

# Which predicates are factive? An empirical investigation\*

Author(s)

April 4, 2019

## Abstract

It has long been assumed that properties of the content of the clausal complement distinguish factive predicates like *know* from non-factive predicates like *think*. There is, however, disagreement about which properties define factive predicates and whether the content of the complement of particular predicates exhibits the properties attributed to factive predicates. The resulting disagreement about which predicates are factive is vexatious because the distinction between factive and non-factive predicates has played a central role in semantic and pragmatic theorizing, especially in the study of presuppositions. This paper investigates properties of the content of the complement of clause-embedding predicates, with the goal of understanding how such predicates can be classified. We argue that factive predicates are more heterogeneous than previously assumed and that there is little empirical support for the assumed categorical distinction between factive and non-factive predicates. We conclude by discussing the implications of our findings for analyses of presuppositions and other projective content.

## 1 Introduction

A long-standing and widely held assumption is that factive predicates like *know* differ from non-factive predicates like *think* in properties of the content of the clausal complement.<sup>1</sup> There is, however, disagreement in the literature about which properties define factive predicates and whether the content of the complement of particular predicates exhibits the properties attributed to factive predicates. As a consequence, there is disagreement about which predicates are factive.

Consider first the defining properties of a factive predicate. In Kiparsky and Kiparsky 1970, the content of the complement of a factive predicate was taken to be presupposed: “the speaker presupposes that the embedded sentence expresses a true proposition” (p.147).<sup>2</sup> By contrast, no such presupposition was taken to arise from the complement of a non-factive predicate. Diagnostic for whether the content of the complement was presupposed was projection of the content from under negation or out of a polar question (p.150f): *regret*, for instance, was taken to be a factive predicate because (1a) “presupposes...that the door is closed” (p.150), whereas *think* was taken to be a non-factive predicate because the content of the complement does not project over negation in (1b).<sup>3</sup>

(1) a. John doesn’t regret that the door is closed.

(Kiparsky and Kiparsky 1970:150)

---

\*Acknowledgments removed for review process

<sup>1</sup>Syntactic properties have also been taken to distinguish factive and non-factive predicates; see, e.g., Kiparsky and Kiparsky 1970, White et al. 2018 and references in the latter. We ignore syntax here because it is not clear that syntactic properties correlate with semantic/pragmatic ones, as already pointed out in Kiparsky and Kiparsky 1970:fn.3; see also White and Rawlins 2018.

<sup>2</sup>They say sentences presuppose but also talk about speakers presupposing. NEED TO ADDRESS THIS.

<sup>3</sup>Predicates identified as factive in Kiparsky and Kiparsky 1970 include *be significant*, *be odd*, *be tragic*, *be exciting*, *be relevant*, *matters*, *counts*, *makes sense*, *suffices*, *amuses*, *bothers* (p.143), *regret*, *be aware*, *grasp*, *comprehend*, *take into consideration*, *take into account*, *bear in mind*, *ignore*, *make clear*, *mind*, *forget*, *deplore*, *resent*, *care* (p.145) as well as *know* and *realize* (p.147).

- b. John doesn't think that the door is closed.

More recently, however, researchers have assumed that the content of the complement of a factive predicate is not merely presupposed but also entailed (e.g., Gazdar 1979a:119-123, Chierchia and McConnell-Ginet 1990:355, van der Sandt 1992:345, Abbott 2006:3, Schlenker 2010:139, Anand and Hacquard 2014:77, Spector and Egré 2015:fn.7). Beaver (2001:66f.), for instance, notes that Gazdar (1979a:119-123) “describes the inferences associated with factive verbs...as being indefeasible in simple affirmative sentences”, i.e., “entailments”. Chierchia and McConnell-Ginet (1990:355) illustrate with *realize* that “[a] sentence can both entail and presuppose another sentence”: *Joan realizes that syntax deals with sentence structure* “both entails and presupposes” that syntax deals with sentence structure. And Schlenker (2010:139), in discussing contents of clausal complements, points to “the pattern of inference which is characteristic of presuppositions: an entailment of the positive sentence is preserved under negation and in questions”.

This characterization of the properties of a factive predicate differs from that assumed in Kiparsky and Kiparsky 1970: it requires the content of the complement to be true in all worlds in which the affirmative sentence with the clause-embedding predicate is true rather than merely requiring that a speaker who utters an affirmative sentence with the predicate presupposes the truth of the content. Predicates on which the two definitions come apart include *take into account*, *bear in mind* and *make clear*. These are considered factive in Kiparsky and Kiparsky 1970 because utterances by Rahim of the polar question in (2a) and the affirmative sentence in (2b) lead listeners to take Rahim to presuppose the truth of the content of the complement, that the groom is unreliable. But true affirmative sentences with these predicates do not entail the content of the complement, as shown by the acceptability of (2c). Thus, the predicates are not factive on the more recent characterization of factive predicates.

- (2) a. Rahim: “*Did Joan { take into account / bear in mind / make clear } that the groom is unreliable?*”  
 b. Rahim: “*Joan { took into account / bore in mind / made clear } that the groom is unreliable.*”  
 c. Rahim said that Joan { took into account / bore in mind / made clear } that the groom is unreliable, but she was wrong to do so because the groom is a very reliable person.

The literature not only disagrees about the defining properties of a factive predicate, but also about whether the content of the complement of particular predicates exhibits the properties taken to characterize factive predicates.

- Consider first the question of whether the content of the complement is entailed.
- Schlenker (2010:139) assumed that the content of the complement of *inform* is entailed, i.e., that Mary being pregnant necessarily follows from (3a). Anand and Hacquard (2014:76), however, pointed out that *inform* can be modified by *falsely*, as in the naturally occurring example in (3b), which shows that the content of the complement does not necessarily follow from a true affirmative sentence with *inform*.

- (3) a. Mary has informed her parents that she's pregnant.  
 b. Doctors falsely informed my parents that I was dying.<sup>4</sup>

- Similarly, both Wyse (2010) and Swanson (2012) took the content of the complement of *establish* to be entailed, but the naturally occurring example in (4), in which *establish* is modified by *falsely*, again shows that this content does not necessarily follow from unembedded matrix sentences with *establish*.

<sup>4</sup><https://www.quora.com/Has-a-doctor-ever-falsely-informed-you-that-you-re-dying>

- (4) The testimony is uncontradicted that Jew Sing committed at least six perjuries and a subornation of perjury to conceal his wrongful entry into the country, and falsely established that he was an American citizen. https:

[//law.justia.com/cases/federal/appellate-courts/F2/202/715/216831](https://law.justia.com/cases/federal/appellate-courts/F2/202/715/216831)

- There is also disagreement about whether the content of the complement of emotive predicates is entailed. For instance, as noted in Abrusán 2011:514, “from [(6a)] one tends to infer that it is raining”, suggesting that the content of the complement is entailed. Second, a speaker who utters the sentence in (6b) is judged to contradict themselves (as indicated by the hashmark): it does not seem possible for the speaker to be committed to the proposition that it is not raining and to the proposition that John regrets that it is raining.

- (5) Abrusán 2011:514

- a. I doubt that John regrets that it's raining.
- b. #It's not raining but John regrets that it's raining.

It has, however, long been noted that emotive predicates merely introduce the presupposition that the subject of the attitude believes in the truth of the complement (e.g., Heim 1992:203), not that the content of the complement is true. Thus, the contents of the complements of *regret* in (7a) and of *be enraged* in (7b) do not follow from these affirmative sentences. The existence of such examples has led some authors, including Klein (1975), Giannakidou (1998), Schlenker (2003a) and Egré (2008), to conclude that the content of the complement of emotive predicates is not entailed. **FACTIVE ON NOBODYS DEF?**

- (6) a. Falsely believing that he had inflicted a fatal wound, Oedipus regretted killing the stranger on the road to Thebes (Klein 1975)  
 b. John wrongly believes that Mary got married, and he is enraged that she is no longer single. (adapted from Egré 2008, based on Schlenker 2003b)

Emotive predicates were also taken to be factive by Gazdar 1979a, Abrusán 2011 and Anand and Hacquard 2014, who maintain that the content of the complement of emotive predicates is not merely presupposed but also entailed.

These were taken to be factive Kiparsky and Kiparsky 1970 and Karttunen 1971a,b under the assumption that the content of the complement of factive predicates is presupposed: listeners take Wilfred, the speaker of (5a) and (5b) to presuppose the truth of the content of the complement, that it was raining.

- (7) a. Wilfred: “*Svea { was glad / regretted } that it was raining.*”  
 b. Wilfred: “*{ Was Svea glad / Did Svea regret } that it was raining?*”  
 c. #It's not raining but John regrets that it's raining. (Abrusán 2011:514)

- There is similar disagreement about whether the content of the cognitive predicate *know* is entailed. The examples in (8) from Hazlett 2010 and Abrusán 2011 show that the truth of the content of the complement does not necessarily follow from true affirmative sentences with *know*.<sup>5</sup> Thus, it does not

<sup>5</sup>Some authors, including Gazdar (1979a) and Abrusán (2011), do not take examples with emotive predicates like (7) or with cognitive predicates like (8) to show that the content of the complement is not entailed. They instead suggested that such examples involve free indirect discourse or are understood with a different concept of knowledge. We discuss these proposals in section 3 in the context of the findings of our entailment experiments.

follow from (8a) that ulcers are caused by bacterial infection or from (8b) that the police are listening to John's phone calls. **SUBJECT OF ATTITUDE!**

- (8) a. Everyone knew that stress caused ulcers, before two Australian doctors in the early 80s proved that ulcers are actually caused by bacterial infection. (Hazlett 2010:501)  
b. John suffers from paranoia. He falsely believes that the police is spying on him and what is more he knows they are listening to his phone calls. (Abrusán 2011:514)

*remember* is not a factive predicate for KK and K, but for many semanticists.

In addition to factive and non-factive predicates, which combine with factive and non-factive complements, respectively, Kiparsky and Kiparsky (1970) also assumed a third class of predicates that are ambiguous and "take complements of both types" (p.163), including *acknowledge*, *announce*, *deduce* and *admit*.<sup>6</sup> For instance, even though the content of the complement appears to follow from (9a) and may project from (9b), *acknowledge* is not a factive predicate because the content of the complement is not entailed, as shown by the naturally occurring example in (9c).

- (9) a. Joan acknowledged that Dan is unreliable.  
b. Did Joan acknowledge that Dan is unreliable?  
c. By signing the waiver, Mr. Lyttle falsely acknowledged that he was a citizen of Mexico and that he agreed to be voluntarily deported to Mexico, despite the fact that Mr. Lyttle was and is a United States citizen.<sup>7</sup>

The idea that the content of the complement is not merely presupposed but also entailed seems to have been around for a while. Karttunen 1971a:342f, for instance, pointed out that the content of the complement of factive and implicative predicates is entailed from affirmative main sentences with such predicates: "the truth of the affirmative main sentence implies the truth of the complement". It is sentences with negation that differentiate the two: "[i]t is characteristic of factive verbs that negation in the main sentence does not affect the presupposition expressed in the complement".

*regret* is factive according to Karttunen (1971b:55) and Kiparsky and Kiparsky 1970

Karttunen:

semi-factive: discover, realize, find out, see, notice

factive: regret, forget, resent (know)

- Factive, non-factive and no specification (KK and Karttunen); nowadays: factive versus non-factive  
Karttunen 1971a:340: "There are also verbs that are not inherently factive or non-factive, e.g., *report*, *announce*, *remember*. They can be used with or without presupposition of the truth of their complement sentence.

---

<sup>6</sup>Kiparsky and Kiparsky 1970:163: "So far, for clarity of exposition, only predicates which are either factive or non-factive have been examined. For this set of cases, the factive and non-factive complement paradigms are in complementary distribution. But there are numerous predicates that take complements of both types. [...] we find verbs which occur indifferently with factive and non-factive complements, e.g., *anticipate*, *acknowledge*, *suspect*, *report*, *remember*, *emphasize*, *announce*, *admit*, *deduce*. Such verbs have no specification in the lexicon as to whether their complements are factive. On a deeper level, their semantic representations include no specification as to whether their complement sentences represent presuppositions by the speaker or not. Syntactically, these predicates participate in both complement paradigms."

<sup>7</sup>[https://www.acluga.org/sites/default/files/field\\_documents/georgia\\_initial\\_complaint.pdf](https://www.acluga.org/sites/default/files/field_documents/georgia_initial_complaint.pdf)

- So far we've only looked at factive versus non-factive. But of course more fine-grained distinctions.

And whereas Karttunen (1971b) distinguished semi-factive from factive predicates

Karttunen 1971a:340: "There are also verbs that are not inherently factive or non-factive, e.g., *report*, *announce*, *remember*. They can be used with or without presupposition of the truth of their complement sentence.

Karttunen (2016) divided the factive predicates of Kiparsky and Kiparsky 1970 into five separate classes

- how do we make progress: diagnose
- diagnostic for presupposition: projection
- More recent literature: increasing recognition that projection is not restricted to presuppositional contents (see Potts on conventional implicature, refs discussing projection associated with non-presuppositional vbs i.e. non-factive projective). At the same time, we have increasing evidence that projection is not all-or-nothing, but comes in degrees (cite Judith, Judith, David)

The distinction has played a critical role in formal analyses of the projectivity of the content of the complement of clause-embedding predicates. The distinction has also been invoked in analyses of the interpretation of embedded questions, as in *John knows whether it is raining*. For instance, Spector and Egré (2015) proposed that factive predicates and those for which the content of the complement is merely entailed are those that express a relation to the actual true answer, as opposed to a potential answer. Our findings are thus also relevant for such analyses, though we limit the discussion to implications for projection analyses.

Finally, for a third subset of non-factive predicates, including *announce*, *tell*, *confess*, *establish*, *inform*, *acknowledge*, *admit* and *confirm*, the content of the complement is projective but not entailed; for comments to this effect see, e.g., Reis 1973, Melvold 1991, Schultz 2003, Swanson 2012, Anand and Hacquard 2014, Spector and Egré 2015, Karttunen 2016, Tonhauser et al. 2018. We refer to these non-factive predicates as projective non-factive. To illustrate, consider the sentence in (10a) with *announce*. As pointed out by Schlenker (2010:139), if Mary is a responsible 30-year old woman, then speakers who utter (10a) may very well be taken to be committed to the content of the complement, that Mary is pregnant, i.e., this content is projective. That the content of the complement of *announce* is not entailed, and *announce* therefore is a non-factive predicate, can be seen from the fact that Mary being pregnant does not necessarily follow from (10b), for instance if Mary is a 7-year old girl.

- (10) a. Has Mary announced to her parents that she is pregnant?  
 b. Mary has announced to her parents that she is pregnant. (Schlenker 2010:139)

Similarly, Spector and Egré (2015) suggested that it is possible to take a speaker who utters (11a) with the non-factive predicate *tell* to be committed to the content of the complement, that Fred is the culprit. That this content is not entailed, and *tell* therefore is a non-factive predicate, is shown by the fact that the content does not follow from the unembedded variant in (11a) if Sue is known to be an unreliable gossip.

- (11) a. Sue didn't tell Jack that Fred is the culprit.  
 b. Sue told Jack that Fred is the culprit. (Spector and Egré 2015:1739)

In sum, entailment and projectivity are assumed to jointly distinguish factive predicates from non-factive ones: as shown in Figure 1, the content of their complement of factive predicates is defined to be both entailed and projective, whereas that of non-factive predicates is not both entailed and projective.

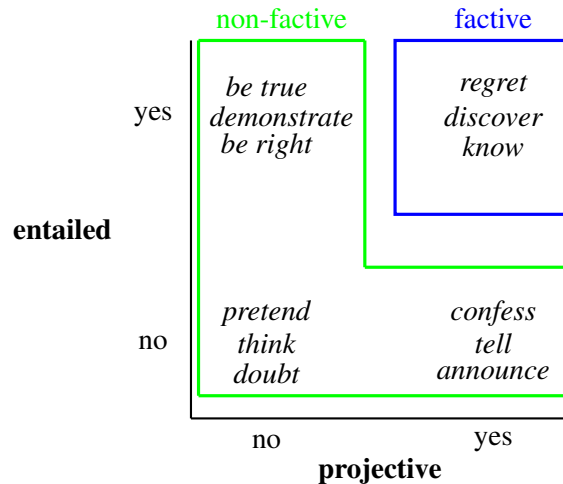


Figure 1: Illustration of the standardly-assumed division of English clause-embedding predicates into factive and non-factive ones based on whether the content of the clausal complement is entailed or projective

The distinction between factive and non-factive predicates has played a critical role in formal analyses of the projectivity of the content of the complement of clause-embedding predicates.<sup>8</sup> First, the empirical

<sup>8</sup>The distinction has also been invoked in analyses of the interpretation of embedded questions, as in *John knows whether it is raining*. For instance, Spector and Egré (2015) proposed that factive predicates and those for which the content of the complement is merely entailed are those that express a relation to the actual true answer, as opposed to a potential answer. Our findings are thus also relevant for such analyses, though we limit the discussion to implications for projection analyses.

purview of such analyses has been defined by the distinction: to date, formal analyses are largely limited to the content of the complement of factive predicates (some exceptions will be discussed later), despite the fact that the content of the complement of some non-factive predicates is also projective. This is primarily due to the fact that early analyses, like that of Heim 1983 and van der Sandt 1992, were limited to the content of the complement of factive predicates. Specifically, in these ‘lexicalist’ analyses, the content of the complement of factive predicates is lexically specified as a presupposition, thereby predicting that the content of the complement is projective over entailment-canceling operators. But because the lexical specification of the content of the complement as a presupposition also predicts that the content follows from unembedded matrix sentences, lexicalist analyses were limited to factive predicates, for which the content of the complement is both projective and entailed.

A second way in which the distinction between factive and non-factive predicates has played a critical role in formal projection analyses is that more recent analyses derive the projectivity of the content of the complement of factive predicates from its status as entailed content (e.g., Abrusán 2011, 2016, Romoli 2015, Simons et al. 2017). For instance, on Abrusán’s (2011, 2016) analysis, entailed content is by default backgrounded and therefore projective, and on Simons et al.’s (2017) analysis, questions addressed by utterances of sentences that entail the content of the complement may also entail that content, which is thereby projective. Thus, in contrast to lexicalist analyses, on which entailment and projectivity jointly define content that is lexically specified to be a presupposition, entailment is a necessary and sufficient condition to deriving projectivity on these ‘entailment-based’ analyses. (As discussed below, these analyses do not typically consider predictions made for veridical non-factive predicates.) In short, the distinction between factive predicates, for which the content of the complement is both entailed and projective, and non-factive ones, for which it is not, has critically shaped the empirical purview and nature of formal projection analyses.

In this paper, we investigate the assumed categorical distinction between factive and non-factive predicates on the basis of experiments designed to explore entailment and projectivity. The impetus for our investigation came from observations in prior literature that challenge the assumed distinction. Consider first projectivity. It has long been recognized that there is variability in the projectivity of the content of the complement of factive predicates. Karttunen (1971b), for instance, already suggested that the content of the complement of *regret* in (12a) is more projective than that of *discover* in (12b):

- (12) a. John didn’t regret that he had not told the truth.  
       b. John didn’t discover that he had not told the truth. (Karttunen 1971b:63)

More recently, Tonhauser, Beaver, and Degen (2018) showed that there is even more projection variability than previously assumed, including among presumed factive predicates. In particular, their Experiment 1b showed that the content of the complement of *be annoyed*, *be amused*, *notice*, *be aware*, *realize*, *see*, *find out* and *learn* was more projective than that of *reveal*, which in turn was more projective than that of *discover*. Because the definition of factive and non-factive predicates in (??) assumes a categorical division between projective and non-projective content, the observed projection variability raises the question of whether a distinction between projective and non-projective content can be maintained and, if yes, how projective the content of the complement of a clause-embedding predicate needs to be in order for the predicate to be considered factive. In this paper, we systematically explore the projectivity of the content of the complement of a range of English predicates presumed to be factive and non-factive.

In sum, there is disagreement in the literature about which clause-embedding predicates are such that the content of their complement is entailed.<sup>9</sup> Despite the central importance of entailment in the identification of factive predicates, there has, to date, not been any systematic investigation of the question of which

<sup>9</sup>Sudo 2012 proposed that not all presuppositions are entailments based on the observation that not all presuppositions obligatorily restrict the domain of the quantifier *exactly one*. For instance, the pre-state content of *stop* is a presupposition and an entailment because (ia) means that exactly one student who used Mac stopped using Mac, but the gender presupposition of *she* in (ib) is not an entailment because (ib) does not mean that exactly one female student criticized herself, but that exactly one student criticized



predicates are such that the content of the clausal complement is entailed.<sup>10</sup> This paper fills this empirical lacuna by exploring whether the content of the complement of a broad range of English predicates presumed to be factive and non-factive is entailed.

Our experimental exploration of projectivity and entailment for English clause-embedding predicates serves to assess whether the long-standing and widely-assumed division between factive and non-factive predicates is empirically supported. In section 2, we present the findings of Experiment 1, which was designed to explore the projectivity of the content of the complement of 20 English clause-embedding factive and non-factive predicates. In section 3, two standard diagnostics for entailment are applied in Experiments 2a and 2b, respectively, to the content of the complement of the same 20 clause-embedding predicates. On the basis of the findings from Exps. 1 and 2, we show in section 4 that factive predicates are more heterogeneous than previously assumed and argue that, contrary to assumption, the categorical distinction between factive and non-factive predicates is not empirically supported. We conclude that the empirical purview of projection analyses is broader than previously assumed and that, in particular, it is not empirically justified for analyses of the projectivity of the content of the complement of clause-embedding predicates to be limited to factive predicates.

---

herself (see Zehr and Schwarz 2016, 2018 for some experimental support). According to this diagnostic, the content of the complement of *know* is entailed content: (ic) means that exactly one student who got accepted by MIT knows that they got accepted by MIT. Whether the *exactly one* diagnostic can help settle the debate about which contents of complements of clause-embedding predicates are entailed is a question we leave for future research.

- |     |   |                |
|-----|---|----------------|
| (i) | a. Exactly one student stopped using Mac.                   | (Sudo 2012:59) |
|     | b. Exactly one student criticized herself.                  | (Sudo 2012:61) |
|     | c. Exactly one student knows that they got accepted by MIT. | (Sudo 2012:64) |

<sup>10</sup>Djäv et al. (2016) set out to investigate whether emotive and cognitive factive predicates differ with respect to whether the content of the clausal complement is entailed. However, because the factive predicates investigated in their experiments were realized in an entailment-canceling environment, namely polar questions (e.g., *Is Maria aware that Mike is moving back to Chicago?*), the content of the complement is not entailed in their experiment stimuli. Consequently, the observed differences cannot directly bear on the question of whether the two types of predicates differ in whether the content of the complement is entailed.

## 2 Experiment 1: Projection

Exp. 1 was designed to explore the projectivity of the content of the complement of 20 English clause-embedding predicates. **Following Tonhauser et al. 2018, we assume that projectivity is a gradient property of utterance content: an utterance content projects to the extent that the reader (or listener) takes the author (or speaker) to be committed to the truth of the content.**<sup>11</sup> Projectivity was explored in Exp. 1 with the ‘certain that’ diagnostic for projectivity, which has been used in previous work on projectivity (see, e.g., Tonhauser 2016, Djärv and Bacovcin 2017, Stevens et al. 2017, Tonhauser et al. 2018).<sup>12</sup> On this diagnostic, participants are presented with an utterance in which the content to be investigated occurs in an entailment-canceling environment, like a polar question, as illustrated in (13) for the content of the nominal appositive, that Martha’s new car is a BMW.

(13) Patrick asks: *Was Martha’s new car, a BMW, expensive?*

To assess whether the relevant content projects, participants are asked whether the speaker is certain of the content; for instance, in (13), participants are asked whether Patrick is certain that Martha’s new car is a BMW. We assume that the more the participant takes the speaker to be certain of the content, the more the speaker is taken to be committed to the content, i.e., the more projective the content is.

### 2.1 Methods

**Participants** 300 participants with U.S. IP addresses and at least 99% of previous HITs approved were recruited on Amazon’s Mechanical Turk platform (ages: 20-71, median: 36; 136 female, 157 male, 2 other, 3 undeclared). They were paid 75 cents for participating in the experiment.

**Materials** Exp. 1 explored the projectivity of content of the complement of the 20 clause-embedding predicates in (14). The 5 factive predicates in (14a) include the cognitive predicate *know*, the change-of-state predicates *discover* and *reveal*, the sensory predicate *see*, and the emotive predicate *be annoyed*. The 15 non-factive predicates in (14b) are divided into 3 classes, following the discussion in section 1. First, the content of the complement of the 4 plain non-factive predicates in (14b.i) is typically taken to be neither entailed nor projective. Second, the 2 predicates in (14b.ii) are veridical non-factive, i.e., the content of the complement is typically taken to be entailed (see, e.g., Anand and Hacquard 2014) but not projective. Third, the content of the complement of the 9 projective non-factive predicates in (14b.iii) has been suggested to be projective though not entailed. We included *establish* in this class, despite Wyse (2010) classifying it as factive and Swanson (2012) classifying it as veridical non-factive, because of its ability to co-occur with *falsely*, as shown in (4) and Tonhauser et al.’s (2018) observation that the content of its complement is projective. We also included *hear* in this class: even though it is often considered a factive sensory predicate (e.g., Beaver 2010, Anand and Hacquard 2014), it has been observed that *hear* also has a non-factive reportative evidential sense (see, e.g., Anderson 1986, Simons 2007), especially when it is combined with complements that describe events that cannot be auditorily perceived, as is the case in Exp. 1.

(14) 20 clause-embedding predicates

a. factive (5): *be annoyed, discover, know, reveal, see*

<sup>11</sup>We further assume that projectivity is an empirical property of utterance content that can be diagnosed with theoretically untrained native speakers. This position seems to contrast with that of Anand and Hacquard (2014), who suggested that examples in which the speaker is taken to be committed to the content of the complement of a non-factive predicate merely give rise to the “illusion of projection” (p.76).

<sup>12</sup>For other diagnostics for projectivity see, e.g., Smith and Hall 2011, Xue and Onea 2011 and Tonhauser et al. 2013; see also the discussion in Tonhauser et al. 2018.

- b. non-factive (15)
  - i. plain non-factive (4): *pretend, suggest, say, think*
  - ii. veridical non-factive (2): *be right, demonstrate*
  - iii. projective non-factive (9): *acknowledge, admit, announce, confess, confirm, establish, hear, inform, prove*

The clausal complements of the 20 predicates were realized by 20 clauses (provided in Appendix B), for a total of 400 predicate/clause combinations. These 400 predicate/clause combinations were realized as polar questions by combining them with random proper name subjects, as shown in the sample target stimuli in (15), which realize the complement clause *Julian dances salsa*. Eventive predicates, like *discover* and *hear*, were realized in the past tense and stative predicates, like *know* and *be annoyed*, were realized in the present tense. The direct object of *inform* was realized by the proper name *Sam*. The speaker of the target stimuli was realized by a proper name and displayed in bold-face. The proper names that realized the speakers, the subjects of the clause-embedding predicates and the subjects of the complement clauses were all unique.

- (15) a. **Carol asks:** Did Sandra discover that Julian dances salsa?
- b. **Paul asks:** Does Edward think that Julian dances salsa?

To assess whether participants were attending to the task, the experiment also included the 6 polar question control stimuli shown in (16). (They were also presented to the participants as utterances by a named speaker.) The content whose projectivity was explored was the main clause content: for instance, in (16a), participants were asked whether the speaker was certain that Zack is coming to the meeting tomorrow. We expected participants to rate the projectivity of these contents as very low because speakers are typically not taken to be certain about content they are asking about.

- (16) a. Is Zack coming to the meeting tomorrow?
- b. Is Mary's aunt sick?
- c. Did Todd play football in high school?
- d. Is Vanessa good at math?
- e. Did Madison have a baby?
- f. Was Hendrick's car expensive?

Each participant saw a random set of 26 stimuli: each set contained one target stimulus for each of the 20 clause-embedding predicates (each with a unique complement clause) and the same 6 control stimuli. Trial order was randomized.

**Procedure** Participants were told to imagine that they are at a party and that, on walking into the kitchen, they overhear somebody ask somebody else a question. Participants were asked to rate whether the speaker was certain of the content of the clausal complement. They gave their responses on a slider marked 'no' at one end (coded as 0) and 'yes' at the other (coded as 1), as shown in Figure 2.

After completing the experiment, participants filled out a short, optional survey about their age, their gender, their native language(s) and, if English is their native language, whether they are a speaker of American English (as opposed to, e.g., Australian or Indian English). To encourage them to respond truthfully, participants were told that they would be paid no matter what answers they gave in the survey.

**Helen asks:** *"Did Amanda discover that Danny ate the last cupcake?"*

Is Helen certain that Danny ate the last cupcake?

no
yes

Next

Figure 2: A sample trial in Experiment 1

**Data exclusion** Prior to analysis, the data from 13 participants who did not self-identify as native speakers of American English were excluded. To assess whether the remaining 287 participants attended to the task, we inspected their responses to the 6 control stimuli, for which we expected low responses. We excluded the data from 16 participants whose response means on the controls were more than 2 standard deviations above the group mean. We furthermore identified 5 participants whose variance in overall response distribution was more than 2 standard deviations below the mean by-participant variance: these participants always selected roughly the same point on the response scale. We excluded the data from these 5 participants, too, leaving the data from 266 participants (ages 20-71; median: 36; 118 female, 143 male, 2 other, 3 undeclared).

## 2.2 Results

Each predicate received 266 certainty ratings and each predicate/clause combination received between 4 and 25 ratings (mean: 13.3). The plot in Figure 3 shows the mean certainty ratings for the target stimuli by predicate, collapsing over complement clauses, as well as the mean certainty rating of the main clause control stimuli (abbreviated ‘MC’). Participants’ individual ratings are given by light gray dots. Factive predicates are given in violet, veridical non-factive predicates in blue, projective non-factive predicates in orange and plain non-factive predicates in gray. (This color scheme is used throughout the remainder of the paper.) The mean certainty ratings are consistent with impressionistic judgments reported in the literature. First, the ratings for main clause content are lowest overall, as expected for non-projective content. Second, the ratings for factive predicates are among the highest overall, suggesting comparatively high projectivity of the content of the complement. Third, the ratings for plain non-factive predicates are among the lowest overall, suggesting comparatively low projectivity of the content of the complement. Finally, the mean certainty ratings of many of the projective non-factives are higher than those of main clauses as well as of plain non-factives, confirming the intuitions of authors like Schlenker (2010), Anand and Hacquard (2014) and Spector and Egré (2015) that the content of the complement of some non-factive predicates may project.

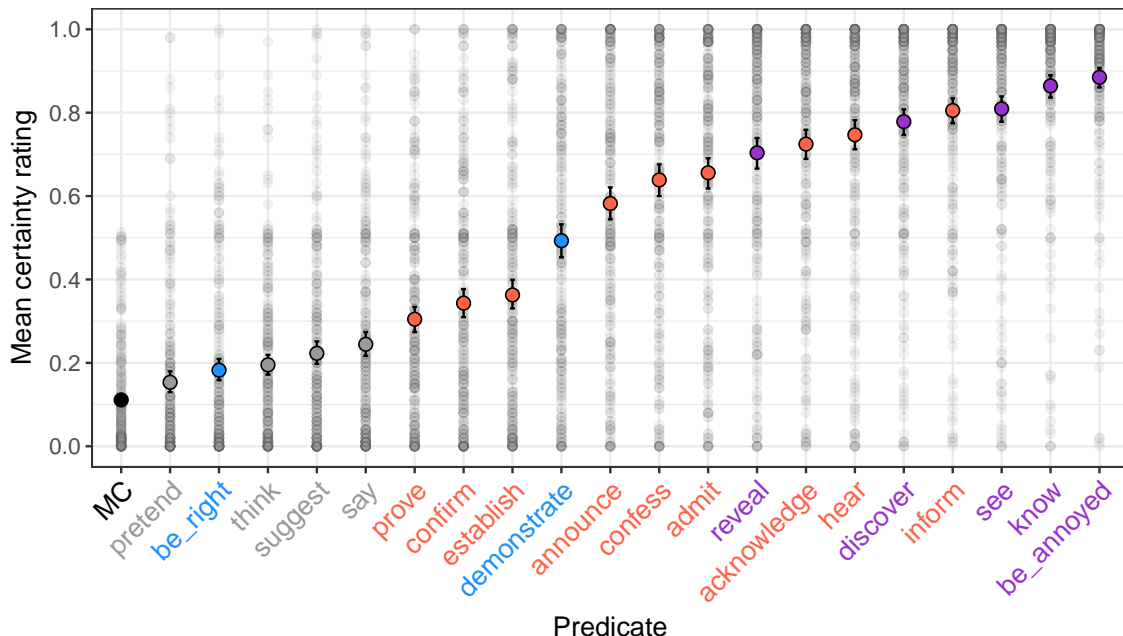


Figure 3: Mean certainty ratings by predicate, collapsing over complement clauses, with 95% confidence intervals. Factive predicates are given in violet, veridical non-factive ones in blue, plain non-factive ones in gray and projective non-factive ones in orange. Light gray dots indicate individual participants’ responses. ‘MC’ abbreviates main clause controls.

Recall from section 1 that the distinction between factive and non-factive predicates relies on two assumptions: i) content is either projective or not, and ii) the content of the complement of factive predicates is projective whereas that of non-factive predicates, with the exception of projective non-factive predicates, is not. Neither of these assumptions is empirically supported by the findings of Exp. 1. Regarding the first assumption, the certainty ratings plotted in Figure 3 provide further support for Tonhauser et al.’s (2018) finding that projectivity is a gradient property of utterance content, not a binary categorical one. With respect to by-predicate projection variability, we observe, for instance, that the mean certainty ratings for the contents of the complements of the 20 predicates are spread out from .17 for *pretend* to .86 for *be annoyed*, suggesting gradience in average projectivity. We also observe gradient within-predicate projection variability: for many predicates, the certainty ratings range from floor to ceiling, suggesting gradience in the extent to which participants took speakers to be committed to the content of the complement of particular clause-embedding predicates. For the factive predicate *reveal*, for instance, Figure 3 shows that the middle 50% of certainty ratings fell in the range from .5 to .95. In sum, projectivity is a gradient property of utterance content, not a binary categorical one.

As discussed in section 1, the finding that projectivity is a gradient property of utterance content gives rise to the question of whether projectivity distinguishes factive predicates from non-factive ones (modulo projective non-factive ones) and how projective the content of the complement of a clause-embedding predicate needs to be in order for the predicate to be considered factive.

**UPDATED ALREADY:** To determine which clause-embedding predicates differed from each other in the projectivity of the content of the complement, we conducted post hoc pairwise comparisons using Tukey’s method (allowing for by-participant and by-item variability), using the *lsmeans* package (Hothorn et al. 2008) in R (R Core Team 2016). We included the main clause controls in the analysis to identify predicates for which the content of the complement was not projective (6916 data points total). The *p*-values for each comparison pair are shown in Table 1.

	MC	<i>pretend</i>	<i>be right</i>	<i>think</i>	<i>suggest</i>	<i>say</i>	<i>prove</i>	<i>confirm</i>	<i>establish</i>	<i>demonstrate</i>	<i>announce</i>	<i>confess</i>	<i>admit</i>	<i>reveal</i>	<i>acknowledge</i>	<i>hear</i>	<i>discover</i>	<i>inform</i>	<i>see</i>	<i>know</i>
<i>pretend</i>	n.s.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>be right</i>	.	n.s.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>think</i>	**	n.s.	n.s.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>suggest</i>	***	n.s.	n.s.	n.s.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>say</i>	***	**	n.s.	n.s.	n.s.	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>prove</i>	***	***	***	**	*	n.s.	-	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>confirm</i>	***	***	***	***	***	**	n.s.	-	-	-	-	-	-	-	-	-	-	-	-	-
<i>establish</i>	***	***	***	***	***	***	n.s.	n.s.	-	-	-	-	-	-	-	-	-	-	-	-
<i>demonstrate</i>	***	***	***	***	***	***	***	***	***	-	-	-	-	-	-	-	-	-	-	-
<i>announce</i>	***	***	***	***	***	***	***	***	***	***	-	-	-	-	-	-	-	-	-	-
<i>confess</i>	***	***	***	***	***	***	***	***	***	***	n.s.	-	-	-	-	-	-	-	-	-
<i>admit</i>	***	***	***	***	***	***	***	***	***	***	.	n.s.	-	-	-	-	-	-	-	-
<i>reveal</i>	***	***	***	***	***	***	***	***	***	***	***	n.s.	n.s.	-	-	-	-	-	-	-
<i>acknowledge</i>	***	***	***	***	***	***	***	***	***	***	***	*	n.s.	n.s.	-	-	-	-	-	-
<i>hear</i>	***	***	***	***	***	***	***	***	***	***	***	**	**	n.s.	n.s.	-	-	-	-	-
<i>discover</i>	***	***	***	***	***	***	***	***	***	***	***	***	***	.	n.s.	n.s.	-	-	-	-
<i>inform</i>	***	***	***	***	***	***	***	***	***	***	***	***	***	**	*	n.s.	n.s.	-	-	-
<i>see</i>	***	***	***	***	***	***	***	***	***	***	***	***	***	***	*	n.s.	n.s.	n.s.	-	-
<i>know</i>	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	*	n.s.	n.s.	-
<i>be annoyed</i>	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	***	**	*	.	ns

Table 1: P-values associated with pairwise comparison of certainty ratings of the 20 clause-embedding predicates and the main clause controls using Tukey’s method. ‘\*\*\*’ indicates significance at .0001, ‘\*\*’ at .01, ‘\*’ at .05, ‘.’ marginal significance at .1, and ‘n.s.’ indicates no significant difference in means. ‘MC’ abbreviates main clause controls. factive predicates are given in darker blue, veridical non-factive ones in lighter blue, plain non-factive ones in brown and projective non-factive ones in black.

The analysis revealed that, for the most part, the content of the complement of plain non-factive and veridical non-factive predicates is less projective than that of factive and projective non-factive predicates. Critically, however, the observed by-predicate differences do not align with the assumed categorical distinction between, on the one hand, factive and projective non-factive predicates and, on the other hand, plain non-factive and veridical non-factive predicates. First, there are pairs of predicates that do not pattern as expected: the content of the complement of the veridical non-factive predicate *demonstrate* is more projective than that of the projective non-factive predicates *establish*, *confirm* and *prove*, and the content of the complement of the projective non-factive predicate *prove* is indistinguishable in its projectivity from that of the plain non-factive predicates *say* and *suggest*. Second, because of the observed by-predicate variability, veridical non-factive and plain non-factive predicates do not form a natural class with respect to the projectivity of the content of the complement, and neither do factive and projective non-factive predicates. That is, for the 6 veridical non-factive and plain non-factive predicates, the content of the complement of *demonstrate* is more projective than that of *be right* and of the 4 plain non-factive predicates. Likewise, for the 14 factive and projective non-factive predicates, there is widespread by-predicate variability: for instance, the content of the complement of *be annoyed* and *know* is more projective than that of *discover* and *hear*, that of *inform* is more projective than that of *acknowledge*, *reveal* and *admit*, that of *discover* is more pro-

jective than that of *reveal*, *admit* and *confess*, and that of *confess* is more projective than that of *establish*.<sup>13</sup> Thus, the findings of Exp. 1 do not support the assumed binary categorical distinction between, on the one hand, factive and projective non-factive predicates, and, on the other hand, plain non-factive and veridical non-factive predicates.

Do the findings of Exp. 1 support an alternative binary categorical distinction between the 20 predicates with respect to the projectivity of the content of the complement? We argue that this is not the case. To illustrate, we note that the pairwise comparison reveals that only the content of the complement of *pretend* is indistinguishable in projectivity from main clause content. This finding suggests that only this content is not projective and that the content of the complement of the remaining 19 predicates, including that of the other 3 plain non-factive and the 2 veridical non-factive predicates, can be considered at least weakly projective.<sup>14</sup> Now consider an alternative binary categorical division among the 20 clause-embedding predicates: the first set consists of *pretend*, *be right*, *think*, *suggest* and *say*, i.e., the 5 predicates for which the content of the complement is least projective, and the second set of the remaining 15 predicates. Neither of the two sets of predicates forms a natural class: the first set includes predicates for which the content of the complement is not projective as well as predicates for which it is weakly projective, and the second set includes predicates that exhibit significant differences in the projectivity of the content of the complement. Furthermore, the projectivity of the content of the complement of *say* and *suggest* (first set) is indistinguishable from that of *prove* (second set), which illustrates the arbitrariness of the dividing line. More generally, we argue that any binary categorical distinction between the 20 predicates is bound to be arbitrary because of the rampant by-predicate variability and the observation that the content of the complement of most of these 20 predicates is at least weakly projective.

In sum, Exp. 1 confirmed Tonhauser et al.’s (2018) finding that projectivity is a gradient property of utterance content and showed that projectivity does not categorically distinguish the content of the complement of factive (and projective non-factive) predicates from that of ‘(plain and veridical) non-factive’ predicates. The experiments described in the next section were designed to explore whether the content of the complement of the same 20 clause-embedding predicates is entailed, to identify whether entailment and projectivity jointly distinguish factive and non-factive predicates, as per the definition in (??).

## 2.3 Discussion

- projectivity alone does not distinguish factive from non-factive predicates: no non-arbitrary cut-off point

## 3 Experiment 2: Entailment

The two experiments described in this section were designed to explore whether the content of the complement is entailed by the content of unembedded matrix sentences with the 20 clause-embedding predicates. We

<sup>13</sup>These findings align with those of Tonhauser et al. 2018, who also observed that the content of the complement of *reveal* was less projective than that of *discover*, which in turn was less projective than that of *know* and *be annoyed*, and that the content of *establish* was more projective than that of *confess*. There are two differences between the findings of Exp. 1 and of Tonhauser et al.’s (2018) Exps. 1a and 1b. First, the projectivity of the content of the complement of *reveal* was significantly higher than that of *confess* in their Exp. 1b, but it is only numerically higher in Exp. 1. Second, the certainty ratings were overall higher in Tonhauser et al.’s (2018) Exps. 1 than in our Exp. 1. For instance, the content of the complement of *be annoyed*, which was the most projective in both sets of experiments, received mean certainty ratings of .96 and .92 in their Exps. 1a and 1b, respectively, but only a .86 in our Exp. 1. Similarly, the content of the complement of *discover* received mean certainty ratings of .86 and .85 in their Exps. 1a and 1b, respectively, but only a .76 in our Exp. 1. Both of these differences may be due to the fact that Tonhauser et al. (2018) primarily explored highly projective content whereas our Exp. 1 also included a wide range of less projective content: it is possible that participants’ certainty ratings were influenced by the overall projectivity of the contents explored. We leave this matter to future research.

<sup>14</sup>Spector and Egré (2015) suggested that it is not clear whether the content of the complement of *say* is projective (p.1739) but that it is definitely less projective than that of *tell*. The findings of Exp. 1 suggest that it is very weakly projective.

assume a standard definition of entailment, according to which entailment is a binary, categorical relation between two contents: given sentences  $\phi$  and  $\psi$ , the content of  $\phi$  entails the content of  $\psi$  if and only if every world in which the content of  $\phi$  is true is also a world in which the content of  $\psi$  is true. Thus, if there is even one world in which  $\phi$  is true but  $\psi$  is false, then the content of  $\phi$  does not entail the content of  $\psi$ .

Given this standard definition of the entailment relation, it is straightforward to establish that the contents of two natural language sentences  $\phi$  and  $\psi$  do not stand in an entailment relation, namely by identifying a world in which the content of  $\phi$  is true and the content of  $\psi$  is false. In this paper, however, we are also interested in establishing that the contents of two sentences  $\phi$  and  $\psi$  stand in an entailment relation. Establishing this is straightforward if the content of some expression in  $\phi$  denotes a subset of the content of some expression in  $\psi$ : for instance, in (17a), the content of  $\phi$  entails the content of  $\psi$  because the denotation of *black cat* in  $\phi$  is a subset of the denotation of *cat* in  $\psi$ . No such structural relation exists, however, between the content of a sentence with a clause-embedding predicate and the content of the complement, as in (17b). Furthermore, it is impossible to empirically verify that the content of  $\phi$  entails the content of  $\psi$  in (17b) because it is impossible to verify that  $\psi$  is true in all worlds in which  $\phi$  is true.

- (17) a.  $\phi$ : Sam owns a black cat.                       $\psi$ : Sam owns a cat.  
       b.  $\phi$ : Sam knows that it's raining.               $\psi$ : It's raining.

Thus, to establish whether the content of an unembedded matrix sentence with a clause-embedding predicate entails the content of the clausal complement, we resort to diagnostics that are taken to provide evidence for an entailment relation. Two standard diagnostics are given in (18); these were already illustrated with the emotive predicate *regret* in section 1. According to the ‘inference diagnostic’ in (18a), the content of an unembedded matrix sentence with a clause-embedding predicate entails the content of the complement if and only if the content of the complement definitely follows from a true statement of the unembedded matrix sentence. **ONLY A NECESSARY CONDITION, NOT SUFFICIENT** According to the ‘contradictoriness diagnostic’ in (18b), such an entailment relationship holds if and only an utterance of the unembedded matrix sentence that is followed by a denial of the content of the complement is contradictory.

- (18) Two diagnostics for entailment                      (see, e.g., Chierchia and McConnell-Ginet 1990:§3.1)  
       a. Inference diagnostic  
            $\phi$  entails  $\psi$  if and only if, if  $\phi$  is true, then the truth of  $\psi$  definitely follows.  
       b. Contradictoriness diagnostic  
            $\phi$  entails  $\psi$  if and only if sentences of the form  $\phi$  *and not*  $\psi$  are contradictory.

These two diagnostics for entailment are applied to the content of the complement of the 20 clause-embedding predicates in Exps. 2a and 2b, respectively. Both experiments include control stimuli for which the relevant inference definitely follows (Exp. 2a) and that are definitely contradictory (Exp. 2b). We assume that the content of the clausal complement of sentences with a clause-embedding predicate is not an entailment if responses to the content of the clausal complement are significantly different from responses to the control stimuli. According to the diagnostics, identifying contents of clausal complements that are entailments, however, requires arguing from a null result: the content of the complement of sentences with a clause-embedding predicate is an entailment if responses to the content of the complement are not significantly different from responses to the control stimuli. We present the findings of the two experiments in sections 3.1 and 3.2, and then discuss the findings jointly in section 3.3.

### 3.1 Experiment 2a: Inference diagnostic for entailment

This experiment explored whether the content of the clausal complement of the 20 clause-embedding predicates is an entailment based on the inference diagnostic for entailment in (18a). In contrast to the entailment



relationship, which is binary and categorical, the inference relationship between the contents of two sentences  $\phi$  and  $\psi$  is gradient: given a true statement  $\phi$ , the truth of  $\psi$  follows to varying degrees. To illustrate, assume that Chris wears his pajamas whenever he is at home and that he doesn't wear pajamas outside of his home. Now consider the contents of the sentences  $\phi_1$  to  $\phi_5$  in (19). Given the aforementioned assumption, the (truth of the) content of the statement  $\psi$  *Chris is wearing his pajamas* definitely follows from a true statement of  $\phi_1$ , it follows to increasingly lower degrees from true statements of  $\phi_2$  to  $\phi_4$  and it definitely does not follow from a true statement of  $\phi_5$ .

- (19)  $\phi_1$  Chris is at home.  
 $\phi_2$  Chris is likely at home.  
 $\phi_3$  Chris is possibly at home.  
 $\phi_4$  There is a slight possibility that Chris is at home.  
 $\phi_5$  Chris is not at home.

In applying the inference diagnostic for entailment in Exp. 2a, we assume that if the content of  $\phi$  entails the content of  $\psi$ , then the (truth of the) content of  $\psi$  definitely follows from true statements of  $\phi$ , i.e., the strength of the inference is at ceiling. If the content of  $\phi$  does not entail the content of  $\psi$ , then the strength of the inference is somewhere below ceiling. Exp. 2a investigated the strength of the inference from true statements of unembedded matrix sentences with the 20 clause-embedding predicates to the contents of the complements by collecting gradient inference ratings.

### 3.1.1 Methods

**Participants** 300 participants with U.S. IP addresses and at least 99% of previous HITs approved were recruited on Amazon's Mechanical Turk platform (ages: 19-69, median: 36; 152 female, 148 male). They were paid 75 cents for participating in the experiment.

**Materials** The target stimuli were 400 unembedded matrix sentences that were constructed by combining the 400 predicate/clause combinations of Exp. 1 with a random proper name subject whose gender differed from the gender of the proper name subject of the embedded clause. As illustrated in the sample stimuli in (20), these matrix sentences were presented to participants as true statements. For each of the 400 target stimuli, the inference that was tested was the inference to the content of the clausal complement: for instance, participants were asked about (20a) whether it follows that Danny ate the last cupcake and about (20b) whether it follows that Emma studied on Saturday morning.

- (20) a. **What is true:** Melissa knows that Danny ate the last cupcake.  
b. **What is true:** Jerry pretended that Emma studied on Saturday morning.

The experiment also included eight control stimuli that were used to assess whether participants were attending to the task, whether they interpreted the task correctly, as well as to identify inference ratings that were at ceiling and at floor. For the four entailing control stimuli in (21), the tested inferences are lexical entailment of the true statements: for instance, that Frederick solved the problem is a lexical entailment of the true statement in (21a) that Frederick managed to solve the problem and that the window broke is a lexical entailment of the true statement in (21b) that Tara broke the window with a bat. We expected the inference ratings for these lexical entailments to be at ceiling, i.e., to be entailments. In contrast, the inferences tested with the four non-entailing control stimuli in (22) definitely do not follow from these true statements: in (22a), that Dana wears a wig does not follow from Dana watching a movie last night, and that Madison closed the window is the negation of what actually follows from the true statement in (22c). We expected the inference ratings for these four control stimuli to be at floor.

(21) Entailing control stimuli

- a. **What is true:** Zack bought himself a car this morning. (Tested inference: Zack owns a car.)
- b. **What is true:** Tara broke the window with a bat. (Tested inference: The window broke.)
- c. **What is true:** Frederick managed to solve the problem. (Tested inference: Frederick solved the problem.)
- d. **What is true:** Vanessa happened to look into the mirror. (Tested inference: Vanessa looked into the mirror.)

(22) Non-entailing control stimuli

- a. **What is true:** Dana watched a movie last night. (Tested inference: Dana wears a wig.)
- b. **What is true:** Hendrick is renting an apartment. (Tested inference: The apartment has a balcony.)
- c. **What is true:** Madison was unsuccessful in closing the window. (Tested inference: Madison closed the window.)
- d. **What is true:** Sebastian failed the exam. (Tested inference: Sebastian did really well on the exam.)

Each participant saw a random set of 28 stimuli: each set contained one target stimulus for each of the 20 predicates (each with a unique complement clause) and the same 8 control stimuli. Trial order was randomized.

**Procedure** Participants were told that they would read true statements and that they would be asked to assess whether a second statement follows from the first, true statement. On each trial, participants read the true statement and were asked to respond to the question *Does it follow that...* where the continuation was realized by the complement of the clause-embedding predicate. Participants gave their responses on a sliding scale from ‘definitely doesn’t follow’ (coded as 0) to ‘definitely follows’ (coded as 1), as shown in Figure 4.

**What is true:** Edward proved that Grace visited her sister.

Does it follow that Grace visited her sister?

definitely doesn't follow  definitely follows

Figure 4: A sample trial in Experiment 2a

To familiarize participants with the task, they first responded to the two familiarization stimuli in (23), where the inference tested was the inference to the content of the clausal complement. Participants who rated (23a) in the lower half of the sliding scale or (23b) in the upper half of the sliding scale were given an explanation for why their answer was wrong. Participants could only advance to the 28 stimuli if they gave

a plausible rating to the two familiarization stimuli, i.e., a rating in the upper half of the sliding scale for (23a) and in the lower half for (23b).

- (23) a. **What is true:** Drew is correct that Patty lives in Canada.  
 b. **What is true:** Drew believes that Patty lives in Canada.

After responding to the 28 stimuli, participants filled out a short, optional survey about their age, their gender, their native language(s) and, if English is their native language, whether they are a speaker of American English (as opposed to, e.g., Australian or Indian English). To encourage them to respond truthfully, participants were told that they would be paid no matter what answers they gave in the survey.

**Data exclusion** Prior to analysis, the data from 14 participants who did not self-identify as native speakers of American English were excluded. For the remaining 286 participants, we inspected their responses to the 8 control stimuli. The group means for the four control stimuli in (21), for which the strength of the inference was expected to be at ceiling, was .94. The group mean for the four control stimuli in (22), for which the strength of the inference was expected to be at floor, was .05. These group means suggest that, overall, participants attended to and understood the task. The response means of 27 participants were more than 2 standard deviations below the group mean for the control stimuli in (21) or above the group mean for the control stimuli in (22). Closer inspection revealed that these participants' responses to the control stimuli were systematically higher or lower, respectively, suggesting that these participants did not attend to the task or interpreted the task differently. The data from these 27 participants were also excluded. In this experiment, we did not identify any participants who always selected roughly the same point on the response scale. The results presented below are based on data from 259 participants (ages 19-69; median: 36; 132 female, 128 male); the group means were .96 for the control stimuli in (21) and .03 for the control stimuli in (22).

### 3.1.2 Results

Each of the 20 clause-embedding predicates received 259 ratings and each of the 400 predicate/clause combinations received between 3 and 23 ratings (mean: 13). The plot in Figure 7 shows the mean inference ratings for the target stimuli by predicate, collapsing over complement clauses, as well as the mean inference ratings of the non-entailing controls (abbreviated 'non-ent. C') and the entailing controls (abbreviated 'entailing C'). Participants' individual ratings are given by light gray dots. Factive predicates are given in violet, veridical non-factive predicates in blue, projective non-factive predicates in orange and plain non-factive predicates in gray. The plain non-factive predicates *pretend*, *think*, *suggest* and *say* received among the lowest ratings overall, as expected from the assumption that the content of the complement of these predicates is not entailed. The mean inference ratings of several predicates were at or close to ceiling (recall the group mean of .96 for the entailing control stimuli in (21)), including the factive predicates *see* (.95), *discover* (.94), *know* (.93), *be annoyed* (.92) and *reveal* (.9), as well as the veridical non-factive predicate *be right* (.96), but also the projective non-factive predicates *prove* (.96), *establish* (.90), *admit* (.91), and *acknowledge* (.91). This finding suggests that participants took the truth of the content of the clausal complement of these predicates to (mostly) follow from the true statements of unembedded matrix sentences with these predicates. **NECESSARY, NOT SUFFICIENT** The same holds for the purported veridical non-factive predicate *demonstrate*, whose mean inference rating was not at ceiling (.85).

To determine which clause-embedding predicates are such that the content of the complement is entailed, i.e., for which clause-embedding predicates the strength of the inference to the content of the complement is at ceiling, we fit a Bayesian model predicting rating from predicate. Random effects for item and participant. slope for verb for the by-participant random effect

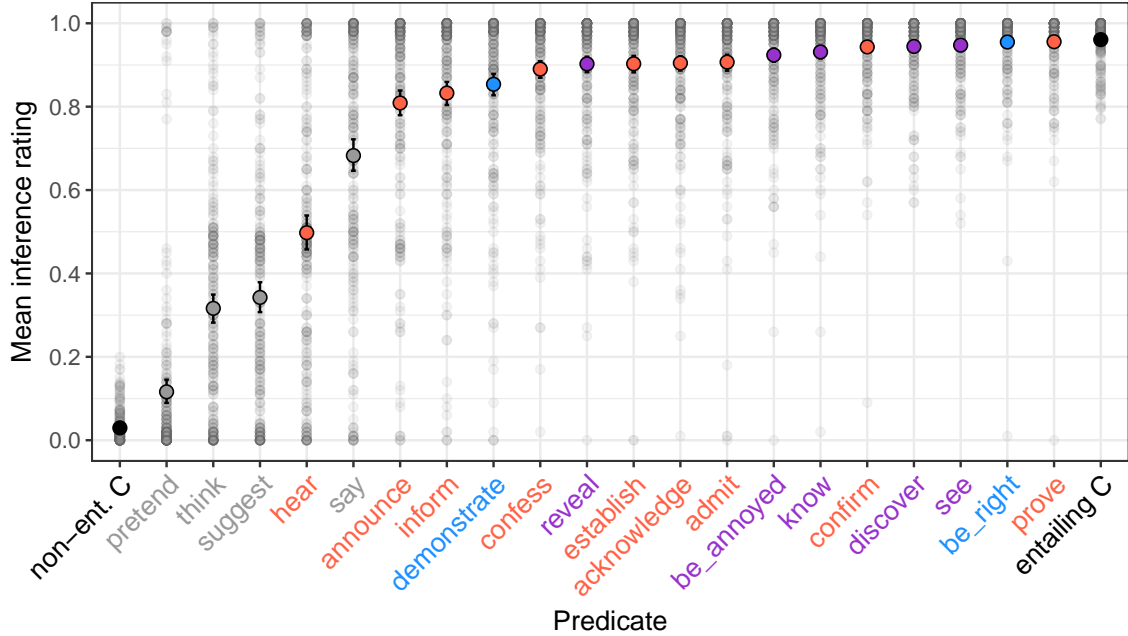


Figure 5: Mean inference rating by predicate, collapsing over complements, including the non-entailing controls (‘non-ent. C’) and the entailing controls (‘entailing C’). Error bars indicate 95% confidence intervals. Light gray dots indicate individual participants’ ratings. Factive predicates are given in violet, veridical non-factive ones in blue, projective non-factive ones in orange and plain non-factive ones in gray.

Results: which ones are indistinguishable from entailing controls?

Not significantly different from entailing controls: *be right, confirm, discover, prove, see*

### 3.2 Experiment 2b: Contradictoriness diagnostic for entailment

This experiment explored whether the content of the complement of the 20 clause-embedding predicates is an entailment based on the contradictoriness diagnostic for entailment in (18b). Participants rated the contradictoriness of utterances of English sentences of the form  $\phi$  but not  $\psi$ , where  $\phi$  is an unembedded matrix sentence with a clause-embedding predicate and  $\psi$  is its clausal complement. We collected gradient rather than binary contradictoriness ratings because we anticipated that participants’ assessment of the contradictoriness of a speaker’s utterance may depend on contextual factors: for instance, we assume that the extent to which an utterance of the sentence in (24) with the projective non-factive predicate *announce* is contradictory may depend on how trustworthy Mary is.

(24) Mary announced that she is pregnant, but she’s not.

We assume that when the content of  $\phi$  entails the content of  $\psi$ , the contradictoriness of an utterance of the form  $\phi$  but not  $\psi$  is not mitigated by contextual factors, and participants’ contradictoriness ratings will be at ceiling. If, however, the contradictoriness of an utterance of the form  $\phi$  but not  $\psi$  is mitigated by contextual factors, then the content of  $\phi$  does not entail the content of  $\psi$ , and participants’ contradictoriness ratings will be somewhere below ceiling.

### 3.2.1 Methods

**Participants** 300 participants with U.S. IP addresses and at least 99% of previous HITs approved were recruited on Amazon’s Mechanical Turk platform (ages: 18-72, median: 35; 137 female, 162 male, 1 other). They were paid 75 cents for participating in the experiment.

**Materials** In the target stimuli, the 400 predicate/clause combinations were combined with a random proper name subject and a *but*-clause that denied the truth of the content of the clausal complement. As shown by the sample stimuli in (25), the target stimuli were presented to participants as utterances by named speakers. The proper names that realized the speakers, the subjects of the 20 predicates and the subjects of the complement clauses were all unique. The gender of the proper name subject of the predicate was distinct from the gender of the proper name in the complement clause, to ensure that the pronoun in the elliptical *but*-clause unambiguously referred to the individual referred to in the complement clause of the predicate.

- (25) a. **Christopher:** “*Melissa knows that Danny ate the last cupcake, but he didn’t.*”  
b. **Susan:** “*Jerry pretended that Emma studied on Saturday morning, but she didn’t.*”

The experiment also included eight control stimuli that were used to assess whether participants were attending to the task, whether they interpreted the task correctly, as well as to identify inference ratings that were at ceiling and at floor. (These stimuli were also presented to the participants as utterances by named speakers.) For the four non-contradictory control stimuli in (26), we expected contradictoriness ratings at floor because the content of the second clause does not contradict the content of the first clause: for instance, in (26d), the content that Vanessa is good at math does not entail the negation of the content that the speaker is not good at math. In contrast, we expected contradictoriness ratings at ceiling for the four contradictory control stimuli in (27) because the first clause of each of these stimuli entails the negation of the second clause: in (27c), for instance, the content of *Madison laughed loudly* entails that Madison laughed, i.e., the negation of *she didn’t laugh*.

- (26) Non-contradictory control stimuli
- a. Zack believes that I’m married, but I’m actually single.
  - b. Tara wants me to cook for her and I’m a terrific cook.
  - c. Frederick is both smarter and taller than I am.
  - d. Vanessa is really good at math, but I’m not.
- (27) Contradictory control stimuli
- a. Dana has never smoked in her life and she stopped smoking recently.
  - b. Hendrick’s car is completely red and his car is not red.
  - c. Madison laughed loudly and she didn’t laugh.
  - d. Sebastian lives in the USA and has never been to the USA.

Each participant saw a random set of 28 stimuli: each set contained one target stimulus for each of the 20 predicates (each with a unique complement clause) and the same 8 control stimuli. Trial order was randomized.

**Procedure.** Participants were told that they would read utterances made by a speaker and were asked to assess whether the speaker’s utterance is contradictory. On each trial, participants read the speaker’s utterance and then gave their response on a slider marked ‘definitely no’ at one end (coded as 0) and ‘definitely yes’ at the other (coded as 1), as shown in Figure 6.

**Margaret:** *"Edward heard that Mary is pregnant, but she isn't."*

Is Margaret's utterance contradictory?

**definitely no****definitely yes**

Figure 6: A sample trial in Experiment 2b

To familiarize participants with the task, they first responded to the two familiarization stimuli in (28): participants who rated (28a) in the lower half of the scale or (28b) in the upper half of the scale were given an explanation for why their answer was wrong. Participants could only advance to the 28 stimuli if they gave a plausible rating to the two training stimuli, i.e., a rating in the upper half of the scale for (28b) and a rating in the lower half of the scale for (28b).

- (28) a. Drew is aware that Patty lives in Canada, but she doesn't.  
 b. Drew thinks that Patty lives in Canada, but she doesn't.

After responding to the 28 stimuli, participants filled out a short, optional survey about their age, their gender, their native language(s) and, if English is their native language, whether they are a speaker of American English (as opposed to, e.g., Australian or Indian English). To encourage them to respond truthfully, participants were told that they would be paid no matter what answers they gave in the survey.

**Data exclusion** Prior to analysis, the data from 19 participants who did not self-identify as native speakers of American English were excluded. For the remaining 281 participants, we inspected their responses to the 8 control stimuli: as expected, the group mean rating for the non-contradictory control stimuli in (26) was at floor (.08)<sup>15</sup> and the group mean rating for the contradictory control stimuli in (27) was at ceiling (.94). This finding shows that, overall, participants attended to and understood the task. The mean ratings of 18 participants were more than 2 standard deviations above the group mean for the non-contradictory control stimuli or below the group mean for the contradictory control stimuli. Closer inspection revealed that these participants' responses to the control stimuli were systematically higher or lower, suggesting that they did not attend to the task or interpreted the task differently. The data from these 18 participants were also excluded. In this experiment, we did not identify any participants who always selected roughly the same point on the response scale. The results presented below are based on data from 263 participants (ages 18-72; median: 36; 126 female, 136 male, 1 other) and resulting in new group means of .06 for the non-contradictory control stimuli in (27) and .96 for the contradictory control stimuli in (26b).

<sup>15</sup>The mean contradictoriness rating of the non-contradictory control stimulus that was of the same form as the target stimuli, namely (26a) *Zack believes that I'm married, but I'm actually single*, was higher, at .17, than the mean ratings of the remaining three non-contradictory control stimuli (means: .05 or .06) and identical to that of the target stimuli with *think*.

### 3.2.2 Results

Each predicate received 263 ratings and each predicate/clause combination received between 3 and 22 ratings (mean: 13.2). The plot in Figure 7 shows the mean contradictoriness ratings by predicate, collapsing over complement clauses, as well as the mean contradictoriness ratings of the non-contradictory controls (abbreviated ‘non-contrd. C’) and of the contradictory controls (‘contradictory C’). Participants’ individual ratings are given by light gray dots. The plain non-factive predicates *pretend*, *think*, *suggest* and *say* received among the lowest ratings overall, as expected from the assumption that the content of the complement of these predicates is not entailed. The only predicate whose mean contradictoriness rating is at ceiling is the veridical non-factive predicate *be right* (mean: .94); recall the group mean of .96 for the contradictory control stimuli in (27). This finding suggests that participants took target stimuli with *be right* to be highly contradictory. The mean contradictoriness rating for all other predicates is lower than that of the contradictory controls and *be right*, suggesting that participants did not consistently assess the target stimuli with these predicates to be contradictory.

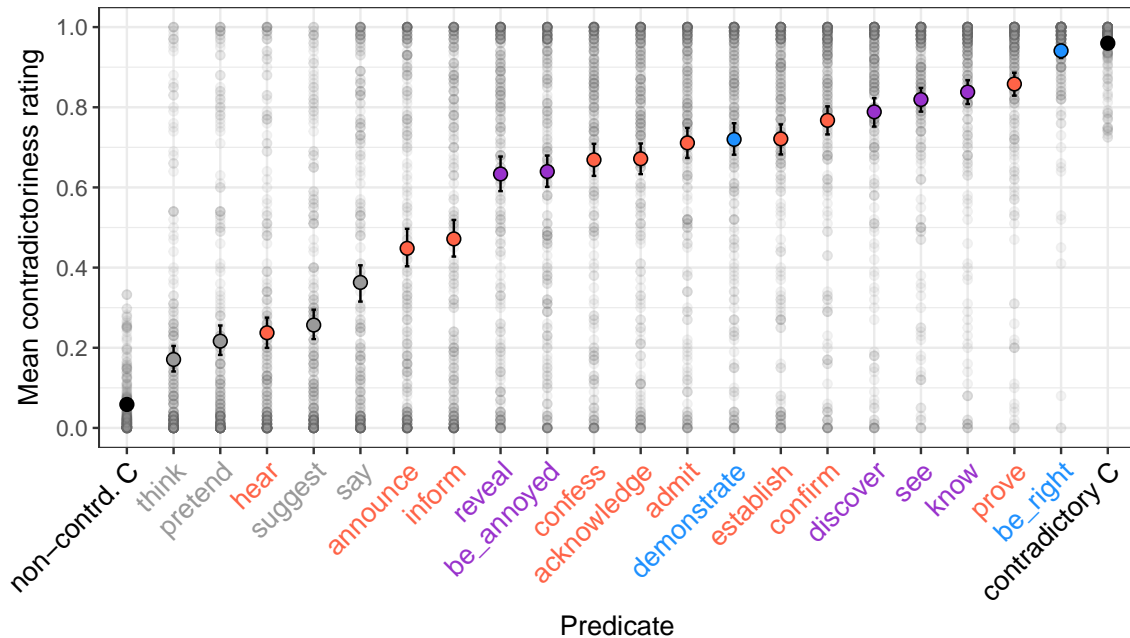


Figure 7: Mean contradictoriness rating by predicate, collapsing over complements, including the non-contradictory controls (‘non-contrd. C’) and the contradictory controls (‘contradictory C’). Error bars indicate 95% confidence intervals. Light gray dots indicate individual participants’ ratings. Factive predicates are given in violet, veridical non-factive ones in blue, projective non-factive ones in orange and plain non-factive ones in gray.

To determine which clause-embedding predicates are such that the content of the complement is entailed, i.e., for which clause-embedding predicates the strength of the inference to the content of the complement is at ceiling, we fit a Bayesian model predicting rating from predicate. Random effects for item and participant. SLOPE?

Results: which ones are indistinguishable from entailing controls?

### 3.3 Discussion

In Experiments 2 we applied two standard diagnostics for entailment to the contents of the complements of 20 clause-embedding predicates with the goal of identifying whether the contents are entailed. Given the definition of entailment, we assumed that if the content of the complement of a predicate is entailed, responses under either diagnostic are at ceiling, i.e., statistically indistinguishable from the relevant controls; otherwise, the content is not entailed.<sup>16</sup> Table 2 summarizes the findings of the two experiments: ‘✓’ indicates that the strength of the inference to the content of the complement (Exp. 2a) or the contradictoriness rating (Exp. 2b) was at ceiling; ‘×’ indicates that the strength of the inference or the contradictoriness rating was significantly lower than ceiling. As shown, the veridical non-factive predicate *be right* is the only predicate for which the responses were at ceiling in both experiments. For the factive predicates *see*, *discover*, *know* and *be annoyed*, and for the projective non-factive predicates *prove* and *confirm*, the responses in Exp. 2b were significantly lower than ceiling. For all other predicates, responses in both experiments were significantly lower than ceiling

Exp.	<i>be right</i>	<i>see</i>	<i>discover</i>	<i>know</i>	<i>be annoyed</i>	<i>prove</i>	<i>confirm</i>	<i>establish</i>	<i>admit</i>	<i>acknowledge</i>	<i>demonstrate</i>	<i>reveal</i>	<i>hear</i>	<i>confess</i>	<i>inform</i>	<i>announce</i>	<i>say</i>	<i>suggest</i>	<i>pretend</i>	<i>think</i>
2a	✓	✓	✓	✓	✓	✓	✓	×	×	×	×	×	×	×	×	×	×	×	×	×
2b	✓	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×	×

Table 2: Findings of Exps. 2a and 2b for the 20 clause-embedding predicates. factive predicates are given in darker blue, veridical non-factive ones in lighter blue, plain non-factive ones in brown and projective non-factive ones in black. ‘✓’ indicates that responses were at ceiling; ‘×’ indicates that responses were significantly lower than ceiling.

For 14 of the 20 clause-embedding predicates, the findings of the two experiments converge. First, the findings of both experiments support an analysis of the content of the complement of the veridical non-factive predicate *be right* as entailed because responses in both experiments to stimuli with this predicates were at ceiling. Second, the findings of both experiments support an analysis of 13 predicates as not entailed because responses in both experiments to stimuli with these predicates were lower than ceiling: these 13 predicates include the plain non-factive predicates *say*, *suggest*, *pretend* and *think*, the projective non-factive predicates *establish*, *admit*, *acknowledge*, *hear*, *confess*, *inform* and *announce*, but also the purportedly factive predicate *reveal* and the purportedly veridical non-factive predicate *demonstrate*. Thus, for *reveal* and *demonstrate*, the findings do not support the prevailing assumption that the content of the complement of these predicates is entailed.

For the remaining 6 clause-embedding predicates, the findings of the two experiments differ: for the factive predicates *see*, *discover*, *know* and *be annoyed* and the projective non-factive predicates *prove* and *confirm*, the inference to the content of the complement was at ceiling in Exp. 2a, supporting an analysis of

<sup>16</sup>One can, of course, imagine other linking functions for the two entailment experiments. For instance, one could assume a linking function according to which the content of the complement of a predicate is entailed if and only if the mean response to stimuli with the predicate is above .8. However, such a linking function is incompatible with the definition of entailment according to which the content of  $\phi$  does not entail the content of  $\psi$  if there is one world in which the content of  $\phi$  is true and the content of  $\psi$  is false. Furthermore, it is unclear how a non-arbitrary threshold could be chosen: why .8 rather than .85 or .75? And, finally, such a linking function would still not support the standard assumption that the content of the complement of factive and veridical non-factive predicates is entailed and that of plain non-factive and projective non-factive predicates is not: with a threshold of .8 on Exp. 2b, for instance, the content of the complement of the factive predicates *reveal*, *discover* and *be annoyed* and of the veridical non-factive predicate *demonstrate* would not be considered to be entailed, but that of the projective non-factive predicate *prove* would be.



the content of the complement as entailed, but the contradictoriness ratings were below ceiling in Exp. 2b, supporting an analysis as not entailed. Thus, for these 6 factive and projective non-factive predicates, the findings of the entailment experiments are weaker but nevertheless critical with respect to the question of which clause-embedding predicates are such that the content of their complement is entailed: the experiments do not support the standard assumption that the contents of some of these predicates are entailed and others are not entailed. Rather, these 6 predicates are indistinguishable under the two entailment diagnostics.

This finding also bears on the debate that was introduced in section 1, about whether the content of the complement of the emotive predicate *be annoyed* and the cognitive predicate *know* is entailed. Some authors, like Klein (1975), Giannakidou (1998), Schlenker (2003a) and Egré (2008), took examples like (7a) and (8b), repeated below for convenience, to show that the content is not entailed. Other authors, like Gazdar (1979a) and Abrusán (2011), did not take such examples to counter-exemplify an analysis of the content of the complement as entailed. These authors instead suggested that such examples are cases of free indirect discourse or are understood with a different concept of knowledge. Under this proposal, (7a), for instance, “reports an attitude of a subject towards facts as perceived by him” (Abrusán 2011:514). As a consequence, “the implications of [(7a)] do not have to be shared by the speaker” (*ibid.*), which accounts for the observation that the content of the complement does not follow from (7a). Similarly, Abrusán (2011:515) suggested that speakers or writers of examples like (8b) “are understood as reporting a belief or feeling of “knowledge” on somebody’s part”, but not necessarily on part of the speaker.

- (7a) Falsely believing that he had inflicted a fatal wound, Oedipus regretted killing the stranger on the road to Thebes (Klein 1975)
- (8b) Everyone knew that stress caused ulcers, before two Australian doctors in the early 80s proved that ulcers are actually caused by bacterial infection. (Hazlett 2010:501)

The findings of our experiments replicate the two positions of the debate. On the one hand, the findings of Exp. 2a show that the strength of the inference to the truth of the content of the complement of *be annoyed* and *know* is at ceiling. On the other hand, the findings of Exp. 2b show that speakers who utter sentences with such predicates need not be (taken to be) committed to the truth of the content of the complement. Advocates of the position that the content of the complement of *be annoyed* and *know* is entailed might argue that the findings of Exp. 2b do not counter-exemplify the analysis of the content as entailed because the speakers in Exp. 2b were presented as not being committed to the truth of the content of the complement (e.g., Christopher: “*Melissa knows/is annoyed that Danny ate the last cupcake, but he didn’t.*”). This, in turn, may have resulted in participants understanding these utterances as “reporting a belief or feeling of knowledge” on the attitude holder’s part. We agree that this is a possible interpretation and, therefore, the findings of our entailment experiments do not resolve the question of whether the content of the complement of *be annoyed* and *know* is entailed. Importantly, however, this interpretation would apply not only to the factive predicates we investigated in Exp. 2b, but also to the non-factive predicates *prove* and *confirm*, which patterned like the factive predicates *see*, *discover*, *know* and *be annoyed* on both entailment experiments. Thus, this interpretation would not support the standard distinction between the purportedly factive predicates *see*, *discover*, *know* and *be annoyed*, on the one hand, and the purportedly non-factive ones *prove* and *confirm*, on the other.

The two entailment diagnostics gave different results for the contents of the complements of the 20 clause-embedding predicates. Specifically, we observe that there were more lower-than-ceiling responses under the contradictoriness diagnostics in Exp. 2b than under the inference diagnostic in Exp. 2a. As a consequence, the findings of Exp. 2a support an analysis of the content of the complement as entailed for more predicates than the findings of Exp. 2b. We hypothesize that this difference is due to a difference in task demand. In order to give a not-at-ceiling response under the inference diagnostic, participants had to imagine a world in which the statement that was presented as true (e.g., *Mark informed Sam / established that Sophie*

*got a tattoo*) is true and in which the content of the complement (e.g., *Sophie got a tattoo*) is false. Evidence that this task is challenging comes from the observation that even trained researchers can fail to imagine such a world (see section 1). By contrast, to give a not-at-ceiling response under the contradictoriness diagnostic, participants merely had to assess the speaker’s utterance as not entirely contradictory.

Does the observation that the two entailment diagnostics give different results mean that one of them is more suitable to diagnosing entailment? Entailment is a semantic notion, and the two diagnostics differ in whether they force participants to consider contextual information. Specifically, on the contradictoriness diagnostic, participants must consider utterances by a speaker and the consistency of the epistemic state of this speaker, whereas, on the inference diagnostic, they consider true statements, not utterances, and no other epistemic agent. One might therefore be tempted to assume that the inference diagnostic is more suitable to diagnosing the semantic notion of entailment. However, in view of the aforementioned task demand of the inference diagnostic, it may be that the inference diagnostic may result in false positives (i.e., affirms an entailment relationship where there isn’t one) whereas the contradictoriness diagnostic may result in false negatives (i.e., denies an entailment relationship where there is one). Furthermore, as discussed in de Marneffe et al. 2012, even judgments about what follows from utterances may be influenced by pragmatic information. For instance, the presumed trustworthiness of the attitude holder may influence not just contradictoriness ratings but also inference ratings.

Resolving the question of how to best diagnose entailment with theoretically untrained speakers is not possible in the scope of this paper. We note, however, that resolving this question is of central importance for research that makes use of the notion of entailment in the analysis of the meanings of sentences with clause-embedding predicates, like projection analyses, as we discuss in the next section.

## 4 Discussion

The experiments reported on in the previous two sections were designed to investigate whether the contents of the complements of 20 English clause-embedding predicates are projective (Exp. 1) and entailed (Exps. 2), with the goal of identifying whether the long-standing and widely-assumed categorical division between factive and non-factive predicates is empirically supported. Specifically, our experiments investigated projectivity and entailment for 5 factive predicates, namely *be annoyed*, *know*, *see*, *discover* and *reveal*, as well as for 15 non-factive predicates, namely the plain non-factive predicates *pretend*, *suggest*, *say* and *think*, the veridical non-factive predicates *be right* and *demonstrate*, and the projective non-factive predicates *acknowledge*, *admit*, *announce*, *confess*, *confirm*, *establish*, *hear*, *inform* and *prove*. Given the prevailing classifications of these predicates, we expected the contents of their complements to pattern alike with respect to projectivity and entailment. However, our experiments identified significant within-class variation in both projectivity and entailment:

- **factive predicates:** The content of the complement of these predicates is expected to be projective and entailed. There is, however, no empirical support for the assumption that the content of the complement of *reveal* is entailed, and only partial support for the remaining 4 predicates. Furthermore, there is by-predicate projection variability.
- **plain non-factive predicates:** The content of the complement of these predicates is expected to be neither projective nor entailed. However, only the content of the complement of *pretend* is not projective; that of the other 3 predicates is at least weakly projective.
- **veridical non-factive predicates:** The content of the complement of these predicates is expected to be entailed but not projective. Contrary to expectation, the content of the complement of both predicates investigated is at least weakly projective, and there is no empirical support for the assumption that the content of the complement of *demonstrate* is entailed.

- **projective non-factive predicates:** The content of the complement of these predicates is expected to be projective but not entailed. However, for *prove* and *confirm*, there is support for an analysis of the content of the complement as entailed. Furthermore, there is by-predicate projection variability.

In sum, our experiments did not find empirical support for the assumption that the content of the complement of predicates in each of these four classes patterns alike. With respect to entailment, there is only partial support for many of the factive and veridical non-factive predicates that the content of the complement is entailed; for *reveal* and *demonstrate* there is no support at all, and hence these predicates cannot be assumed to be factive and veridical non-factive, respectively. Furthermore, the experiments identified significant with-class projection variability, which calls into question the naturalness of these four classes of predicates.

Do the experiments support a different classification of the 20 predicates under investigation into factive and (various types of) non-factive ones? To answer this question, consider Figure 8, which plots the mean certainty rating of the contents of the complements in Exp. 1 on the x-axis and, on the y-axis, whether the content of their complement is entailed under, in the left panel (a), the inference diagnostic of Exp. 2a and, in the right panel (b), the contradictoriness diagnostic of Exp. 2b.

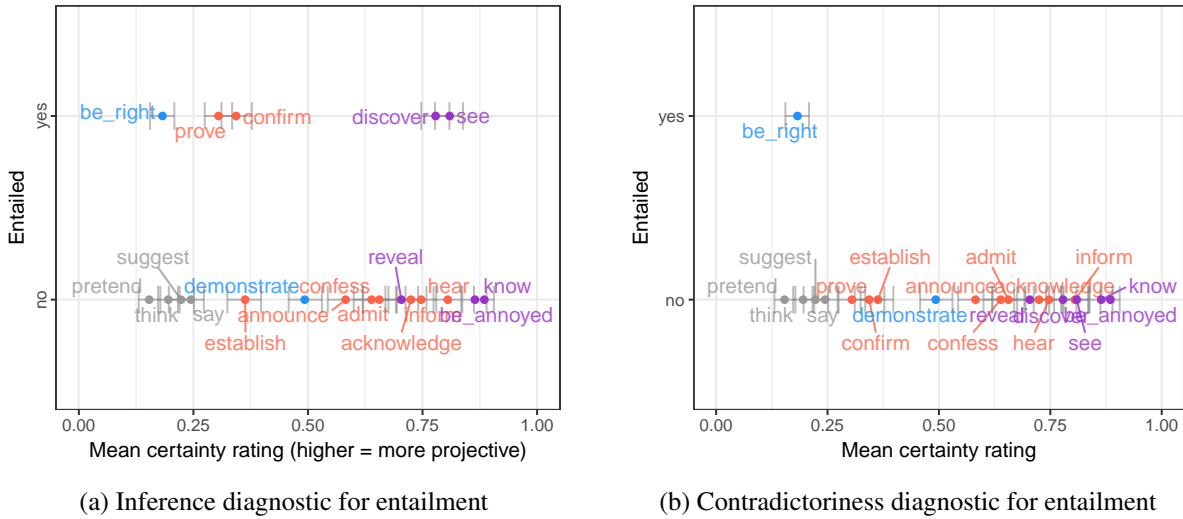


Figure 8: Properties of the contents of the complements of 20 English clause-embedding predicates: on the x-axis, the mean certainty rating (with 95% confidence intervals) and, on the y-axis, whether the content of the clausal complement is entailed under (a) the inference diagnostic of Exp. 2a or (b) the contradictoriness diagnostic of Exp. 2b. factive predicates are given in darker blue, veridical non-factive ones in lighter blue, plain non-factive ones in brown and projective non-factive ones in black.

Does a class of factive predicates emerge from the experiments? Given the findings of the inference diagnostic for entailment (left panel of Figure 8), there are four predicates for which the content of the clausal complement is both entailed and relatively highly projective, namely *discover*, *be annoyed*, *see* and *know*. Whether a predicate like *confirm*, *prove* or *be right* should count as factive depends on how projective the content of the complement needs to be in order for the predicate to be considered factive: the contents of the complements of *confirm*, *prove* or *be right* are significantly less projective than that of *discover*, *be annoyed*, *see* and *know*, but that of *discover* is also significantly less projective than that of *know* and *be annoyed*. It is unclear where a non-arbitrary line could be drawn between predicates, given that the contents of the complements of *confirm*, *prove* and *be right* are not non-projective, just less projective. Furthermore, no class of factive predicates appears to emerge given the findings of the contradictoriness diagnostic for entailment (right panel). In short, our experiments do not provide clear empirical support for a class of

factive predicates: whether there is such a class depends on how entailment is best diagnosed and how projective the content of the complement of a clause-embedding predicate needs to be to count as factive.

Given the finding that the content of the complement of *pretend* is not projective, the experiments can be taken to support a class of plain non-factive predicates for which the content of the complement is neither entailed nor projective. Which other predicates belong to this class is an open question. There is also empirical support for a class of projective non-factive predicates, i.e., predicates for which the content of the complement is projective but not entailed. In fact, this class appears to be quite large, given that the content of the complement of many more predicates than previously assumed is at least weakly projective. Of course, which predicates are projective non-factive ones again depends on the question of how entailment is best diagnosed. Furthermore, this class would be quite heterogeneous, given the rampant by-predicate projection variability, which again calls into question the naturalness of such a class of predicates. Finally, our experiments do not provide support for a class of veridical non-factive predicates, i.e., predicates for which the content of the complement is entailed but not projective. For instance, the content of the complement of *be right* is entailed but also weakly projective, and the content of the complement of *pretend* is not projective (as noted above) but also not entailed.

In sum, the experiments provide straightforward support neither for the standard classification of the 20 clause-embedding predicates into factive and (various types of) non-factive ones nor for an alternative classification. The main issues with supporting such a categorical classification are that it is not clear i) how entailment is best diagnosed and ii) how gradient projectivity can motivate a binary classification.

Resolving these issues is of critical importance for research that builds on the distinction between factive and non-factive predicates, like formal analyses of projection, as introduced in section 1. Both lexicalist analyses (e.g., Heim 1983, van der Sandt 1992) and entailment-based analyses (e.g., Abrusán 2011, 2016, Romoli 2015, Simons et al. 2017) are challenged by our findings. Lexicalist analyses, on which the content of the complement of factive predicates is lexically specified as a presupposition, must resolve which clause-embedding predicates are factive; entailment-based analyses, on which the projectivity of the content of the complement is derived from its status as entailed content, must resolve for which clause-embedding predicates the content of the complement is entailed. Furthermore, entailment-based analyses also need to confront the observation that there may be contents that are entailed but only weakly projective. For instance, given the findings of the inference diagnostic for entailment (see panel (a) of Figure 8), the content of the complement of *be right* and *confirm* is entailed but only weakly (*be right*) or mildly (*confirm*) projective. On Abrusán's (2011, 2016) analysis, entailed content is predicted to not project if projection is incompatible with the asserted content or an implicature of the assertion, or if the content addresses the grammatically signaled background question (Abrusán 2011:511, 532). On Simons et al.'s (2017) analysis, entailed content is predicted to not project if the content is not entailed by the Question Under Discussion. Whether these analyses can systematically account for entailed content that is less projective is a question for future research (see also Tonhauser et al. 2018 on this point).

The findings of Exp. 1 also show that the empirical purview of projection analyses is broader than previously assumed, given that the contents of the complements of almost all of the 20 clause-embedding predicates emerged as projective, albeit to varying degrees. To date, there are only very few formal analyses of the projectivity of the content of the complement of non-factive predicates. One attempt is presented in Schlenker 2010, who proposed that *announce* is a “part-time presupposition trigger” (p.139): “in some contexts, [*announce*, JT&JD] does not entail the truth of its complement; in other contexts, it entails and *presupposes* the truth of the complement” (*ibid.*). The notion of entailment adopted by Schlenker appears to be a context-dependent one, not the standard notion assumed here. We refer to his notion as ‘Schlenker-entailment’: in context *c*, the content of sentence  $\phi$  Schlenker-entails the content of sentence  $\psi$  if and only if the truth of  $\psi$  follows from the truth of  $\phi$  in *c*. Thus, on Schlenker's proposal, a context in which the content of the complement of *announce* projects is a context in which a true assertion of the unembedded sentence with *announce* Schlenker-entails the content. Abstracting over contexts, it follows that if the content of the

complement of *announce* is less projective than that of some other clause-embedding predicate *P*, we expect the content of the complement of *announce* to follow less strongly from true assertions of unembedded sentences with *announce* than from such assertions with *P*. This expectation is not borne out: the content of the complement of *announce* is less projective than that of *hear* (Exp. 1), but on both the inference diagnostic (Exp. 2a) and the contradictoriness diagnostic (Exp. 2b), the content of the complement of *announce* received significantly higher responses than that of *hear*. We conclude that the findings of our experiments are not predicted by Schlenker's (2010) proposal for the projectivity of the content of the complement of non-factive predicates like *announce*.

Another proposal to derive projectivity for non-factive predicates is presented by Spector and Egré (2015), who proposed (p.1736) that predicates like *tell*, *predict* and *announce* are ambiguous between a factive lexical entry on which the content of the complement is entailed and presupposed, and thereby projective, and a non-factive one on which it is not. One concern with this proposal is that it is not clear where to draw lines between predicates: which predicates only have a factive lexical entry, which ones have both a factive and a non-factive one, which ones have only have a non-factive one? Given our finding that *pretend* is the only one of the 20 clause-embedding predicates explored for which the content of the complement is not projective, a second concern with this proposal is that it may require the assumption of widespread ambiguity among clause-embedding predicates. Finally, it is not clear whether the proposal can account for the observed by-predicate projection variability. For instance, the assumption that *discover* and *be annoyed* only have a factive lexical entry does not suffice to predict that the content of the complement of *discover* is less projective than that of *be annoyed*; likewise, the assumption that *inform* and *announce* have both a factive and a non-factive lexical entry does not suffice to predict that the content of the complement of *announce* is less projective than that of *inform*.

A third proposal is sketched by Anand and Hacquard (2014): these authors suggested that speakers may be taken to be committed to the content of the complement of so-called assertive predicates like *acknowledge*, *admit* and *confirm* “because of the kind of discourse moves that these predicates report” (p.74). Specifically, utterances of sentences with these predicates can be used to report the success of an uptake in a reported common ground: for example, an utterance of a sentence like *Kim confirmed that Mary is the murderer* can report that the content *p* of the complement, that Mary is the murderer, is accepted “into the common ground of the reported discourse” (*ibid.*). The authors suggested that “[t]his acceptance of *p* can easily bleed into the actual common ground, under the assumption that no subsequent move removed *p* from the common ground” (p.74f). The authors briefly entertained an extension of this proposal to utterances of sentences with *inform* and *announce*, which convey that the content of the complement is part of the reported common ground: “to the extent that we identify ourselves with that context” (p.76), the speaker and the addressee are committed to the truth of the content of the complement. While this proposal would need to be fleshed out in more detail to be predictive, we are sympathetic to the idea that the discourse functions of utterances of sentences with clause-embedding predicates plays a role in deriving the projectivity of the content of the complement. The findings of Exp. 1 suggest, however, that more fine-grained distinctions between predicates and their discourse functions would need to be made because there are significant differences between predicates that Anand and Hacquard (2014) suggested share discourse functions: the content of the complement of *confirm* is less projective than that of *admit* and *acknowledge*, and that of *announce* is less projective than that of *inform*. In sum, there is currently no projection analysis on the market that adequately accounts for the projectivity of the content of the complement of non-factive predicates.

Where do we go from here? The two main findings of this paper are that i) projectivity and entailment, as diagnosed in this paper, do not straightforwardly support a categorical classification of predicates into factive and non-factive ones, and ii) the empirical purview of projection analyses is broader than previously assumed because the contents of the complements of many non-factive clause-embedding predicates are projective. These main findings motivate at least two fruitful ways forward, ways that are entirely compatible with one another. The first way forward is to identify whether other experimental implementations of the two

standard entailment diagnostics or an application of a different entailment diagnostic altogether can provide better empirical support for the prevailing classification of predicates into ones for which the content of the complement is entailed and ones for which it is not. If such support can be provided, it is possible to continue to derive the projectivity of entailed content through a lexicalist or an entailment-based projection analysis (with the caveat, as discussed above and in Tonhauser et al. 2018, that current analyses do not predict the observed projection variability among factive predicates).

The second way forward is to investigate how factors other than entailment are implicated in the projectivity of the content of the complement of clause-embedding predicates (and of other utterance content more generally). Factors that have already been identified in the literature as influencing the projectivity of utterance content are summarized in (29):

- (29) Factors influencing the projectivity of utterance content
- a. Contextual information, including common ground (e.g., Gazdar 1979a,b, Tonhauser et al. ms)
  - b. Prior probability of events / lexical content (e.g., Tonhauser et al. 2018)
  - c. Lexical meaning and discourse function of the attitude predicate (e.g., Anand and Hacquard 2014, Tonhauser et al. 2018)
  - d. Information structure, as marked for instance by prosody (e.g., Cummins and Rohde 2015, Tonhauser 2016, Djärv and Bacovcin 2017)
  - e. At-issueness (e.g., Simons et al. 2010, 2017, Abrusán 2011, Tonhauser et al. 2018, Mahler 2019)
  - f. Perceived degree of reliability and trustworthiness of the attitude holder (e.g., Schlenker 2010, de Marneffe et al. 2012)

As mentioned above, research on projectivity has been largely limited to entailed content. Many of the factors in (29), however, have already been shown to influence the projectivity of content regardless of whether it is entailed. Tonhauser et al. (2018), for instance, investigated the projectivity of the content of the complement of many factive predicates, but also of *establish*, *confess* and *reveal* (having assumed that the content is entailed), and found that the factors prior probability, lexical meaning and at-issueness influence projectivity. Furthermore, as discussed above, Anand and Hacquard (2014) assumed that the lexical meaning of a clause-embedding predicate influences the projectivity of the content of the complement of some non-factive predicates, and Schlenker 2010 and de Marneffe et al. 2012 assumed that the trustworthiness of the attitude holder matters for non-factive predicates. These observations give rise to the following research questions: Which factors influence the projectivity of utterance content, including the contents of complements of clause-embedding predicates, and do they influence entailed and non-entailed content differently? Addressing these questions is an exciting avenue for future research and the answers to these questions hold the key to understanding whether distinct projection analyses are needed for entailed and non-entailed content.

## 5 Conclusions

This paper investigated the long-standing and widely-adopted distinction between factive and non-factive predicates on the basis of experiments designed to explore projectivity and entailment for the contents of the complements of 20 English clause-embedding predicates. These experiments revealed that i) projectivity and entailment do not straightforwardly support a categorical classification of predicates into factive and non-factive ones, and that ii) the empirical purview of projection analyses is broader than previously assumed. We also observed that current projection analyses do not adequately predict the projectivity of the contents of the complements of non-factive predicates, primarily because such analyses have limited their attention almost entirely to the contents of the complements of factive predicates. Our research thereby

provides impetus for empirically-driven investigations of factors implicated in the projectivity of entailed as well as non-entailed content and the development of projection analyses for non-entailed content.

## A 20 complement clauses

The following clauses realized the complements of the predicates in the three experiments:

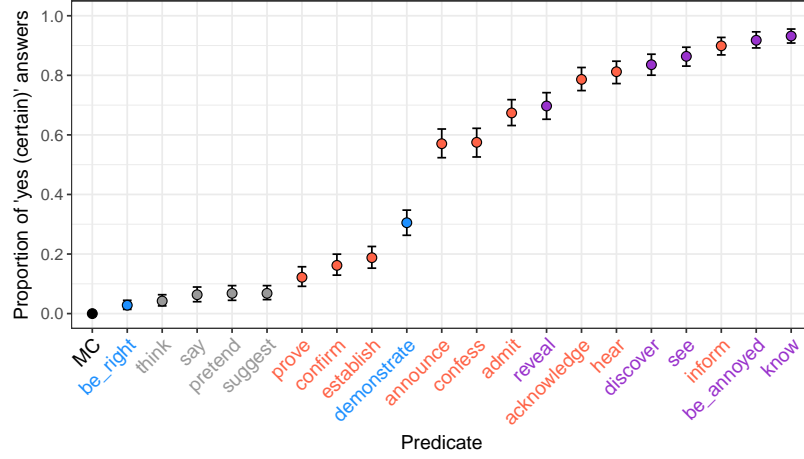
- |                                      |  |
|--------------------------------------|--|
| 1. Mary is pregnant.                 | 11. Danny ate the last cupcake.            |
| 2. Josie went on vacation to France. | 12. Frank got a cat.                       |
| 3. Emma studied on Saturday morning. | 13. Jackson ran 10 miles.                  |
| 4. Olivia sleeps until noon.         | 14. Jayden rented a car.                   |
| 5. Sophia got a tattoo.              | 15. Tony had a drink last night.           |
| 6. Mia drank 2 cocktails last night. | 16. Josh learned to ride a bike yesterday. |
| 7. Isabella ate a steak on Sunday.   | 17. Owen shoveled snow last winter.        |
| 8. Emily bought a car yesterday.     | 18. Julian dances salsa.                   |
| 9. Grace visited her sister.         | 19. Jon walks to work.                     |
| 10. Zoe calculated the tip.          | 20. Charley speaks Spanish.                |

## B Experiments with binary response tasks

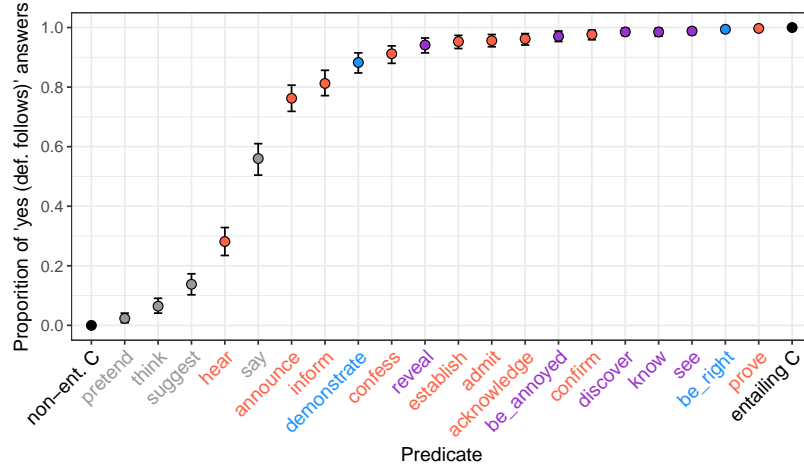
The three experiments reported on in this Appendix are identical to the three experiments in sections 2 and 3, except that two-alternative forced choice ratings ('yes', 'no') were collected rather than gradient ones. For each of the experiments, we recruited 600 participants with U.S. IP addresses and at least 99% of previous HITs approved on Amazon's Mechanical Turk platform. They were paid \$1 for participating.

Prior to analysis, we excluded the data from participants who did not self-identify as native speakers of American English. If a participant took an experiment more than once, we only kept the data from the first time they took the experiment. We also excluded the data from participants who gave a wrong rating to at least one control: a rating was considered wrong if it was a 'yes' rating on a projection control, a non-entailing control or a non-contradictory control, or a 'no' rating on an entailing control or a contradictory control. The results presented below are based on data from 426 participants (ages 18-81; median: 37; 200 female, 220 male, 2 other) in the projection experiment, from 341 participants (ages 18-73; median: 38; 180 female, 161 male) in the inference experiment and from 302 participants (ages 18-73; median: 37; 150 female, 152 male) in the contradictoriness experiment.

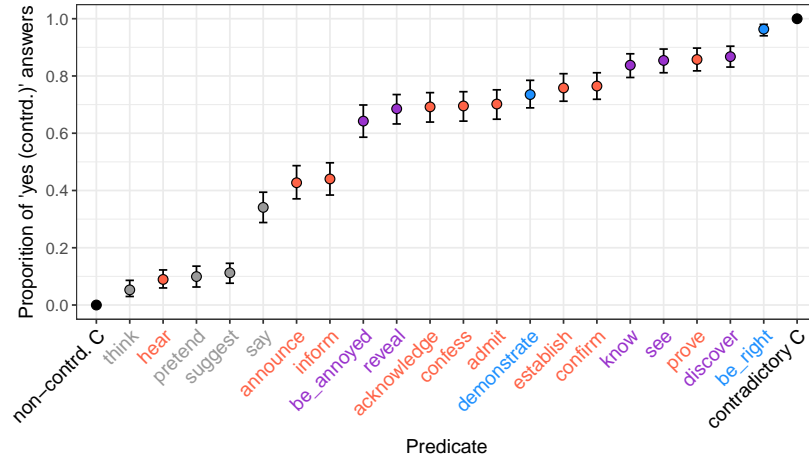
The findings of the experiments with the two-alternative forced choice task are shown in Figure A1, which show the proportion of 'yes' ratings by predicate, collapsing over complement clauses, in the projection experiment (panel a), in the entailment experiment with the inference diagnostic (panel b) and in the entailment experiment with the contradictoriness diagnostic (panel c). **SAY SOMETHING ABOUT BMRS MODELS HERE**



(a) 'certain that' diagnostic for projection.



(b) Inference diagnostic for entailment.



(c) Contradictoriness diagnostic for entailment

Figure A1: Proportion of 'yes' ratings with 95% confidence intervals, collapsing over complement clauses. Controls are given in black, factive predicates in purple, veridical non-factive ones in blue, projective non-factive ones in orange and plain non-factive ones in gray.



Figure A2 allows for a comparison of the findings of the experiments with the gradient scales in sections 2 and 3 and the findings of the experiments with the two-alternative forced choice: these plots show the mean ratings on the gradient scales by the proportion of ‘yes’ ratings in the two-alternative forced choice in the projection experiment (panel a), the entailment experiment with the inference diagnostic (panel b) and the entailment experiment with the contradictoriness diagnostic (panel c). Controls are given in black, factive predicates in purple, veridical non-factive ones in light blue, projective non-factive ones in orange and plain non-factive ones in gray. **SAY SOMETHING ABOUT HOW SIMILAR FINDINGS ARE**

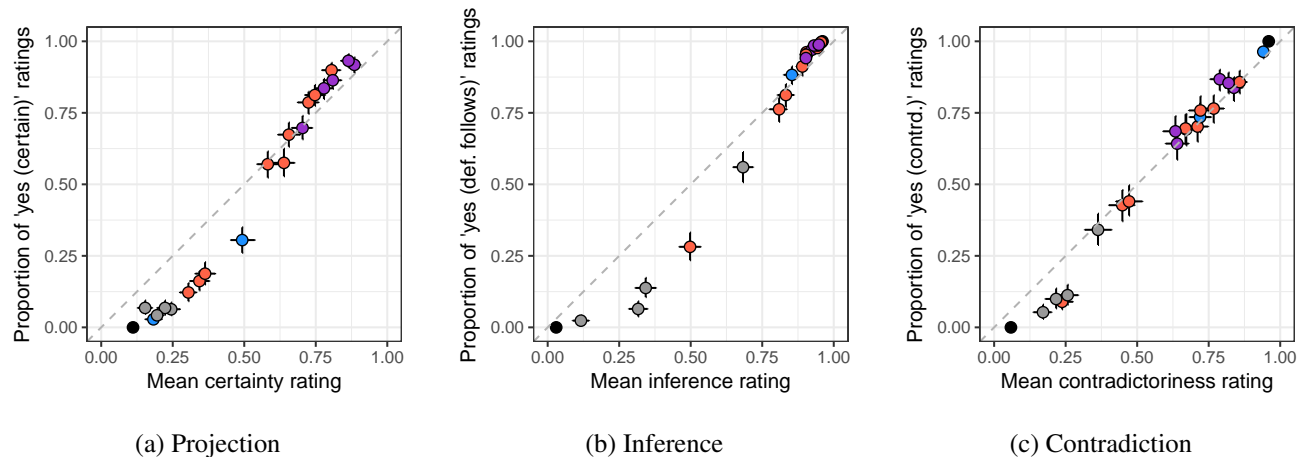


Figure A2: Mean ratings on gradient scale by proportion of ‘yes’ ratings in two-alternative forced choice, both with 95% confidence intervals, collapsing over complement clauses. Controls are given in black, factive predicates in purple, veridical non-factive ones in light blue, projective non-factive ones in orange and plain non-factive ones in gray.

### **SPELL THIS OUT IN SUFFICIENT DETAIL FOR NON-EXPERIMENTAL PEOPLE:**

Spearman rank correlations (on target stimuli):

projection: 0.983

inference: 0.996

contradictoriness: 0.985

## **References**

- Abbott, Barbara. 2006. Where have some of the presuppositions gone? In B. J. Birner and G. Ward, eds., *Drawing the Boundaries of Meaning: Neo-Gricean Studies in Pragmatics and Semantics in Honor of Laurence R. Horn*, pages 1–20. Philadelphia: John Benjamins.
- Abrusán, Márta. 2011. Predicting the presuppositions of soft triggers. *Linguistics & Philosophy* 34:491–535.
- Abrusán, Márta. 2016. Presupposition cancellation: Explaining the ‘soft-hard’ trigger distinction. *Natural Language Semantics* 24:165–202.
- Anand, Pranav and Valentine Hacquard. 2014. Factivity, belief and discourse. In *The Art and Craft of Semantics: A Festschrift for Irene Heim*, pages 69–90. MIT Working Papers in Linguistics.
- Anderson, Lloyd B. 1986. Evidentials, paths of change, and mental maps: Typologically regular asymmetries. In W. Chafe and J. Nichols, eds., *Evidentiality: The Linguistic Coding of Epistemology*, pages 273–312. New Jersey: Ablex Publishing Corporation.
- Beaver, David. 2001. *Presupposition and Assertion in Dynamic Semantics*. Stanford, CA: CSLI Publications.

- Beaver, David. 2010. Have you noticed that your belly button lint colour is related to the colour of your clothing? In R. Bäuerle, U. Reyle, and E. Zimmermann, eds., *Presuppositions and Discourse: Essays Offered to Hans Kamp*, pages 65–99. Oxford: Elsevier.
- Chierchia, Gennaro and Sally McConnell-Ginet. 1990. *Meaning and Grammar*. Cambridge, MA: MIT Press.
- Cummins, Chris and Hannah Rohde. 2015. Evoking context with contrastive stress: Effects on pragmatic enrichment. *Frontiers in Psychology* 6:Article 1779.
- de Marneffe, Marie-Catherine, Christopher D. Manning, and Christopher Potts. 2012. Did it happen? The pragmatic complexity of veridicality assessment. *Computational Linguistics* 38:301–333.
- Djärv, Kajsa and Hezekiah Akiva Bacovcin. 2017. Prosodic effects on factive presupposition projection. In *Semantics and Linguistic Theory (SALT) XXVII*, pages 116–133. Ithaca, NY: CLC Publications.
- Djärv, Kajsa, Jérémy Zehr, and Florian Schwarz. 2016. Cognitive vs. emotive factives: An experimental differentiation. In *Proceedings of Sinn und Bedeutung 21*, pages 367–385.
- Egré, Paul. 2008. Question-embedding and factivity. In L. Lihoreau, ed., *Grazer Philosophische Studien*, pages 85–125. Amsterdam: Rodopi.
- Gazdar, Gerald. 1979a. *Pragmatics: Implicature, Presuppositions, and Logical Form*. New York: Academic Press.
- Gazdar, Gerald. 1979b. A solution to the projection problem. In C.-K. Oh and D. Dineed, eds., *Syntax and Semantics 11: Presupposition*, pages 57–89. New York: Academic Press.
- Giannakidou, Anastasia, ed. 1998. *Polarity sensitivity as (non)veridical dependency*. Dordrecht: John Benjamins.
- Hazlett, Allan. 2010. The myth of factive verbs. *Philosophy and Phenomenological Research* 80:497–522.
- Heim, Irene. 1983. On the projection problem for presuppositions. In M. Barlow, D. Flickinger, and M. Westcoat, eds., *West Coast Conference on Formal Linguistics (WCCFL) 2*, pages 114–125.
- Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9:183–221.
- Hothorn, Torsten, Frank Bretz, and Peter Westfall. 2008. Simultaneous inference in general parametric models. *Biometrical Journal* 50(3):346–363.
- Karttunen, Lauri. 1971a. Implicative verbs. *Language* 47:340–358.
- Karttunen, Lauri. 1971b. Some observations on factivity. *Papers in Linguistics* 4:55–69.
- Karttunen, Lauri. 2016. Presupposition: What went wrong? In *Proceedings of Semantics and Linguistic Theory (SALT) XXVI*, pages 705–731. Ithaca, NY: CLC Publications.
- Kiparsky, Paul and Carol Kiparsky. 1970. Fact. *Progress in Linguistics* pages 143–173.
- Klein, Ewan. 1975. Two sorts of factive predicates. *Pragmatics Microfiche* it 1.1. frames B5C14.
- Mahler, Taylor. 2019. Does at-issueness predict projection? It's complicated! In *49th Annual Meeting of the North East Linguistic Society*.
- Melvold, Janis. 1991. Factivity and definiteness. In L. Cheng and H. Demirdache, eds., *More Papers on Wh-movement*, vol. 5, pages 97–117. MIT Working Papers in Linguistics.
- R Core Team. 2016. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Reis, Marga. 1973. Entanglement on factives. *Linguistic Inquiry* 4:261–271.
- Romoli, Jacopo. 2015. The presuppositions of soft triggers are obligatory scalar implicatures. *Journal of Semantics* 32:173–291.
- Schlenker, Philippe. 2003a. The lazy Frenchman's approach to the subjunctive. In *Proceedings of Romance Languages and Linguistic Theory*, pages 269–309.
- Schlenker, Philippe. 2003b. A plea for monsters. *Linguistics & Philosophy* 26:29–120.
- Schlenker, Philippe. 2010. Local contexts and local meanings. *Philosophical Studies* 151:115–142.
- Schultz, Petra. 2003. *Factivity: Its Nature and Acquisition*. Tübingen: Niemeyer.

- Simons, Mandy. 2007. Observations on embedding verbs, evidentiality, and presupposition. *Lingua* 117(6):1034–1056.
- Simons, Mandy, David Beaver, Craige Roberts, and Judith Tonhauser. 2017. The Best Question: Explaining the projection behavior of factive verbs. *Discourse Processes* 3:187–206.
- Simons, Mandy, Judith Tonhauser, David Beaver, and Craige Roberts. 2010. What projects and why. In *Semantics and Linguistic Theory (SALT) XXI*, pages 309–327. Ithaca, NY: CLC Publications.
- Smith, E. Allyn and Kathleen Currie Hall. 2011. Projection diversity: Experimental evidence. In *Proceedings of the 2011 ESSLLI Workshop on Projective Content*, pages 156–170.
- Spector, Benjamin and Paul Egré. 2015. A uniform semantics for embedding interrogatives: An answer, not necessarily the answer. *Synthese* 192:1729–1784.
- Stevens, Jon Scott, Marie-Catherine de Marneffe, Shari R. Speer, and Judith Tonhauser. 2017. Rational use of prosody predicts projectivity in manner adverb utterances. In *39th Annual Meeting of the Cognitive Science Society*, pages 1144–1149.
- Sudo, Yasutada. 2012. *On the semantics of phi-features on pronouns*. Ph.D. thesis, MIT.
- Swanson, Eric. 2012. Propositional attitudes. In *Semantics: An International Handbook of Natural Language Meaning*, vol. 2, pages 1538–1561. Berlin: Mouton de Gruyter.
- Tonhauser, Judith. 2016. Prosodic cues to speaker commitment. In *Semantics and Linguistic Theory (SALT) XXVI*, pages 934–960. Ithaca, NY: CLC Publications.
- Tonhauser, Judith, David Beaver, and Judith Degen. 2018. How projective is projective content? Gradience in projectivity and at-issueness. *Journal of Semantics* 35:495–542.
- Tonhauser, Judith, David Beaver, Craige Roberts, and Mandy Simons. 2013. Toward a taxonomy of projective content. *Language* 89:66–109.
- Tonhauser, Judith, Judith Degen, Marie-Catherine de Marneffe, and Mandy Simons. ms. Evaluative adjective sentences: A question-based analysis of projection. Ms., The Ohio State University, Stanford University, Carnegie Mellon University.
- van der Sandt, Rob. 1992. Presupposition projection as anaphora resolution. *Journal of Semantics* 9:333–377.
- White, Aaron S., Valentine Hacquard, and Jeffrey Lidz. 2018. Semantic information and the syntax of propositional attitude verbs. *Cognitive Science* 42:416–456.
- White, Aaron S. and Kyle Rawlins. 2018. The role of veridicality and factivity in clause selection. In *48th Meeting of the North East Linguistic Society*, page ??
- Wyse, Brendan. 2010. *Factive/non-factive predicate recognition within Question Generation systems*. Master’s thesis, Open University.
- Xue, Jingyang and Edgar Onea. 2011. Correlation between projective meaning and at-issueness: An empirical study. In *Proceedings of the 2011 ESSLLI Workshop on Projective Content*, pages 171–184.
- Zehr, Jérémy and Florian Schwarz. 2016. Entailed vs. non-entailed presuppositions: An experimental assessment. In *Proceedings of North East Linguistic Society 46*, pages 319–328.
- Zehr, Jérémy and Florian Schwarz. 2018. Returning to non-entailed presuppositions again. In *Proceedings of Sinn und Bedeutung 22*, pages 463–480.