in questions, is sensitive enough to detect fine-grained differences between contents, the smaller range of response means for the other diagnostics could suggest the need for a more sensitive diagnostic for contents embedded in declarative assertions. Additional comparison Syrett and Koev 2015 (details omitted in the abstract) points to potential effects of the response task and the speech act of the utterance embedding the tested content.

Figure 1: Mean responses by content for the QUD-diagnostic (a), 'asking whether' diagnostic (b), and 'direct dissent' diagnostic (c). Error bars indicate 95% bootstrapped confidence intervals. Violin plots indicate the kernel probability density of the individual participants' ratings, which were given on a 0–1 scale, by adjusting a slider.

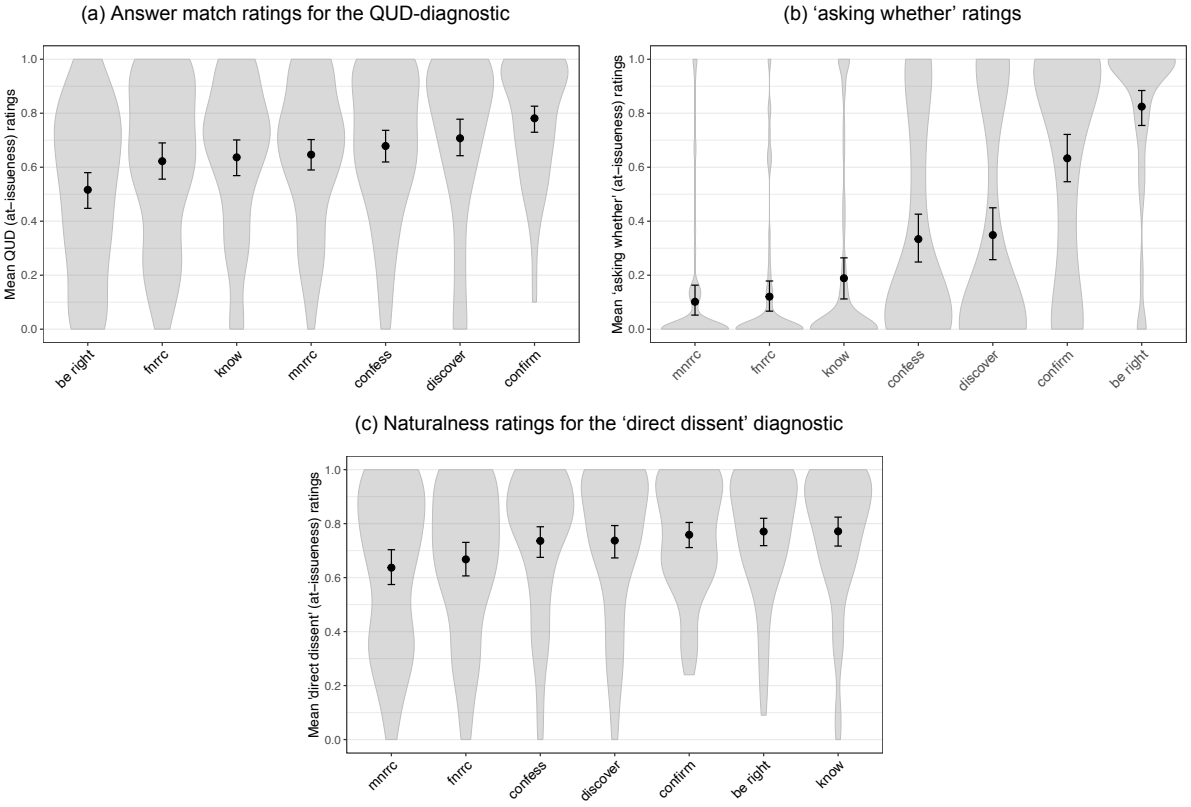
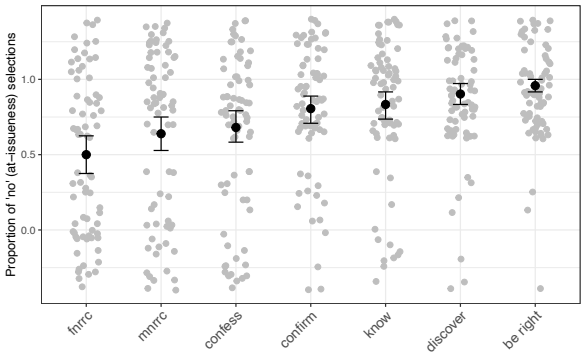


Figure 2: Proportion of 'no' choices by content for the 'yes but' diagnostic. Error bars indicate 95% bootstrapped confidence intervals. Gray dots indicate individual participant responses (either 'no' or one of the 'yes'-responses, jittered vertically and horizontally for legibility).



References. Yuqiu Chen. Presuppositions at the semantics-pragmatics interface. 2024. • Emilie Destruel, Edgar Onea, Daniel Velleman, Dylan Bumford, and David Beaver. A cross-linguistic study of the non-at-issueness of exhaustive inferences. In Floran Schwarz, editor, Experimental approaches to presupposition, pages 135–156. Springer, 2015. • Katharina Drozdov. Projection and at-issueness of nonrestrictive relative clauses. Unpublished manuscript, 2024. • Todor Koev. Notions of at-issueness. Language and Linguistics Compass, 12(12):e12306, 2018. • Torgrim Solstad and Oliver Bott. Cataphoric resolution of projective content: The case of occasion verbs. Semantics and Pragmatics, 17(11):1–66, July 2024. • Kristen Syrett and Todor Koev. Experimental evidence for the truth conditional contribution and shifting information status of appositives. Journal of Semantics, 32(3):525–577, 2015. • Judith Tonhauser. Diagnosing (not-) at-issue content. Proceedings of Semantics of Under-represented Languages of the Americas (SULA), 6:239–254, 2012. • Judith Tonhauser, David I Beaver, and Judith Degen. How projective is projective content? gradience in projectivity and at-issueness. Journal of Semantics, 35(3):495–542, 2018. • Jingyang Xue and Edgar Onea. Correlation between presupposition projection and at-issueness: An empirical study. In Proceedings of the ESSLLI 2011 workshop on projective meaning, pages 171–184, 2011.