Why So Hard?
A Developmental Study of Why-Questions

DANIELLA PRESTI

New York University
Department of Linguistics
May 2020
# Table of Contents

Abstract 3

1 Introduction 4

2 Why in the Syntax: Theoretical Bases 6
   2.1 Why Cross-linguistically 6
   2.2 Why-Suggestions and Polarity Items 8
   2.3 Why and How Come 18

3 Adult Pilot 23
   3.1 Background 23
   3.2 Methods 24
      3.2.1 Comprehension Task 25
      3.2.2 Grammaticality Judgments 27
   3.3 Results 29
      3.3.1 Part 1 29
      3.3.2 Part 2 31
   3.4 Discussion 33

4 Why-questions and Children: Developmental Bases 35
   4.1 Background 35
   4.2 Corpus work 36
   4.3 Discussion 38

5 Child Experimental Study 39
   5.1 Background 39
   5.2 Methods 40
      5.2.1 Comprehension 40
      5.2.2 Production 42

6 Conclusion 43

Acknowledgments 45

References 46

Appendix 49
   A: Experiment 1 Items 49
   B: Experiment 2 Items 59
Abstract

Syntacticians disagree over whether why-questions undergo the same wh-movement as other wh-questions, or whether they merge higher in the left periphery. Additionally, for those who adhere to the latter theory, there is disagreement about where exactly why merges and whether or not it undergoes further movement. Why-questions have not been subject to experimental investigation in either adults or children. This thesis first reviews recent syntactic analyses of why, as well as a longitudinal study of first language acquisition of why-questions, both in contrast to other wh-questions. New diagnostics are presented to provide a more precise analysis of why. Based on these newly posited theoretical assumptions, as well as corpus investigations, I present two experimental studies. The first tests arguments from Collins (1991) that why-questions with quantifiers and embedded clauses are ambiguous, whereas how come questions of the same kind are not, suggesting that the two questions have distinct syntactic analyses. The first study provides evidence that, contrary to previous theoretical claims (and to the hypothesis), adults do not have fewer available interpretations of how come questions in contrast to why-questions with quantifiers and embedded clauses. In fact, there is no discernible systematic difference between the two questions at all in the tested contexts. The second study targets the acquisition of the same ambiguous type of why-questions as were tested in adults, and also attempts to capture a more holistic snapshot of the language acquisition process by testing children’s comprehension and production abilities concurrently. This thesis thus provides insight into question acquisition, as well as into how the syntax of why may be represented across development and in adulthood.
1 Introduction

This Honors Thesis investigates the syntax and interpretation of why compared to other wh- and how come questions in adults and children. English-speaking adults are examined to determine whether ambiguous why- and how come questions have distinct interpretations, testing claims made in theoretical literature that why and how come have distinct syntactic analyses which affects possible interpretations, and which raises questions of the development of these constructions. For example, questions like (1) which include quantifiers like everyone could be asking for a collective reason, like in (1a), or individual reasons, like in (1b), for doing something, and questions like (2) which include an embedded predicate could be asking for a reason about the matrix verb, as in (2a), or the embedded clause, as in (2b).

(1) Why did everyone go to the beach?
   (a) They wanted to go swimming.  COLLECTIVE
   (b) John wanted to take pictures of nature, Mary wanted to sit in the sun, Sandy wanted to go for a walk…

(2) Why did Mary say she was tired?
   (a) Her friend asked her.  MATRIX
   (b) She didn’t sleep well last night.  EMBEDDED

Novel adult experimental results are used to lay the foundation for a developmental study investigating children’s comprehension of ambiguities in why-questions. Such questions are of interest because children produce adult-like why-questions far later than they do other wh-questions (Thornton 2008, Thornton 2016, Wode 1971), and cross-linguistic analyses suggest that why-questions are syntactically different from other wh-questions (Ko 2005, Rizzi
2001, Shlonsky and Soare 2011), perhaps because they ask for an entire proposition as a response, as in (5a), in contrast to other wh-questions which instead ask for phrasal responses, like DPs in (3a) and (4a).

(3) What do you want?
   (a) A coffee.

(4) Where are you going?
   (a) The store.

(5) Why do you want a coffee?
   (a) Coffee wakes me up.

This paper will begin with an overview of the theoretical literature about: (a) cross-linguistic examples which indicate that why- does not have the same surface-structure as other wh-questions, like (6), and how they interact with polarity items, (b) a syntactic comparison of why- and how come questions, (c) why-questions, discussing why-suggestions like (6), and (d) child data which demonstrates that some children persistently produce non-adult-like why-questions even after attaining adult-like patterns for the rest of their questions. All of these examples provide insight into the syntax of why-questions, which are explicitly contrasted with other wh-questions throughout.

(6) Why don’t we go get some coffee?

A syntactic analysis is presented in section 2.2, based on the aforementioned theoretical claims, which will be assumed as a basis for experimental investigation into adults’ and children’s interpretation of why-questions. However, Collins (1991), which discusses the difference in why- and how come questions provides a small sample of evidence consisting of
judgments by linguists. The additional need to have an adult baseline before engaging in analysis of children’s performance required an adult pilot study which investigated whether *how come* questions have more limited interpretations than *why*-questions in two contexts: pair-list/collective reason contexts such as (1), and embedded question contexts such as (2). These contexts are discussed in further detail in sections 2.3 and 3.1. The results of the adult pilot, presented in section 3.3, indicate that adults do not seem to have the distinctions which were argued for in the literature (i.e. non-ambiguity in *how come* questions and ambiguity in *why*-questions), although they confirm ambiguity in *why*-questions. Therefore, the child study tests whether children share these adult patterns.

The proposed child study and experimental design are presented, as well as a qualitative discussion of some child pilot data. The child study consists of two parts: a comprehension experiment and a production experiment, in order to gain insight into the acquisition path for these two sides of the same coin. Particularly, it is of interest whether children reach adult-like comprehension capabilities before they produce adult-like utterances in the same domain.

2 *Why in the Syntax: Theoretical Bases*

2.1 *Why* Cross-linguistically

To distinguish *why* from other wh-words, we can first turn to its semantic character, which differs from that of other wh-words. *Why* minimally asks for a proposition such as (7a), whereas other wh-words minimally ask for smaller phrases like DPs, as seen in (8a), or VPs.

(7) Why did James run away?

(a) He was afraid.
(8) Where did James go?

(a) The lodge.

Cross-linguistically, we see evidence that why-questions may differ from other wh-questions. Thornton presents examples from Ko (2005) to illustrate how why differs from other wh-words in wh-in-situ languages. In Japanese, question words cannot generally be preceded by scope-bearing elements, such as a negative polarity items or only phrases. (9) illustrates the ungrammaticality of this word order, and (10) shows the correct word order.

(9) *Taroo-sika nani-o yoma-nakat-ta no?  (Ko 2005)
   Taroo-only what-ACC read-NEG-PST Q
   ‘What did only Taroo read?’

(10) Nani-o Taroo-sika yoma-nakat-ta no?  (Ko 2005)
    what-ACC Taroo-only read-NEG-PST Q
    ‘What did only Taroo read?’

However, why-questions are grammatical with either word order, as in (11) and (12).

(11) Taroo-sika naze sono hon-o yoma-nakat-ta no?  (Ko 2005)
    Taroo-only why that book-ACC read-NEG-PST Q
    ‘Why did only Taroo read that book?’

    why Taroo-only that book-ACC read-NEG-PST Q
    ‘Why did only Taroo read that book?’
In Italian, T-to-C movement is obligatory for all wh-words except why, for which it is optional, as in (15). This illustrates another asymmetry between wh-words and why\(^1\).

(13) Che cosa ha fatto Gianni? (Rizzi 2001)

what has done Gianni

‘What did Gianni do?’

(14) *Che cosa Gianni ha fatto? (Rizzi 2001)

what Gianni has done

‘What did Gianni do?’

(15) Perché Gianni è venuto? (Rizzi 2001)

why Gianni has come

‘Why did Gianni come?’

(16) Perché è venuto Gianni? (Rizzi 2001)

why has come Gianni

‘Why did Gianni come?’

These examples serve to motivate the hypothesis that why-questions are syntactically distinct from other wh-questions in multiple languages.

2.2 *Why*-Suggestions and Polarity Items

The way that polarity items interact with *why*-suggestions, a particular subtype of *why*-questions, like (17), helps to illustrate a problem for analyses of *why*-questions in contrast to those of other wh-questions with the same polarity items.

\(^{1}\) Note that the following data from Rizzi (2001) includes different verbs, (*have* in (13) and (14) and *be* in (15) and (16)), so the data does not consist of minimal pairs.
(17) Why don’t we go out to dinner tonight?

Positive polarity items such as *something* in (18) are licensed in generally affirmative contexts, whereas negative polarity items such as *anything* in (19) are licensed in negative contexts. Though it is an oversimplification, for the purposes of this discussion we can more specifically define NPIs as scoping under negation, and as PPIs being unable to do so (see Szabolcsi 2004: 1).

(18) Laura hid something from Donna.

(19) Laura didn’t hide anything from Donna.

A simple definition of negation will be useful here as well. The negated version of a proposition can be understood to have the reverse truth value of its non-negated version (Miestamo 2007: 1). For example, the truth value of the proposition that Agent Cooper likes Norma’s cherry pie is true in (20) and false in (21).

(20) Agent Cooper likes Norma’s cherry pie.

(21) Agent Cooper does not like Norma’s cherry pie.

Polarity items and negation do not act as straightforwardly in interrogative contexts as they do in declarative contexts. First let’s focus on just negation in questions. Interrogative statements do not have truth values, but one approach to this problem views interrogative meanings in terms of the truth values of their answers (Guerzoni 2003, Hamblin 1973, Karttunen 1977). If we adopt this point of view, it is clear that a negated interrogative statement will not necessarily have an answer that is opposite in truth value from a non-negated interrogative statement. For example, question (23) is the negated version of question (22), but one would expect both of them to be
answered affirmatively if the responder does indeed like cherry pie, and negatively if the responder does not.

(22) Do you like cherry pie?

(23) Don’t you like cherry pie?²

Rather, a contrast between the two sentences exists in terms of what they presuppose. Sentence (22) does not presuppose anything about the status of the responder’s taste for cherry pie, whereas (23) presupposes that the responder does like cherry pie (see Siddiqui 1977: 61), and is asking for either confirmation or denial. Thus, there is a clear contrast between negation in declarative and interrogative contexts: while negation in declarative contexts reverses the truth value of a proposition, negation in interrogative contexts need not have as strong of a semantic effect, and may rather have a subtle pragmatic effect.

Similarly to how negated and non-negated interrogative sentences can become quite close to each other in meaning, in interrogative contexts both PPIs and NPIs can occur in the same context and be well-formed, without much difference in meaning (Siddiqui 1977: 87). For example, sentences (24) and (25) seem to have the same (or a very similar) meaning, despite having a PPI and an NPI, respectively. Sentences (26) and (27) are negated, and we see the same pattern in which there is not a clear contrast between each of the sentences. Moreover, sentences (24)-(27) are all approximately identical in meaning, although (26) and (27) have the additional presupposition that the proposition is true, as we already discussed of sentence (23).

(24) Did Audrey overhear something secret last night?

(25) Did Audrey overhear anything secret last night?

² Note that when the negation is not cliticized, e.g. ‘Do you not like cherry pie?’ the interpretation of the sentence may not include the presupposition on the part of the speaker that the responder does like cherry pie.
Didn’t Audrey overhear something secret last night?

Didn’t Audrey overhear anything secret last night?

Thus far, I have discussed the patterning of polarity items in the context of questions that ask for a yes or no answer, but polarity items can also occur in wh-questions. Unlike non-wh-questions, such as the ones in (22)-(27), there is a clear contrast in the meaning of negated and non-negated wh-questions. Potentially due to the unique syntactic character of why-questions (to be discussed shortly), there is an interesting puzzle in the way that negated why-questions pair with polarity items.

Why-questions presuppose propositions with truth values, and this time there is a contrast in meaning between the negated and non-negated versions, stemming from what the questions presuppose. Sentence (28) presupposes that Harry and Andy had coffee and donuts for breakfast, and sentence (29) presupposes that they did not. Each question asks for a reason as to why the presupposed proposition has the truth value that it does.

Why-questions can pair with PPIs when they presuppose affirmative propositions as in (30), and NPIs when the presuppose negative propositions as in (31).

Why-questions can pair with either NPIs or PPIs without much contrast in meaning, as in (30) and (31), but non-negated why-questions can only appear with PPIs such as (30), not with NPIs such as (33).
(32) Why didn’t Leo hide something from Shelly?
(33) */? Why did Leo hide anything from Shelly?

The following observation complicates things further: although we have already established that negated why-questions do not have a contrast in meaning dependent on the polarity item that they pair with, as in (31) and (32), there actually is a context in which we see a clear contrast in meaning among negated why-questions with different polarity items. (34) is a question similar to the ones that we have already seen. However, (35) is a suggestion, and is distinct in meaning from (34). While both questions presuppose that you are inactive in some way, (34) is asking for a reason for the inaction and (35) is suggesting that you stop being inactive. The observation that these two sentences differ in meaning is interesting because the polarity item appears to be changing the interpretation of the sentences.

(34) Why don’t you do anything (about it)?
(35) Why don’t you do something (about it)?

In addition, it is curious that sentence (34) is not a suggestion like (35) if we compare them with sentences (36) and (37), which do not have the word why. Both sentences pragmatically approximate suggestions or requests, and share similar meanings to each other despite pairing with different polarity items.

(36) Won’t you do something (about it)?
(37) Won’t you do anything (about it)?

The distinction also does not apply to other kinds of wh-questions. Consider examples (38)-(41), which are the non-why versions of (30)-(33), respectively. The crucial sentences to look at are

---

3 This is okay to me in the context of already knowing that Leo hid something (or things) that he shouldn’t have, contrary to expectations.
the two sentences with negation, (39) and (41). There is not an obvious semantic contrast with these sentences as there is with (35) and (36), and no suggestions arise from any of these examples.

(38) Who hid something from Shelly?
(39) Who didn’t hide anything from Shelly?
(40) *Who hid anything from Shelly?
(41) Who didn’t hide something from Shelly?

Additionally, why-suggestions also occur in Italian with the same types of morphemes as in English, namely, why, a negative morpheme, and a word which functions like a PPI. The same asymmetry with these why-suggestions and with why-questions also shows up if we include a Neg-word, which would correspond to an NPI. Examples (42)-(45) illustrate this, and also further demonstrate the optionality of T-to-C movement with why in Italian.

(42) Perché tu non fai qualcosa?
   why you no do something
   ‘Why don’t you do something?’
(43) Perché non fai qualcosa tu?
   why no do something you
   ‘Why don’t you do something?’
(44) perché tu non fai niente?

---

4 My informant clarified that intonation matters a lot in the interpretation of (42) and (43); depending on if it has a declarative intonation or an interrogative intonation, it can be interpreted as either a suggestion or a question. I set this aside for now, and just focus on the fact that the suggestion reading is possible in (42) and (43), whereas (44) and (45) are unambiguously questions. However, given that the ambiguity seems to be less apparent (to me at least) in both English and Spanish, it is worth collecting more judgments from more speakers and from more languages to determine the extent to which different languages exhibit the same patterns.
why you no do nothing

‘Why aren’t you doing anything?’

(45) perché non fai niente tu?

why no do nothing you

‘Why aren’t you doing anything?’

The same pattern that exists with why-suggestions and why-questions and the types of polarity items that they pair with in English and Italian exists in many other languages as well, such as Spanish, French, Russian, and Dutch (based on native speaker judgments that I have collected so far). When taken with the observation that in many of these languages, why behaves differently from other wh-words, the contrast in meaning between polarity items in why-questions and why-suggestions can lead us to draw two conclusions: first, the syntax of why-questions is different from other wh-questions, and second, the syntax of why-suggestions must differ from why-questions.

Prior analyses of why (Rizzi 2001, Shlonsky and Soare 2011, and Thornton 2008) all argue that rather than undergoing standard wh-movement from within the TP, why externally merges into a specifier in the left periphery. Before presenting data to illustrate this argument, I will briefly turn to Thornton’s discussion of children’s acquisition of why. According to Thornton, many children who are otherwise adult-like in their production of questions will not perform T-to-C movement when asking information-seeking why-questions, even 7-8 months after they consistently do so with other wh-questions. However, they can reliably produce why-suggestions with T-to-C movement. Though Thornton does not further discuss why-suggestions, this evidence may suggest that the why in suggestions and the why in
information-seeking questions merge differently. While child evidence may not be sufficient on its own to conclusively support the proposal that why-suggestions are different from why-questions in adult speech, we have already seen evidence of the abundance of asymmetries cross-linguistically between why-questions and other wh-questions; therefore it would not be so strange if these asymmetries carried over into suggestions and information-seeking why-questions. If why-questions are syntactically different from other wh-questions, then it may be more difficult for child learners to map the correct derivation onto them. Two different derivations would explain why “why don’t you” phrases can have a contrast in meaning between NPIs and PPIs, whereas other questions don’t (presumably because other questions have the same derivations). Thus, even though the same words may occur, the syntax and intended readings vary. This is something that a child learner needs to work out.

So, the puzzle here is clear: why is there a contrast in sentences with the particular constructions of why-questions like (34) and (35), contrary to the patterns of polarity items that we see among other interrogative sentences? I argue that the unique syntax of both why and how come, as opposed to other wh-words, can account for the phenomenon.

Now that we have established that why is plausibly different from other wh-words, I will present Shlonsky and Soare’s analysis of why, which is a response to Rizzi (2001)’s proposal. Rizzi suggests that why externally merges into Spec, IntP in the left periphery. However, Shlonsky and Soare argue that it actually merges in the specifier of a projection ReasonP, and then moves up to Spec, IntP (or FocusP in other cases which are not relevant here). Crucially, why needs to merge higher up because of criterial freezing, which limits (movement) chains to be between a semantically selected position where the specifier externally merges, and a scope
position, or *criterial* position to which movement occurs. To account for the multiple positions in which *why* can occur in, e.g., Italian and Japanese, or for differing interpretations of *why*-questions depending on the polarity items which they surface with, there must be the option for *why* to move further up from its merge position. This is the analysis that I will assume is correct for *why*-questions. Figure 1 illustrates Shlonsky and Soare’s proposal, where *why* first merges above the TP and then moves up even further. Note that both Rizzi and Shlonsky and Soare assume that wh-movement from within the TP holds for other wh-words, and that these words move up to FocusP in the left periphery.
Figure 1 Tree of a why-question, following Shlonsky and Soare (2011)
2.3 Why and How Come

*How come* questions provide another contrast to *why*-questions, as the following data will demonstrate. While *why* can take on the form of question or suggestion, *how come* is only available as a question as in (46a), not a suggestion as in (46b). Note that additionally, *how come* questions do not exhibit T-to-C movement like other English questions do.

(46) How come we don’t go to the store?

(a) ‘Why aren’t we going to the store?’

(b) #‘Let’s go to the store.’

Shlonsky and Soare present an analysis for *how come* which differs from that of *why*. They suggest that *how come* merges into Spec, IntP rather than moving from Spec, ReasonP, because of diagnostics in Collins (1991) which distinguish the two phrases from each other syntactically, although they seem to have the same semantic meaning in that they both ask for explanations. Collins provided some diagnostics, which Shlonsky and Soare ultimately used to refine their analysis of *how come*.

The first diagnostic involves short and long distance construals. (47a) and (47b) should both be viable interpretations of (47), but only (48a) should be a reasonable interpretation of (48). The ‘distance’ refers to the distance of the question word to the verb it is asking about. In (47) that verb is either *say* or *left*. *Say* is closer to *why*, hence ‘short distance,’ and *left* is further, hence ‘long distance.’ So, according to Collins (1991), *how come* cannot have a long distance construal, and it is hypothesized to be precisely because *how come* merges higher up in the derivation, never interacting with the lower proposition.
Why did John say Mary left?  
(Collins 1991)

(a) ‘What is the reason for John’s saying?’

(b) ‘What is the reason for Mary’s leaving?’

How come John said Mary left?  
(Collins 1991)

(a) ‘What is the reason for John’s saying?’

(b) # ‘What is the reason for Mary’s leaving?’

The second diagnostic involves quantifier interpretation: (49) should be able to get a collective reading or a pair-list reading (49a) and (49b), respectively, whereas (50) should only be able to get a collective reading, as in (50a) and not (50b). This is argued to be evidence that how come is higher in the derivation, and thus cannot access a lower interpretation of the quantifier.

Why did everybody hate John?  
(Collins 1991)

(a) ‘What is everyone’s shared reason for hating John?’

(b) ‘What are Mary’s, Sally’s, Sue’s… individual reasons for hating John?’

How come everybody hates John?  
(Collins 1991)

(a) What is everyone’s shared reason for hating John?’

(b) # ‘What are Mary’s, Sally’s, Sue’s… individual reasons for hating John?’

To capture these contrasts, Collins posits that why is in Spec, CP and that how come is the head of CP (1991: 30-32). Shlonsky and Soare adapt this proposal into the cartographic framework, as we can see in the tree in Figure 2, where how come merges as Spec, IntP. Note that this projection is higher than ReasonP (included as an empty projection), where why is argued to merge.
I will now propose a syntax of why-suggestions, based primarily on Shlonsky and Soare’s analysis. First, I will assume that interrogative why merges in Spec, ReasonP, following Shlonsky and Soare. My proposal is derived from Collins’s analysis and is supported by further
discussion that he includes in his paper. I propose that *why*-suggestions have the same analysis as *how come*, and merge in Spec, IntP. The following examples illustrate my reasons for arguing that *why*-suggestions syntactically behave more like *how come* questions than like *why*-questions. (51) is ambiguous. When it has the short construal in (51a), it can be a suggestion. When it has the long construal in (51b) it can only be a question. Therefore, the suggestion behaves like *how come*.

(51)  Why don’t you say Lucy is sad?
     (a) What is your reason for not saying it? / You should say it.
     (b) What is the reason for Lucy’s sadness?

Additionally, it is worth noting that Collins states that *how come* is likely its own lexical item or an idiom (Collins 1991: 45-46), rather than compositional *how* + *come*. I think this would also be a reasonable interpretation of *why*-suggestions (e.g. *why* + negation, without an NPI), given that they are not interrogative in nature, and given that the negation does not seem to be performing the same function that it does in other questions with negation and PPIs, for example. The potential idiomatic or lexical quality of these suggestions would be a further similarity to *how come* constructions. The tree in Figure 3 is my proposed analysis for *why*-suggestions. Negation still merges above vP in the tree because I would still like to remain agnostic about whether negation is a component of this derivation and where it originates. However, the crucial part of the proposal is that suggestion *why* merges into IntP, like *how come*, and remains in-situ rather than moving higher up, like *why*-questions are argued to do in order to yield ambiguous meanings at different criterial (scope) positions.
It is important to recognize that this analysis is imperfect in a way that is significant; *how come* is a question, and *why*-suggestions are not. Therefore, it might be the case that *why*-suggestions merge in a similar relative position to *how come*, but in a different specifier;
maybe the specifier of a projection such as SuggestP, a novel proposal based on the evidence which I have presented here.

More cross-linguistic evidence and more English data will be necessary in order to further assess and amend this proposal, but it is clear from the data presented that the asymmetry in meaning between why-questions with different polarity items is likely because they are two distinct derivations. In contrast, other interrogative phrases that do not exhibit the same asymmetry likely have the same derivations, regardless of the polarity items that may occur with them.

3 Adult Pilot

3.1 Background

Collins (1991)’s proposal for why and how come syntactic differences was based on the grammaticality judgments of other linguists. Therefore, in order to take those claims at face-value it is necessary to have a larger dataset, and ideally to subject the claims to experimentation in the wider English-speaking population. Thus, the following pilot was conducted to address whether why-questions can have ambiguous interpretations when they occur with quantifiers (52) or when they include embedded clauses (53).

(52) Why does everybody hate John?
(53) Why did John say Mary left?

Recall that Collins argues that in examples with quantifiers, like (52), why could be asking for a collective reason or individual reasons (pair-list) for hating John. An answer to the
former could be something like, “Everybody hates John because he is disruptive in class.” An answer to the latter could be “Mary hates John because he copied her test, Sally hates John because he insulted her…”

In examples with embedded clauses, like (53), why could be asking about the higher ‘saying’ event, or the lower ‘leaving’ event. The former could be answered with, for example, “John said Mary left because Sally asked him where Mary was,” and the latter with “John said Mary left because she was running late to work.”

Collins argues that how come questions don’t share the same ambiguity as why-questions, and are only subject to one reading. Thus, the questions for this pilot were twofold:

1. Are adults sensitive to ambiguities in why or how come questions?
2. Do adults’ interpretations of ambiguous why vs how come questions differ from each other?

The hypotheses were as follows:

\( H_{10} = \) Adults do not perceive ambiguities in either why or how come questions.

\( H_{1A} = \) Adults perceive ambiguities in why or how come questions.

\( H_{20} = \) Adults’ interpretations of ambiguous why and how come questions do not differ from each other.

\( H_{2A} = \) Adults’ interpretations of ambiguous why and how come questions differ from each other.

3.2 Methods

In order to answer these questions, a multiple choice survey in two parts was designed in Qualtrics. Participants were recruited via Amazon Mechanical Turk. Part 1 was a comprehension
task, and Part 2 was a grammaticality-judgment task. The design was within-subjects and blocked, so all participants first did Part 1 and then did Part 2. Part 2 was always presented second, because it was important to not have the explicit nature of grammaticality judgments reveal the focus of the comprehension task and cause participants to make metalinguistic judgments during Part 1.

3.2.1 Comprehension Task

The comprehension task consisted entirely of multiple choice items based on stories: 4 training items, and 12 why-questions, 12 how come questions, 12 other wh-questions, and one attention check. Multiple choice items consisted of 5 answers total: either one or two correct answers, one or two items which were mentioned in the story but were incorrect answers to the question, one completely unrelated item to be used as exclusion criteria, and the answer “not enough information given.” Multiple choice answers for each item were randomized for order in Qualtrics. The training items were presented first, and the remaining 37 questions were randomized for order automatically in Qualtrics.

The why- and how come questions (i.e. the test items) followed two kinds of short paragraph-style stories: either pair-list/collective stories, or stories with embedded-question-probing scenarios. There were three conditions for each kind of story (pair-list/collective and embedded question-probing) which were determined based on how much information was given in the story: enough information for a SHORT DISTANCE reading, enough for a LONG DISTANCE reading, and enough for an AMBIGUOUS reading. There were 8 items per condition: 2 probing why-questions in pair-list/collective, 2 probing how come in
pair-list/collective, 2 probing why in embedded context, and 2 probing how come in embedded context. In the pair-list/collective stories, the SHORT DISTANCE condition forced the collective reading, the LONG DISTANCE condition forced the pair-list reading, and the AMBIGUOUS condition presented enough information that both readings were available. Similarly, in the embedding-probing stories, the SHORT DISTANCE condition forced the short distance reading, the LONG DISTANCE condition forced the long distance reading, and the AMBIGUOUS condition allowed for both readings to be available. Two versions of the survey allowed the why and how come questions to be counterbalanced for all stories. Table 1 provides example stimuli for all conditions. A full list of stories and multiple choice items can be found in Appendix A.

<table>
<thead>
<tr>
<th>Supported Readings</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SHORT DISTANCE</strong></td>
</tr>
<tr>
<td><strong>PAIR-LIST/ COLLECTIVE</strong></td>
</tr>
<tr>
<td>Why (6 total)</td>
</tr>
<tr>
<td>(6 total)</td>
</tr>
<tr>
<td>How come</td>
</tr>
<tr>
<td><strong>EMBEDDED</strong></td>
</tr>
<tr>
<td>Why (6 total)</td>
</tr>
<tr>
<td>(12 total)</td>
</tr>
<tr>
<td>How come</td>
</tr>
<tr>
<td>(n=2)</td>
</tr>
</tbody>
</table>

*Table 1* Experimental conditions for Part 1 of Study 1
The rationale of forcing readings in the experimental items is to see whether participants choose answers which would be inconsistent with Collins’ interpretation of *how come*, e.g., allowing *how come* to target multiple readings and thus being ambiguous, or underpredicted according to Collins’ interpretation of *why*, indicating that *why* has more restricted contexts than claimed by Collins. By presenting each possible condition with both *why*- and *how come* questions, it is possible to see whether answers for each type of question differ in a systematic way.

### 3.2.2 Grammaticality Judgments

The grammaticality judgment part of the survey consisted of 8 questions. 4 items consisted of ambiguous pair-list/collective items with ambiguous responses, which required the presentation of a story first. Then, the participants were presented with a *why*- or *how come* question, whereby participants were given 3 possible interpretations (worded to probe short distance, long distance, and ambiguous readings) and asked to indicate which one they thought the question was asking. The remaining 4 questions consisted of the same task for embedded questions, which did not require a story first because “the animals” does not need to be specified in order to set up the context for ambiguity. Rather, one character is sufficient to probe ambiguity in embedded questions. Table 2 contains sample stimuli for each condition, as well as sample multiple-choice answers.
**Table 2** Experimental conditions for Part 2 of Study 1

<table>
<thead>
<tr>
<th>Supported Reading Stimuli</th>
<th>Multiple Choice Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMBIGUOUS</td>
<td>Short Distance</td>
</tr>
</tbody>
</table>

**PAIR-LIST/ COLLECTIVE**

<table>
<thead>
<tr>
<th>Why</th>
<th>Why do the animals want to take a break from the puzzle?</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n=2)</td>
<td>Crocodile and Bear are putting a puzzle together. Crocodile wants to take a break because the puzzle is difficult. Bear wants to take a break to have a snack. Crocodile also wants to stop to take a bathroom break, and Bear does too.</td>
</tr>
</tbody>
</table>

**How come**

| Cat and Mouse are on a sailboat. Both Cat and Mouse came because they wanted to enjoy the sunny day. Cat also came because he wanted to learn how to sail, and Mouse also came because she wanted to see the dolphins in the water. |
| (n=2) | How come the animals came on the sailboat? |

**EMBEDDED**

<table>
<thead>
<tr>
<th>Why</th>
<th>Why does Giraffe say sunflowers are her favorite?</th>
</tr>
</thead>
<tbody>
<tr>
<td>(n=2)</td>
<td>Crocodile says pizza is yummy?</td>
</tr>
</tbody>
</table>

**How come**

| The question is asking about the reason that both animals share for taking a break ONLY. | The question is asking about individual animals' distinct reasons for taking a break ONLY. | The question could be asking for individuals' reasons, or for the shared reason for taking a break |
| The question could be asking for Giraffe's reason for saying, or for the reason that sunflowers are Giraffe's favorite. | The question could be asking for Crocodile's reason for saying, or for the reason that pizza is yummy. | The question could be asking for Crocodile's reason for saying, or for the reason that pizza is yummy. |
The goal of including this part of the experiment was to compensate for any problems with salience that might occur in Part 1, particularly for the LONG DISTANCE condition with *how come* questions, where the target answer of ‘not enough information given’ might not be chosen, even if it were a more accurate interpretation than the more salient items in the multiple choice.

3.3 Results

13 participants aged 18-60 responded to the survey. None were excluded. All participants were native English speakers living in the mainland United States. Each participant was paid $7.50 for their participation, which was estimated to take 20 minutes.

3.3.1 Part 1

Figures 4 and 5 show the results of pair-list/collective stimuli, for *why*- and *how come* questions, respectively. Note that the target answers are different in the ambiguous context and the long distance context, since *how come* should only yield short distance (collective) answers. The *why*-questions yielded target responses. However, the data makes apparent that *how come* questions could yield long distance interpretations as well. Note that when an ambiguous context was presented, participants chose the ambiguous answer, or a long distance answer when asked a *how come* question; when a long distance context was presented, participants chose the long distance answer when asked a *how come* question, instead of the expected “not enough information.”
Figures 6 and 7 show the results of embedded-question stimuli, for why- and how come questions, respectively. Again, the target answers are different in the ambiguous context and the long distance context, since how come is predicted to only yield short distance answers. The why-questions yielded target responses in the SHORT DISTANCE and LONG DISTANCE conditions. In the AMBIGUOUS condition, 22 responses exhibited a preference for one interpretation; 14 chose the long distance interpretation, and 8 chose the short distance interpretation, while 4 allowed for ambiguity. These are strictly not incorrect responses, but they differ from the target. However 6/13 participants were not consistent with their two responses, indicating that despite commitment to one reading, both types of readings were indeed possible within the condition.

Turning to the how come questions, the data again makes apparent that how come questions could yield both short distance and long distance interpretations as with the pair-list/collective questions. In the ambiguous context, where the expected answer is short distance, we see a similar pattern to the why-questions. Participants did choose short distance interpretations a plurality of times, but they also chose ambiguous or long distance interpretations. Again, when a
long distance context was presented, participants chose the long distance answer when asked a
*how come* question, instead of the expected “not enough information.”

Figure 6 Embedded *why*-questions

![Graph showing why-question responses](image)

In *how come* questions which provided ambiguous contexts, adults over-accepted according to Collins’ expectations every single time (52/52). In both non-ambiguous embedded-question conditions, 5/26 answers were ‘not enough information’. Among the non-ambiguous pair-list/collective scenarios, adults answered as expected for 25/26 answers.

3.3.2 Part 2

The grammaticality judgments did not yield any data which was inconsistent with the experiment. While some participants indicated preferences for one interpretation over others, the answers were neither internally consistent nor consistent between subjects for any given item.

Figures 8 and 9 show the grammaticality judgments for pair-list/collective stimuli. We see that most participants accepted ambiguity for both *why*- and *how come* questions, and they did so at about the same rate. Next, for both questions, was long distance (which should be

31
impossible for *how come*, according to the theoretical assumptions), and least common was short distance. Crucially, the proportions of answers were nearly identical for *why* and *how come* questions.

![Figure 8](image)

**Figure 8** Pair-list/collective *why* judgments (Target: Ambiguous)

![Figure 9](image)

**Figure 9** Pair-list/collective *how come* judgments (Target: Short distance)

Figures 10 and 11 show grammaticality judgments for embedded-question stimuli. Here, most participants preferred the short distance interpretation for both *why*- and *how come* questions, although again, all interpretations were possible. Again, the proportions of answers were nearly identical for *why* and *how come* questions.
3.4 Discussion

Based on the comprehension task and on grammaticality judgments, adults appear to treat *why*- and *how come* questions as equally ambiguous in both pair-list/collective and embedded contexts. No systematic differences were found between *why*- and *how come* questions, and we now know that adults perceive ambiguities for both kinds of question. Adults preferred to choose a non-ambiguous reading for embedded sentences but were not internally consistent. Therefore, it is apparent that even when there was a preference, both readings were available. There were 5 instances of participants choosing ‘not enough information given’ in the comprehension task for
embedded *why*-questions, but this likely has to do with the lack of salience of ‘saying.’ Recall that in the comprehension task, adults always chose the long-distance reading of *how come* instead of “not enough information.” It is possible that they were accommodating for the task when in reality this distinction exists. However, the fact that they chose long-distance readings 100% of the time in the comprehension task, and that their explicit grammaticality judgments seemed to corroborate this result casts some doubt that the results demonstrated mere accommodation.

These results raise the question of why Collins and his colleagues got the judgments that were published into the 1991 paper. It is possible that they belong to a more conservative dialect, although none of the participants exhibited the same pattern despite the diversity in age and location. It is also possible that this is a distinction which does indeed exist, but perhaps is only perceivable in natural speech, rather than in writing. This would explain why it may only be salient to those with a high level of metalinguistic awareness, but this possibility would need to be investigated further before arriving at any conclusions. Another possibility is that these results demonstrate language change in-progress. Participants from a wide age-range were recruited to determine whether behavior varied by age, with the assumption that if language change had occurred, older participants might pattern differently than younger participants. Though behavior did not pattern by age, it is still possible that the interpretation of *how come* questions is undergoing change for all speakers, regardless of age.

---

5 It is worth noting here that I also initially agreed with the judgments in the Collins paper, as do some other linguists who I consulted. I first hypothesized that perhaps my own judgments were due to confirmation bias, but Lucas Champollion pointed out that Sprouse, Schütze, & Almeida (2013) experimentally validated 95% (±5%) of grammaticality judgments in a syntax journal, with 900 participants. Therefore, it would be prudent to do a follow-up study explicitly testing for confirmation bias by telling some participants to expect a contrast between *why* and *how come.*
The crucial takeaway here is that since experimentally, adults appear to treat why- and how come questions the same, further investigations should not assume that there are salient syntactic differences between the two questions for adults.

4 Why-questions and Children: Developmental Bases

4.1 Background

According to Thornton (2016), children are able to produce questions (including those with relative clauses) in an adult-like way by age 6. According to Wode (1971), what and where are the first questions for which children develop adult-like productions, and when and why are the final questions for which children develop adult-like productions. Thornton (2008) did a longitudinal study on the acquisition of why-questions, and found that some children (but not all) persist in not being adult-like with their productions of why-questions long after they are adult-like with their other wh-questions. The errors of production in why-question consist of children not producing inversion that occurs with T-to-C movement in English questions. (54) is an example of one of these productions, at age 3;2.

(54) Why you can only do it (and I can’t)? [make coffee] (Thornton 2008)

Given that the acquisition of why-questions has not been studied experimentally, especially considering this unusual pattern, more investigation is warranted. Additionally, although children’s productions of why have been studied, their comprehension of the questions, and their answering behavior have never (to my knowledge) been studied. Both of these are of interest, although ultimately the latter will only be investigated indirectly in the experiment which will be discussed shortly, in 5.2.
4.2 Corpus work

To get a sense of how frequent why-questions are relative to other questions in children’s input and productions, it was necessary to investigate Thornton’s claims about children’s productions of why using corpus data, and additionally the adult input. Therefore, analyses were performed on the Adam corpus to get quantitative data on both Adam and adults’ productions of why- and other wh-questions.

Table 3 shows the total number of wh-words, total tokens, and percentage of wh-words by total tokens produced by adults and Adam in the corpus. We can see that the child produced slightly more questions than the adults, but the child had a greater number of tokens. Therefore, the percentage of wh-words by tokens is greater for adults. Interestingly, the total number of wh-words produced by both adults and Adam are nearly identical.

<table>
<thead>
<tr>
<th></th>
<th>Total Wh-Words</th>
<th>Total Tokens</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adults</td>
<td>6034</td>
<td>117882</td>
<td>5.12%</td>
</tr>
<tr>
<td>Adam</td>
<td>6196</td>
<td>163137</td>
<td>3.79%</td>
</tr>
</tbody>
</table>

Table 3 Proportion of question words to total number of words

Figures 11 and 12 show the average monthly frequency of wh-words produced by adults and Adam, respectively. The rationale of examining monthly averages is to capture change in frequency over time within the limitations of the corpus (which has at least one session for almost every month from 2;03-5;02). Both adults and Adam produce what far more than other wh-questions each month. Adam asks more where and who questions than adults. Adults ask
why-questions consistently over time, as does Adam, but Adam asked far more why-questions between 3;0 and 4;0. How come questions occurred extremely infrequently, so they were excluded from the data visualization, but adults asked them 3 times in total. Adam asked them more, a total of 20 times, but 12 of those were in one session at age 4;10.

**Figure 11** Adult average monthly frequency of wh-words at 6-month intervals

**Figure 12** Adam average monthly frequency of wh-words at 6-month intervals
Figures 13 and 14 give a proportional breakdown of adults’ and Adam’s wh-questions in aggregate rather than over time. Note that although what questions make up more of adults’ total questions, both adults and Adam ask what most, followed by where, followed by why, then how⁶ and who, and finally which and when.

**Figure 13** Proportional breakdown of total adult questions

**Figure 14** Proportional breakdown of Adam’s total questions

### 4.3 Discussion

The results of these corpus examinations show that Adam heard and produced why-questions throughout his development. Attention was not paid to the adult-like-ness of his

---

⁶ Counted separately from how come, which is not visible in Figures 13 or 14 because of extremely infrequent use (3 total uses by adults; and 20 total uses by Adam, with 12 in the same session).
questions over time, however this data verifies that it would be possible to test the acquisition of why-questions experimentally from very young ages, since children are hearing these questions, and since the proportion of their usage of these questions is very similar to adults’ usage of them.7

5 Child Experimental Study

5.1 Background

This experimental study investigates whether children are sensitive to the same ambiguities as adults in their comprehension and production of why-questions, and whether grammatical competence with these types of questions shows any asymmetries between comprehension and production.

It is of interest to test both comprehension and production not because of the specific domain of study, but because of previous conventions in the field which lead to unnecessary vagueness. As in most domains of first language acquisition, comprehension of syntax precedes adult-like production of the same type of construction. So when non-adult-like productions of a grammatical construction occur, is it also the case that comprehension of the same construction is not adult-like? Additionally, for why-questions in particular, prior studies (e.g. Thornton 2008) focused only on production, which may not provide a full picture of competence. The hypothesis for this experiment is that children are adult-like in their comprehension of ambiguous why-questions before their production of ambiguous why-questions.

7 Long distance questions are relevant for the purposes of this thesis. The total number of long distance questions in the corpus was not counted. However, we know Adam was exposed to these questions from as early as 2:09.04, e.g. MOT: “How do you know who did it if you don’t ask?” (Brown 1973).
5.2 Methods

The child experiment consists of two components: a comprehension style task, which is similar to the adult study described above, but conducted in person and presented with images and audio stimuli; and a production task which consists of sentence repetition.

5.2.1 Comprehension

Figure 15 is an example of a stimulus image for the comprehension task, and the text in (56) is an example of a pair-list/collective item. There will also be embedded-question items.

![Figure 15 Stimulus image for Part 1 (TVJT) of Study 2](image)

(56) Narrator: “It is horse’s birthday and he is having a party at the park. Mouse went because horse is her best friend, bat went to play with the balloons, and both the animals went because they wanted to eat birthday cake.

Frog: “Why are Bat and Mouse at the party?”

Duck: “Mouse is at the party because Horse is her best friend, and Bat is at the
party to play with the balloons.”

OR

“Mouse and Bat are at the party because they want to eat birthday cake.”

Experimenter: “Is Duck right?”

If yes: “How do you know?”
If no: “What really happened?”

To test comprehension I use a truth-value judgment task consisting of two blocks: a training section with 4 stimuli, and a testing section with 24 stimuli. There are 4 kinds of stimuli in the testing section: 6 stories with pair-list/collective why-questions, with 3 items targeting the pair-list reading and 3 items targeting the collective reading; 6 embedded question items, with 3 items each targeting the long distance reading and 3 items targeting the short distance reading; 4 stories preceding non-why long distance wh-questions to see whether children can handle embedded questions, and 8 stories with short distance why-questions to make sure children are competent at answering simpler why-questions. Table 4 includes examples of the experimental conditions, and a full list of stimuli is included in Appendix B.

<table>
<thead>
<tr>
<th>Interpretation in answer</th>
<th>SHORT DISTANCE</th>
<th>LONG DISTANCE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PAIR-LIST/COLLECTIVE</strong></td>
<td>Mouse and Bat are at the party because they want to eat birthday cake.</td>
<td>Mouse is at the party because Horse is her best friend, and Bat is at the party to play with the balloons.</td>
</tr>
<tr>
<td>(n=3 per condition)</td>
<td>Mouse says that Bat is sad because Cat asked her.</td>
<td>Mouse says that Bat is sad because his bicycle broke.</td>
</tr>
<tr>
<td><strong>EMBEDDED</strong></td>
<td>Mouse says that Bat is sad because Cat asked her.</td>
<td>Mouse says that Bat is sad because his bicycle broke.</td>
</tr>
<tr>
<td>(n=3 per condition)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Table 4* Experimental conditions for Part 1 (TVJT) of Study 2
In the comprehension task, a narrator will introduce two characters: Frog and Duck. Frog likes listening to stories, but sometimes doesn’t hear because she gets distracted chasing yummy flies. Frog’s friend Duck wants to help by answering Frog’s questions, but sometimes gets it wrong. Therefore, the child is asked to answer whether Duck is right. Children will be shown a picture of a scenario, and will listen to the narrator describing the context of the image. Frog will then ask an ambiguous why-question, and will provide a possible answer which offers only one interpretation. When the question has a quantifier, the possible answer will target either a pair-list reading or a collective reading, and when the question has an embedded clause, the answer will target either the higher or lower event. The children will then be asked whether the answer is correct based on the scenario described in the picture. Crucially, all of the answers of the long distance why-questions should be true for adults; they offer a singular, limited interpretation, but one that should be available regardless. We are looking for a systematic pattern of rejections by children to determine whether they differ from adults by having more limited interpretations. If children are comprehending in a completely adult-like way, we may test younger children.

 Initially, children from ages 5-7 will be tested. By this age, they should be able to produce why-questions in an adult-like way (according to Thornton 2008). They should also have competence with embedding (Kidd and Bavin 2002 claim that children can comprehend embedded questions by age 5, and can produce them as early as 3;3).

### 5.2.2 Production

The production task uses a simple repetition task. Klem et al. (2015) demonstrated that
when children are asked to repeat sentences, they will only repeat them accurately if they have grammatical competence with the construction being tested. Thus, repetition is a measure of linguistic competence and not merely working memory. So in the production task, children will be asked to repeat simple why-questions, why-questions with quantifiers, and why-questions with embedded clauses. We are looking for systematic patterns of degradation in ability to accurately repeat sentences to ascertain the extent to which children display grammatical competence in their production.

6 Conclusion

In this thesis, we have explored multiple aspects of why-questions in adults and children. I have presented data which poses a problem for the traditional analysis of why-questions as being syntactically identical to other wh-questions. The novel data included about why-suggestions was incorporated into an analytical framework provided by Shlonsky and Soare (2011) and Collins (1991).

An adult pilot was conducted to test diagnostics in Collins (1991) which were used to argue that why and how come questions are syntactically distinct from each other. Those results appear to contradict Collins’s claims entirely. This is an interesting case study of a theoretical argument which seems to hold up based on individual grammaticality judgments, but which is disproven experimentally. However, it is possible that these results are due to language change in-progress rather than inaccurate judgments. As Sprouse, Schütze, and Almeida (2013) demonstrated, grammaticality judgments tend to be accurate most of the time (95% ± 5), so the

---

8 Repetition tasks have frequently been used to assess SLI, especially in bilinguals, in comparison to typically-developing children (e.g. Clay 1971, Marinis & Armon-Lotem 2015, Riches 2012).
judgments in Collins (1991) should not be quickly dismissed as unrepresentative of *how come* questions in 1991. These results can perhaps serve as a motivation to empirically test some high-stakes theoretical claims based on the grammaticality judgments of few individuals (particularly as time from publication elapses, to account for possible language change), which is not currently standard for the field.

An examination of the Adam corpus looked at the frequency of wh-questions over time in the input that a child receives, to get an idea of how questions are proportionally distributed. The goal was to determine whether children may acquire *why*-questions later because they are skewed to a more frequent distribution later in the input, but this was not the case. The corpus results verified that all questions are consistently available in both the adult input and in child productions at all time intervals.

Finally a child experiment was proposed which would do something new: test children’s comprehension of *why*-questions, by asking them to perform a Truth Value Judgment Task to determine which answers were acceptable and unacceptable answers to *why*-questions. Children’s ability to produce these questions would also be concurrently tested, to get an idea of how certain un-adult-like comprehension might pattern with un-adult-like production.
Acknowledgments

There are so many people without whom this thesis would not have been possible. Firstly, I would like to thank my advisor, Dr. Ailís Cournane. We arrived at NYU at the same time, and she taught my first linguistics course in the fall of my freshman year. When I approached her as a sophomore with an interest in doing acquisition research, she took a chance on me, and the rest, as they say, is history. Ailís has always challenged me to do more than I thought I could, and she has made herself available to me for countless hours over the past three years. I am deeply indebted to her for her guidance and support.

Thanks to the faculty and graduate students who provided extensive feedback as I developed my ideas, especially Dr. Stephanie Harves, Dr. Chris Collins, Dr. Lucas Champollion, Dr. Renée Blake, Dr. Zachary Jaggers, Alicia Parrish, Maxime Tulling, and Alicia Chatten. Thanks to all the other faculty and students in the department who have been so kind, generous, and receptive. The support I have received in this department has gone far beyond academics and for that I will always be grateful.

Thanks to the Child Language Lab, Dr. Sudha Arunachalam, and the LEARN Lab for providing invaluable support, and to Michael Marinaccio, Chiara Repetti-Ludlow, and Melissa Rojas for lending me their voices for stimuli.

Mom, Dad, Izzy, and Abuela: you are the best family I could have ever asked for. Thank you for your constant encouragement. CJ and Alex, I feel so lucky to have had you both in my life over the past 4 years. Thanks for keeping me sane through all this! Finally, biggest thanks to B and C, the kids who inspired all of this!
References


https://doi.org/10.1111/desc.12202


Appendix

A: Experiment 1 Items

<table>
<thead>
<tr>
<th>Story</th>
<th>Question</th>
<th>Multiple Choice Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Hippo and Crocodile are best friends. They are spending the whole day together. In the morning, they will go to school together. They are going to study at the library in the afternoon. Then, they are going to go to the store tonight to get food to cook dinner.</td>
<td>When are Hippo and Crocodile going to the store? You should choose the option that best answers the question. Please choose &quot;Tonight&quot;.</td>
<td>Tonight, Yesterday, In the morning, In the afternoon, Not enough information given</td>
</tr>
<tr>
<td>2. Bear likes to go to the park and play. On his way to the park, Bear meets his friend Crocodile. Crocodile likes to jump rope. Crocodile and Bear are playing at the park together now. Bear is swinging on the monkey bars.</td>
<td>What is Bear doing? Sometimes more than one answer will be correct. Please choose &quot;Swinging on the monkey bars&quot; AND &quot;Playing at the park with Crocodile&quot;.</td>
<td>Swinging on the monkey bars, Jumping rope, Eating a sandwich, Playing at the park with Crocodile, Not enough information given</td>
</tr>
<tr>
<td>3. Mouse and Giraffe are in a band together. They rehearse at Giraffe's house sometimes and at Mouse's house other times. They like to perform for their friends often. They are going to sing a song on stage today.</td>
<td>Where are Mouse and Giraffe going to sing a song?</td>
<td>On stage, At Mouse’s house, At Giraffe’s house, At school, Not enough information given</td>
</tr>
</tbody>
</table>

Table 5 Experiment 1 training stimuli

9 For all multiple choice items in Experiment 1, the expected answer is shaded. The orders of all multiple choice items except “not enough information given”, and of all items in the comprehension task after training, are randomized in Qualtrics. “Not enough information given” is always the last multiple-choice item presented.
<table>
<thead>
<tr>
<th>Story</th>
<th>Question</th>
<th>Multiple Choice Answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Today is Bear's birthday, and he is having a party at the park. There will be a big piñata with lots of candy inside. Both Mouse and Crocodile are going to the party because they want to play with the piñata. Mouse is especially excited to go to the party because Bear is her best friend, and Crocodile is excited to go to the party because he wants to play outside.</td>
<td>Why are the animals going to the party?</td>
<td>Best friends with Bear</td>
</tr>
<tr>
<td></td>
<td>How come the animals are going to the party?</td>
<td>Best friends with Bear</td>
</tr>
<tr>
<td>2. Crocodile, Cat and Giraffe are swimming in the ocean. Crocodile asks his friends if they want to sit on the beach instead, and they both say yes. Giraffe wants to sit on the beach because the water is too cold. Cat wants to sit on the beach because he is tired of swimming. Cat also wants to sit on the beach so he can keep building their big sand castle, and Giraffe does too.</td>
<td>Why do the animals want to sit on the beach?</td>
<td>To keep building the sand castle</td>
</tr>
<tr>
<td></td>
<td>How come the animals want to sit on the beach?</td>
<td>To keep building the sand castle</td>
</tr>
<tr>
<td>3. Cat and Mouse are on a sailboat. Both Cat and Mouse came because they wanted to enjoy the sunny day. Cat also came because he wanted to learn how to sail, and Mouse also came because she wanted to see the dolphins in the water.</td>
<td>How come the animals came on the sailboat?</td>
<td>To learn how to sail</td>
</tr>
<tr>
<td></td>
<td>Why did the animals come on the sailboat?</td>
<td>To learn how to sail</td>
</tr>
</tbody>
</table>

10 Pair-list/collective stories are counterbalanced for order of information presentation. The counterbalanced alternative for this story is “Today is Bear’s birthday, and he is having a party at the park. Mouse is especially excited to go to the party because Bear is her best friend, and Crocodile is excited to go to the party because he wants to play outside. Both Mouse and Crocodile are also going to the party because they want to play with the piñata.” All other alternatives follow the same order.

11 For each story, participants are either presented with a why or how come question. Participants are presented with an equal number of why and how come questions.
<table>
<thead>
<tr>
<th></th>
<th>How come the animals want to take a break from the puzzle?</th>
<th>The puzzle is difficult</th>
<th>To have a snack</th>
<th>To take a bathroom break</th>
<th>To go home</th>
<th>Not enough information given</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.</td>
<td>Crocodile and Bear are putting a puzzle together. Crocodile wants to take a break because the puzzle is difficult. Bear wants to take a break to have a snack. Crocodile also wants to stop to take a bathroom break, and Bear does too.</td>
<td>Why do the animals want to take a break from the puzzle?</td>
<td>The puzzle is difficult</td>
<td>To have a snack</td>
<td>To take a bathroom break</td>
<td>To go home</td>
</tr>
<tr>
<td>5.</td>
<td>Bear and Mouse are eating some yummy pizza. Bear likes pizza with veggies more than pepperoni pizza, and so does Mouse, because pepperoni is too spicy.</td>
<td>Why do the animals like pizza with veggies more than pepperoni pizza?</td>
<td>Pepperoni is too spicy</td>
<td>The pizza with veggies is cold</td>
<td>They are eating pizza</td>
<td>The pizza is yummy</td>
</tr>
<tr>
<td>6.</td>
<td>Cat and Crocodile want to surprise Hippo with a gift, since Hippo is such a good friend to them! Both Cat and Crocodile want to give her a hat instead of a cape, because Hippo’s head is cold.</td>
<td>Why do the animals want to give Hippo a hat instead of a cape?</td>
<td>Hippo is a good friend</td>
<td>Hippo's head is cold</td>
<td>Don't like how capes look</td>
<td>Want it to be a surprise</td>
</tr>
<tr>
<td>7.</td>
<td>Hippo and Mouse are walking to school this morning instead of taking the bus. Hippo is walking because the weather is really warm, and so is Mouse.</td>
<td>How come the animals are walking to school?</td>
<td>The weather is warm</td>
<td>The bus is broken</td>
<td>They are together</td>
<td>It is morning</td>
</tr>
<tr>
<td>8.</td>
<td>Bear and Crocodile both play instruments in the orchestra. Both Bear and Crocodile play trumpet instead of violin because the trumpet is louder and they like loud sounds.</td>
<td>How do the animals play trumpet instead of violin?</td>
<td>Trumpet is louder</td>
<td>They are in the orchestra</td>
<td>They like violins more</td>
<td>They play instruments</td>
</tr>
</tbody>
</table>
9. Cat and Mouse are watching their friends play baseball at the field. Cat stands up to cheer for his friends. Mouse stands up because she can’t see when she’s sitting down.

Why do the animals stand up?
- Can’t see
- To cheer for friends
- To get a snack
- They are at the field
- Not enough information given

How come the animals stand up?
- Can’t see
- To cheer for friends
- To get a snack
- They are at the field
- Not enough information given

10. Hippo and Bear are riding the biggest roller coaster at the amusement park. Hippo screams because she is having a great time. Bear screams because he is afraid.

Why do the animals scream?
- Afraid
- Having a great time
- The rollercoaster is broken
- They are at the amusement park
- Not enough information given

How come the animals scream?
- Afraid
- Having a great time
- The rollercoaster is broken
- They are at the amusement park
- Not enough information given

11. Giraffe and Crocodile are dancing together. Giraffe is dancing because she likes the music. Crocodile is dancing because it is a good form of exercise.

Why are the animals dancing?
- Like the music
- Good exercise
- They are together
- They were bored
- Not enough information given

How come the animals are dancing?
- Like the music
- Good exercise
- They are together
- They were bored
- Not enough information given

12. Hippo and Giraffe are taking pictures of flowers in the park. Hippo is taking pictures of the flowers because the flowers are blooming. Giraffe is taking pictures of the flowers because there are so many bees on them.

Why are the animals taking pictures of the flowers?
- The flowers are blooming
- They have bees on them
- They are in the park
- They have unique colors
- Not enough information given

How come the animals are taking pictures of the flowers?
- The flowers are blooming
- They have bees on them
- They are in the park
- They have unique colors
- Not enough information given

Table 6 Experiment 1 Pair-list/collective comprehension stimuli
<table>
<thead>
<tr>
<th>Question</th>
<th>Scenario</th>
<th>Why</th>
<th>How come</th>
<th>Not enough information given</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. Crocodile and Bear are reading at the library. Bear really likes her book because it has lots of pictures, and makes a thumbs up sign to show Crocodile how much she likes it. Crocodile doesn’t see, so Bear tells Crocodile that she likes the book.</td>
<td>How come Bear says she likes the book?</td>
<td>Crocodile doesn't see her thumbs up</td>
<td>They are at the library</td>
<td>Crocodile wants to know</td>
</tr>
<tr>
<td>4. Giraffe and Mouse are painting. Mouse shows Giraffe her painting and asks if she likes it. Giraffe tells Mouse that the painting is beautiful because it has so many colors.</td>
<td>How come Giraffe says that Mouse’s painting is beautiful?</td>
<td>Mouse asked if she likes it</td>
<td>They are both painting</td>
<td>Mouse told Giraffe her painting is beautiful first</td>
</tr>
<tr>
<td>5. Mouse and Crocodile are watching a movie at Mouse’s house. Crocodile chose the movie, because Mouse chose last time. Mouse wonders if Crocodile is hungry, and asks him if he wants some popcorn.</td>
<td>Why does Mouse ask Crocodile if he wants popcorn?</td>
<td>Mouse wonders if he is hungry</td>
<td>They are at Mouse’s house</td>
<td>Crocodile only eats popcorn</td>
</tr>
<tr>
<td>6. Bear and Giraffe are studying for their math test at the library. Giraffe asks Bear to say what 1+1 is. Bear says that 1+1 is 2.</td>
<td>Why does Bear say 1+1 is 2?</td>
<td>Bear loves math</td>
<td>They are at the library</td>
<td>Bear and Giraffe are together</td>
</tr>
<tr>
<td>7. Mouse and Bear are telling scary stories at the campfire. Bear says that the witch could fly.</td>
<td>How come Bear says that the witch could fly?</td>
<td>The story is scary</td>
<td>Bear is telling a story</td>
<td>Mouse and Bear are at a campfire</td>
</tr>
<tr>
<td>8. Crocodile and Cat are rehearsing for their play which they will perform tonight at school. When it is Crocodile’s turn to say his line, he says “I am so mad at you, Cat!”</td>
<td>How come Crocodile says he is mad?</td>
<td>It is his line in the play</td>
<td>Crocodile doesn't like Cat</td>
<td>The performance is tonight</td>
</tr>
</tbody>
</table>
9. **Mouse and Crocodile are about to take a trip in an airplane. Crocodile says that he is excited to fly because it is his first time.**

**Why does Crocodile say he is excited to fly?**

- It is his first time
- Mouse asked
- They are about to be on an airplane
- They are taking a trip
- Not enough information given

**How come Crocodile says he is excited to fly?**

- It is his first time
- Mouse asked
- They are about to be on an airplane
- They are taking a trip
- Not enough information given

10. **Cat and Giraffe are planting flowers in the garden. Giraffe says that sunflowers are her favorite because they are yellow just like her!**

**Why does Giraffe say sunflowers are her favorite?**

- She wanted to say so
- They are yellow like her
- She is planting flowers with Cat
- They are in the garden
- Not enough information given

**How come Giraffe says sunflowers are her favorite?**

- She wanted to say so
- They are yellow like her
- She is planting flowers with Cat
- They are in the garden
- Not enough information given

11. **Bear and Giraffe are at an art museum with lots of sculptures and paintings. Bear says she is happy to be at the museum because there are so many interesting sculptures!**

**Why does Bear say she is happy to be at the museum?**

- There are many interesting sculptures
- Bear wanted to say so
- Giraffe is with Bear
- There are paintings at the museum
- Not enough information given

**How come Bear says she is happy to be at the museum?**

- There are many interesting sculptures
- Bear wanted to say so
- Giraffe is with Bear
- There are paintings at the museum
- Not enough information given

12. **Hippo and Cat are about to perform in a play. All of their friends and families will be there. Hippo says she is scared because it is her first time on a stage.**

**Why does Hippo say she is scared?**

- It is her first time on a stage
- She is going to perform with Cat
- Her friends will be there
- Not enough information given

**How come Hippo says she is scared?**

- It is her first time on a stage
- She is going to perform with Cat
- Her friends will be there
- Not enough information given

---

Table 7: Experiment 1 Embedded-question comprehension stimuli

<table>
<thead>
<tr>
<th>Story</th>
<th>Question</th>
<th>Multiple choice answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. <strong>Mouse is eating ice cream at the ice cream shop. Giraffe is waiting outside. Mouse is having chocolate, which is her favorite flavor.</strong></td>
<td>Who is eating ice cream?</td>
<td>Mouse</td>
</tr>
<tr>
<td>2. <strong>Cat is tying his shoes with help from Mouse and Crocodile. They are experts and are happy to help!</strong></td>
<td>How is Cat tying his shoes?</td>
<td>With help from Mouse</td>
</tr>
</tbody>
</table>
3. Giraffe and Hippo are on a camping trip. They are building a campfire outside. There is a tent nearby, and a fishing pole.

<table>
<thead>
<tr>
<th>What are Giraffe and Hippo building?</th>
<th>They are building a campfire</th>
<th>They are building a tent</th>
<th>They are building a house</th>
<th>They are building a fishing pole</th>
<th>Not enough information given</th>
</tr>
</thead>
</table>

4. Bear left home to play a game at the arcade. She is playing there now, in the blue room. If she wins, she will play again. Giraffe left home to come and play with Bear.

<table>
<thead>
<tr>
<th>Where is Bear playing a game?</th>
<th>Home</th>
<th>Arcade</th>
<th>Park</th>
<th>Giraffe’s house</th>
<th>Not enough information given</th>
</tr>
</thead>
</table>

5. Crocodile and Giraffe ride the bus in the morning. They need to get to school. In the afternoon, they walk home. In the evening, they ride the bus to work.

<table>
<thead>
<tr>
<th>When do Crocodile and Giraffe ride the bus?</th>
<th>Morning</th>
<th>Afternoon</th>
<th>Evening</th>
<th>Night</th>
<th>Not enough information given</th>
</tr>
</thead>
</table>

6. Hippo and Crocodile are jumping on the trampoline. They want to find out who can jump the highest. Cat is watching.

<table>
<thead>
<tr>
<th>Who is jumping on the trampoline?</th>
<th>Hippo is jumping</th>
<th>Crocodile is jumping</th>
<th>Cat is jumping</th>
<th>Bear is jumping</th>
<th>Not enough information given</th>
</tr>
</thead>
</table>

7. The floor is dirty! Cat cleans it by sweeping up with a broom. Giraffe cleans it with a mop. It is hard work and takes a long time.

<table>
<thead>
<tr>
<th>How does Cat clean the floor?</th>
<th>With a broom</th>
<th>With a mop</th>
<th>With a sponge</th>
<th>Quickly</th>
<th>Not enough information given</th>
</tr>
</thead>
</table>

8. Hippo is going to the doctor because her tummy and head hurt. The doctor examines Hippo, and checks her heart rate also. It will be all better soon!

<table>
<thead>
<tr>
<th>What hurts Hippo?</th>
<th>Tummy</th>
<th>Head</th>
<th>Heart</th>
<th>Ears</th>
<th>Not enough information given</th>
</tr>
</thead>
</table>

9. Crocodile is blowing bubbles by the lake. Crocodile loves to blow the biggest bubbles! Later, he will walk past the river and blow more bubbles on his way home.

<table>
<thead>
<tr>
<th>Where is Crocodile blowing bubbles?</th>
<th>Lake</th>
<th>River</th>
<th>Home</th>
<th>School</th>
<th>Not enough information given</th>
</tr>
</thead>
</table>

10. Bear waters her plants at noon. They need lots of sunlight and water. She checks on them every morning, noon, and night.

<table>
<thead>
<tr>
<th>When does Bear water her plants?</th>
<th>At noon</th>
<th>At night</th>
<th>At bed time</th>
<th>In the morning</th>
<th>Not enough information given</th>
</tr>
</thead>
</table>
Giraffe does too.

Building

Sit

Tired

Wants

Because

Giraffe

Beach

Swimming

Crocodile,

Because he wants to play outside.

Crocodile

Especially

Play

Going

Inside.

Be

Having

Today

You are helping her out!

Mouse is very happy that

Answer, and do not pick

Whether you are paying

Mouse wants to check

Story

Time last night.

Soundly

And for a long

Time last night.

Hippo wakes up bright

And early. She slept

Soundly and for a long

Time last night.

Table 8 Experiment 1 Other wh-question comprehension stimuli

<table>
<thead>
<tr>
<th>Story</th>
<th>Question</th>
<th>Multiple choice answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mouse wants to check whether you are paying attention. To assure her that you are, please choose &quot;Mouse is eating cheese.&quot; Only choose that answer, and do not pick any additional choices. Mouse is very happy that you are helping her out!</td>
<td>What is Mouse doing? Mouse is eating cheese Mouse is making a cake Mouse is walking to the river Mouse is going to sleep</td>
<td>Not enough information given</td>
</tr>
</tbody>
</table>

Table 9 Experiment 1 Attention check

<table>
<thead>
<tr>
<th>Story</th>
<th>Question</th>
<th>Multiple choice answers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Today is Bear's birthday, and he is having a party at the park. There will be a big pinata with lots of candy inside. Both Mouse and Crocodile are going to the party because they want to play with the pinata. Mouse is especially excited to go to the party because Bear is her best friend, and Crocodile is excited to go the party because he wants to play outside.</td>
<td>Why did the animals go to the party? The question is asking about individual animals' distinct reasons for going to the party ONLY.</td>
<td>The question could be asking for individual(s)' reasons, or for the shared reason.</td>
</tr>
<tr>
<td></td>
<td>How come the animals went to the party? The question is asking about individual animals' distinct reasons for going to the party ONLY.</td>
<td>The question could be asking for individual(s)' reasons, or for the shared reason.</td>
</tr>
<tr>
<td>2. Crocodile, Cat and Giraffe are swimming in the ocean. Crocodile asks his friends if they want to sit on the beach instead, and they both say yes. Giraffe wants to sit on the beach because the water is too cold. Cat wants to sit on the beach because he is tired of swimming. Cat also wants to sit on the beach so he can keep building their big sand castle, and Giraffe does too.</td>
<td>Why do the animals want to sit on the beach? The question is asking about individual animals' distinct reasons for wanting to sit on the beach ONLY.</td>
<td>The question could be asking for individual(s)' reasons, or for the shared reason.</td>
</tr>
<tr>
<td></td>
<td>How come the animals want to sit on the beach? The question is asking about individual animals' distinct reasons for wanting to sit on the beach ONLY.</td>
<td>The question could be asking for individual(s)' reasons, or for the shared reason.</td>
</tr>
<tr>
<td>Question</td>
<td>Multiple Choice Answers</td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td><strong>1. Why does Cat say that Giraffe is sad?</strong></td>
<td>The question is asking for Cat's reason for saying that Giraffe is sad ONLY.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The question is asking for Giraffe's reason for being sad ONLY.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The question could be asking for Cat's reason for saying, or for Giraffe's reason for being sad.</td>
<td></td>
</tr>
<tr>
<td>How come Cat says that Giraffe is sad?</td>
<td>The question is asking for Cat's reason for saying that Giraffe is sad ONLY.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The question is asking for Giraffe's reason for being sad ONLY.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The question could be asking for Cat's reason for saying, or for Giraffe's reason for being sad.</td>
<td></td>
</tr>
<tr>
<td><strong>2. Why does Mouse say that Bear is friendly?</strong></td>
<td>The question is asking for Mouse's reason for saying that Bear is friendly ONLY.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The question is asking for the reason that Bear is friendly ONLY.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The question could be asking for Mouse's reason for saying, or for the reason that bear is friendly.</td>
<td></td>
</tr>
<tr>
<td>How come Mouse says that Bear is friendly?</td>
<td>The question is asking for Mouse's reason for saying that Bear is friendly ONLY.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The question is asking for the reason that Bear is friendly ONLY.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The question could be asking for Mouse's reason for saying, or for the reason that bear is friendly.</td>
<td></td>
</tr>
</tbody>
</table>

Table 10 Experiment 1 Pair-list/collective grammaticality judgment stimuli
<table>
<thead>
<tr>
<th></th>
<th>Question</th>
<th>The question could be asking for the reason that sunflowers are Giraffe's favorite.</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>How come Giraffe says sunflowers are her favorite?</td>
<td>Giraffe's reason for saying that sunflowers are her favorite ONLY.</td>
</tr>
<tr>
<td></td>
<td>Why does Giraffe say sunflowers are her favorite?</td>
<td>The question could be asking for the reason that sunflowers are Giraffe's favorite.</td>
</tr>
<tr>
<td>4</td>
<td>How come Crocodile says pizza is yummy?</td>
<td>Crocodile's reason for saying that pizza is yummy ONLY.</td>
</tr>
<tr>
<td></td>
<td>Why does Crocodile say pizza is yummy?</td>
<td>The question could be asking for the reason that pizza is yummy.</td>
</tr>
</tbody>
</table>
### B: Experiment 2 Items

<table>
<thead>
<tr>
<th>Story</th>
<th>Question</th>
<th>Answer</th>
<th>Expected Truth Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Cat is eating a sandwich.</td>
<td>What is Cat eating?</td>
<td>Cat is eating a sandwich.</td>
<td>T¹²</td>
</tr>
<tr>
<td>2. Elephant is playing with cards.</td>
<td>What is Elephant playing with?</td>
<td>Elephant is playing with dominoes.</td>
<td>F¹³</td>
</tr>
<tr>
<td>3. Bat is stargazing at night.</td>
<td>When is Bat stargazing?</td>
<td>Bat is stargazing in the morning.</td>
<td>F</td>
</tr>
<tr>
<td>4. Horse is at the beach.</td>
<td>Where is Horse?</td>
<td>Horse is at the beach.</td>
<td>T</td>
</tr>
</tbody>
</table>

**Table 12** Experiment 2 training stimuli

<table>
<thead>
<tr>
<th>Story</th>
<th>Question</th>
<th>Answer</th>
<th>Target reading</th>
<th>Expected truth value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. It’s Horse’s birthday and he is having a party at the park. Mouse is at the party because Horse is her best friend, and Bat is at the party to play with the balloons. Both Mouse and Bat are at the party because they want to eat birthday cake.¹⁴</td>
<td>Why are Mouse and Bat at the party?</td>
<td>Mouse is at the party because Horse is her best friend, and Bat is at the party to play with the balloons.</td>
<td>PL¹⁵</td>
<td>T</td>
</tr>
<tr>
<td>2. Horse and Mouse are taking pictures of flowers in the park. Horse is taking pictures of the flowers because the flowers are blooming. Mouse is taking pictures of the flowers because she sees more yellow flowers than ever before. Both Horse and Mouse are taking pictures because the flowers are so pretty.</td>
<td>Why are Horse and Mouse taking pictures of the flowers?</td>
<td>Horse is taking pictures of the flowers because the flowers are blooming, and Mouse is taking pictures of the flowers because she sees more yellow flowers than ever before.</td>
<td>PL</td>
<td>T</td>
</tr>
<tr>
<td>3. Mouse and Horse are swimming in the ocean. Mouse wants to go sit on the beach because she is tired of swimming. Horse wants to go sit on the beach because the water is too cold. Both Mouse and Horse want to sit on the beach to keep building the big sand castle.</td>
<td>Why do Mouse and Horse want to sit on the beach?</td>
<td>Mouse wants to sit on the beach because she is tired of swimming, and Horse wants to sit on the beach because the water is too cold.</td>
<td>PL</td>
<td>T</td>
</tr>
</tbody>
</table>

---

¹² True  
¹³ False  
¹⁴ Pair-list/collective stories are counterbalanced for order of information presentation. The counterbalanced alternative for this story is “It’s Horse’s birthday and he is having a party at the park. Mouse and Bat are at the party because they want to eat birthday cake. Mouse is also at the party because Horse is her best friend, and Bat is also there to play with the balloons.” All other alternatives follow the same order.  
¹⁵ Pair-list  
¹⁶ Collective
4. Cat and Bat are on a sailboat. Cat is on the sailboat because he wants to learn how to sail, and Bat is on the sailboat because he wants to see the fish jumping out of the water. Both Cat and Bat are on the sailboat because they want to enjoy the sunny day.

   Why are Cat and Bat on the sailboat?
   Cat is on the sailboat because he wants to learn how to sail, and Bat is on the sailboat because he wants to see the fish jumping out of the water.
   PL | T

   Cat and Bat are on the sailboat because they want to enjoy the sunny day.
   C | T

5. Cat and Elephant are eating some yummy pizza. Cat likes pizza with veggies more than pepperoni pizza because veggies are his favorite food. Elephant likes pizza with veggies more than pepperoni pizza because veggies are healthy. Both Cat and Elephant like pizza with veggies more than pepperoni pizza because pepperoni is too spicy.

   Why do Cat and Elephant like pizza with veggies more than pepperoni pizza?
   Cat likes pizza with veggies more than pepperoni pizza because veggies are his favorite food, and Elephant likes pizza with veggies more than pepperoni pizza because veggies are healthy.
   PL | T

   Cat and Elephant like pizza with veggies more than pepperoni pizza because pepperoni is too spicy.
   C | T

6. Horse and Mouse are dancing together. Horse is dancing because he likes the music. Mouse is dancing because dance is a good form of exercise. Both Horse and Mouse are dancing because they like to wear tutus in dance class.

   Why are Horse and Mouse dancing?
   Horse is dancing because he likes the music, and Mouse is dancing because dance is a good form of exercise.
   PL | T

   Horse and Mouse are dancing because they like to wear tutus in dance class.
   C | T

---

**Table 13 Experiment 2 Pair-list/collective stimuli**

<table>
<thead>
<tr>
<th>Story</th>
<th>Question</th>
<th>Answer</th>
<th>Target reading</th>
<th>Expected truth value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mouse and Bat are at an art museum with lots of sculptures and paintings. Mouse says she is happy, and Bat asks her why. Mouse says she is happy to be at the museum because there are so many interesting sculptures.</td>
<td>Why does Mouse say she is happy to be at the museum?</td>
<td>Mouse says she is happy to be at the museum because Bat asked her.</td>
<td>SD&lt;sup&gt;17&lt;/sup&gt;</td>
<td>T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mouse says she is happy to be at the museum because there are so many interesting sculptures.</td>
<td>LD&lt;sup&gt;18&lt;/sup&gt;</td>
<td>T</td>
</tr>
<tr>
<td>2. Cat is walking home from the park. He sees his friend Mouse standing on the sidewalk. Cat says hello to Mouse, and asks Mouse about their friend Bat. Mouse says that Bat is sad because his bicycle broke.</td>
<td>Why does Mouse say that Bat is sad?</td>
<td>Mouse says that Bat is sad because Cat asked her.</td>
<td>SD</td>
<td>T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Mouse says that Bat is sad because his bicycle broke.</td>
<td>LD</td>
<td>T</td>
</tr>
</tbody>
</table>

<sup>17</sup> Short distance  
<sup>18</sup> Long distance
3. Mouse and Bat are painting. Mouse is painting a car, and Bar is painting flowers. Bat shows Mouse his painting and asks if she likes it. Mouse tells Bat that the painting is beautiful because it has so many colors. Why does Mouse say that Bar’s painting is beautiful? Mouse says that Bar’s painting is beautiful because Bat asked her. SD T

4. Elephant and Horse are on a camping trip. After they go for a hike, they build a campfire and tell stories about a witch. Elephant says that the witch can fly because she has magic powers. Why does Elephant say that the witch can fly? Elephant says that the witch can fly because she has magic powers. LD T

5. Mouse and Elephant are listening to music together. Mouse says that the music is too loud and tells Elephant to turn it down, but Elephant doesn’t hear. So Mouse loudly tells Elephant to turn the music down because it is hurting her ears. Why does Mouse loudly tell Elephant to turn the music down? Mouse loudly tells Elephant to turn the music down because Elephant didn’t hear the first time. SD T

6. Elephant and Bat are reading at the library. Elephant really likes her book because it has lots of pictures, and whispers to Bat that she likes it. Bat doesn’t hear, so Elephant tells Bat that she likes the book because it is about an exciting adventure. Why does Elephant loudly say that she likes her book? Elephant loudly says that she likes her book because Bat didn’t hear her when she whispered. LD T

Table 14 Experiment 2 embedded-question stimuli

<table>
<thead>
<tr>
<th>Story</th>
<th>Question</th>
<th>Answer</th>
<th>Expected truth value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Horse and Cat to go to the park and play. Horse loves to go to the park because there is a playground with swings there. The swings are Horse’s favorite.</td>
<td>Why are Horse and Cat going to the park?</td>
<td>Horse and cat are going to the park because they want to play.</td>
<td>T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Horse and Cat are going to the park because the monkey bars are Horse’s favorite.</td>
<td>F</td>
</tr>
<tr>
<td>2. Cat and Elephant spent the whole day together. They read some books at the library, went swimming in the lake, and then played at Elephant’s house. Now, they are saying goodbye, and Cat is going home to water his plants.</td>
<td>Why is Cat going home?</td>
<td>Cat is going home to water his plants.</td>
<td>T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cat is going home to make dinner.</td>
<td>F</td>
</tr>
<tr>
<td>3. Bat needs to wake up early tomorrow to go to his soccer game. Bat gets his homework done, eats dinner, and takes a bath. He goes to sleep early tonight, so he won’t be tired tomorrow.</td>
<td>Why does Bat go to sleep early?</td>
<td>Bat goes to sleep early so he won’t be tired tomorrow.</td>
<td>T</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bat goes to sleep early because he didn’t sleep well last night.</td>
<td>F</td>
</tr>
</tbody>
</table>
Elephant wants to jump higher and higher.

Why is Elephant waiting to fly his kite?

Bat is waiting to fly his kite because there is no wind outside today.

Bat is waiting to fly his kite because he wants to do it with a friend.

Horse got a new hula hoop, but he doesn’t know how to use it. Bat is really good at hula hooping, so he decides to help horse out. Horse watches bat hula hoop so that he can learn how, too.

Why does Horse watch Bat hula hoop?

Horse watches Bat hula hoop so that he can learn how, too.

Horse watches Bat hula hoop because Bat is showing him a cool trick.

Horse is at the park. He is setting up a picnic for his friends. He brought a basket with snacks inside and a picnic blanket. Horse is at the park early because he wants to surprise his friends.

Why is Horse at the park early?

Horse is at the park early because he wants to surprise his friends.

Horse is at the park early because he forgot what time he was supposed to meet his friends.

Cat is at the apple orchard. He has been picking apples all day. Cat stops picking apples now, because his basket is full.

Why does Cat stop picking apples now?

Cat stops picking apples now because his basket is full.

Cat stops picking apples now because there are no apples left on the trees.

Elephant is getting ready to walk to school. It is cold and snowing outside, so she puts on a hat to keep herself warm and dry.

Why does Elephant put on a hat?

Elephant puts on a hat to keep herself warm and dry.

Elephant puts on a hat because her favorite color is red.

<table>
<thead>
<tr>
<th>Experiment 2 short distance why control stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Story</strong></td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>1. Cat is ordering ice cream at the ice cream shop. He looks at all of the delicious flavors. Cat thinks for a second, and then says that he wants chocolate, which is his favorite flavor.</td>
</tr>
<tr>
<td>2. Elephant is playing a game at the arcade. She plays it again and again, but she can’t beat it yet. It’s really hard! Elephant says she will go home after she wins.</td>
</tr>
<tr>
<td>3. Cat rides the bus in the morning. He needs to get to school. He wants to get off at the park, though, to look at all the beautiful nature there.</td>
</tr>
<tr>
<td>4. Elephant loves to jump. Usually, she just jumps on the floor, but Mouse surprised her with a new gift: her very own trampoline! Elephant is jumping on the new trampoline now. She says that she wants to jump higher and higher.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 16 Experiment 2 long distance other wh-question stimuli</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Story</strong></td>
</tr>
<tr>
<td>-----------</td>
</tr>
<tr>
<td>1. Cat is ordering ice cream at the ice cream shop. He looks at all of the delicious flavors. Cat thinks for a second, and then says that he wants chocolate, which is his favorite flavor.</td>
</tr>
<tr>
<td>2. Elephant is playing a game at the arcade. She plays it again and again, but she can’t beat it yet. It’s really hard! Elephant says she will go home after she wins.</td>
</tr>
<tr>
<td>3. Cat rides the bus in the morning. He needs to get to school. He wants to get off at the park, though, to look at all the beautiful nature there.</td>
</tr>
<tr>
<td>4. Elephant loves to jump. Usually, she just jumps on the floor, but Mouse surprised her with a new gift: her very own trampoline! Elephant is jumping on the new trampoline now. She says that she wants to jump higher and higher.</td>
</tr>
</tbody>
</table>