Quantifying into *wh*-dependencies:
Composing multi-*wh* questions and questions with a quantifier

Yimei Xiang, Rutgers University
(Date: March 19, 2021)

Abstract *Wh*-questions with a quantifier have readings that seemingly involve quantification-into questions (called 'QɪQ' for short). This paper argues to unify the derivation of QɪQ-readings and distinguish these readings from pair-list readings of multi-*wh* questions. I propose that QɪQ-questions and pair-list multi-*wh* questions both involve a *wh*-dependency, namely, that the quantificational/*wh*-subject stands in a dependency relation with the *wh*-object. In particular, in a pair-list multi-*wh* question, the *wh*-subject binds in across an identity condition; in a QɪQ-question, the quantificational subject binds in across a predication condition. The subtle differences between these two quantificational binding operations explain a contrast in domain exhaustivity between multi-*wh* questions and questions with a universal quantifier. I further propose that the seeming QɪQ-effect in a QɪQ-question is derived by extracting one of the minimal proposition sets that satisfy the aforementioned quantificational predication condition. The values of these sets determine whether the QɪQ-reading is available and whether a QɪQ-question admits pair-list answers and/or has a choice flavor.

Keywords Questions, quantifiers, multi-*wh*, pair-list, functionality, uniqueness, domain exhaustivity, quantificational variability, categorial approaches, compositionality

1. Introduction

Wh-questions with a universal (∀-)quantifier (called '∀-questions' henceforth) are sometimes ambiguous between individual readings, functional readings, and pair-list readings (Engdahl 1980, 1986). As exemplified in (1), the three readings call for answers that name an atomic movie, a Skolem function to atomic movies, and a list of boy-movie pairs, respectively.

(1) Which movie did every/each boy watch?
  a. 'Which movie y is s.t. every boy watched y?' '*Spiderman*.' (Individual)
  b. 'Which function f to atomic movies is s.t. every boy x_i watched $f(x_i)$?' (Functional)
  'His$_i$ favorite superhero movie.'
  c. 'For every boy x, [you tell me] which movie did x watch?' (Pair-list)
  'Andy watched *Ironman*, Billy watched *Spiderman*, Clark watched *Hulk*.'

In general, there are two directions to analyze the pair-list reading (1c). One is to assume that this reading involves 'quantification-into questions (QɪQ)' (Groenendijk and Stokhof 1984; Chierchia 1993; a.o.). An informal paraphrase for QɪQ-readings is given in (2), where 'Det' stands for a determiner.

(2) Which movie did Det-boy(s) watch? (QɪQ-reading)
  ≈ 'For Det-boy(s), [you tell me]/[I ask you] which movie did they watch?'

For *wh*-questions with an existential (∃-)quantifier (called '∃-questions'), their QɪQ-readings have a choice flavor (Groenendijk and Stokhof 1984). For example, the choice reading (3b) asks the addressee to choose one/two of the relevant boys and specify the unique movie he/they watched.[1]

[1] In ∃-questions, functional readings are only marginally acceptable. For example, the fragment functional answer (i-a) is

(3) Which movie did one/two of the boys watch?

a. 'Which movie y is s.t. one/two of the boys watched y?' '*Ironman*.' (Individual)

b. 'For one/two of the boys, [you tell me] which movie did he/they watch?' (Choice)
'Andy watched *Ironman*.'/ 'Billy and Clark watched *Spiderman*.'

In contrast, *wh*-questions with a negative quantifier (called 'NO-questions' henceforth) do not have QIQ-readings. For example, (4) cannot be responded by silence.

(4) Which movie did {no boy, none of the boys} watch?

a. 'Which movie y is s.t. no boy watched y?' '*Revengers*.' (Individual)

b. 'Which function f to atomic movies is s.t. no boy x_i watched $f(x_i)$?' (Functional)
'The movie recommended by their$_i$ grandfather.'

c. # 'For no boy, [you tell me] which movie did they watch?' [Slience] (✗QIQ)

The other direction to analyze the pair-list reading (1c) is to treat questions with pair-list readings uniformly. Similarly to the ∀-question in (1), the multi-*wh* question in (5) has a reading that calls for an answer that specifies a list of boy-movie pairs. Accounts adopting this line of thinking either use the same LF to derive the pair-list readings of these two questions (Engdahl 1980, 1986; Dayal 1996, 2017) or analyze these two questions with different structures that yield the same root denotation (Fox 2012a,b).

(5) Which boy watched which movie?

a. 'Which unique boy-x-movie-y pair is s.t. x watched y?' (Single-pair)
'Andy watched *Spiderman*.'

b. 'Which boy-x-movie-y pairs are s.t. x watched only y?' (Pair-list)
'Andy watched *Ironman*, Billy watched *Spiderman*, Clark watched *Hulk*.'

In sum, it remains controversial whether we should treat questions with QIQ-readings (abbreviated as 'QIQ-questions') uniformly or questions with pair-list readings (abbreviated as 'pair-list questions') uniformly. This paper pursues the former option: pair-list readings of ∀-questions and multi-*wh* questions have a contrast in domain exhaustivity (Sect. 2.1), which argues that these two types of pair-list questions have different meanings and different structures. Despite this contrast, similarities among these complex questions in form and meaning argue that their composition procedures are not drastically different.

I propose that QiQ-questions and pair-list multi-*wh* questions both involve a '*wh*-dependency', namely, that the quantificational/*wh*- subject stands in a dependency relation with the *wh*-object. The core analysis is sketched in (6). The complex object-trace t^i_j carries a functional index j bound by the *wh*-object as well as an argument index i bound by the quantificational/*wh*- subject. In the multi-*wh* question (6b), the *wh*-subject binds into the trace of the *wh*-object across an identity (IDENT) condition; in the QIQ-question (6a), the quantificational subject binds into the trace of the *wh*-object across a predication (PRED) condition. Differences between these two operations explain the contrast in domain exhaustivity between these two questions (see Sect. 6).

under-informative; the identity of the boy who watched a movie has to be specified, as in (i-b). I leave this puzzle open.

(i) (Context: Among the relevant boys, only Andy watched a movie, which was his favorite superhero movie *Ironman*.)
Which movie did one of the boys watch?

a. ?? His favorite superhero movie.

b. Andy watched his favorite superhero movie.

(6) A schema of composing complex questions

a. Which boy watched which movie? (Pair-list reading)
... [which-movie$_j$ ... which-boy$_i$ [$_{\text{IDENT}}$ ... [t_i watched t^i_j]]]

b. Which movie did Det-boy(s) watch? (QiQ-reading)
... [which-movie$_j$ ... Det-boy(s)$_i$ [$_{\text{PRED}}$ ... [t_i watched t^i_j]]]

The rest of this paper is organized as follows. Section 2 presents evidence against the view of treating the two types of pair-list questions (i.e., ∀-questions and multi-*wh* questions) uniformly as well as evidence that supports the view of treating QiQ-questions uniformly. Section 3 lays out the technical challenges and relevant facts that this paper aims to account for. Section 4 reviews two influential approaches to composing pair-list questions, including the functionality approach of Dayal (1996, 2017) and the family-of-questions approach of Fox (2012a,b). Section 5 introduces a GB-style categorial approach to composing questions. Section 6 puts forward my central analysis of composing pair-list multi-*wh* questions and QiQ-questions. The denotations and the composition procedures of these two types of complex questions will be presented in tandem. Section 7 accounts for the quantificational (Q-)variability effects in embeddings of pair-list questions. Section 8 concludes. Appendices A and B review another two accounts of composing QiQ-questions.

2. Arguments for unifying the derivation of QiQ-readings

This section argues that pair-list ∀-questions should be composed uniformly as other QiQ-questions, not as their multi-*wh* counterparts. On the one hand, when having pair-list readings, ∀-questions are subject to domain exhaustivity, whereas their multi-*wh* counterparts are not (Sect. 2.1). This contrast argues that these two types of questions should be interpreted and composed differently. On the other hand, QiQ-questions exhibit the same subject–object/adjunct asymmetry. What's more, the distributional pattern of QiQ-readings is preserved in questions where the subject is a coordination of quantifiers (Sect. 2.2). These facts argue that QiQ-questions have a uniform syntax.

2.1. A contrast in domain exhaustivity

It is commonly claimed that pair-list readings of multi-*wh* questions and ∀-questions both exhibit 'domain exhaustivity' (Dayal 1996, 2002; a.o.). For a question with a *wh*/∀-subject and a *wh*-object, the domain exhaustivity condition says that every member of the set quantified over by the *wh*/∀-subject is paired with a member of the set quantified over by the *wh*-object. For instance, in (1) and (5), repeated below, domain exhaustivity requires that <u>every</u> boy watched a (possibly different) movie. Moreover, since the *wh*-object is singular (i.e., the *wh*-complement *movie* is singular), the two questions are also subject to 'point-wise uniqueness', which says that each boy watched <u>at most one</u> movie.

(7) a. Which movie did every/each boy watch?
b. Which boy watched which movie?

The point-wise uniqueness effect is easy to attest, but the domain exhaustivity effect is not so obvious. In the multi-*wh* question (7b), for example, it is unclear which set of boys is quantified over by the *wh*-subject; domain exhaustivity would be trivial if the domain of quantification consists of only the boys who did watch a movie. To remove this confound, Fox (2012a) uses the pair of examples in (8), where the quantification domain of each *wh*-phrase is explicitly specified.[2] Fox claims that (8b) rejects a pair-list reading (in contrast to (8a)), since interpreting this question with a pair-list reading would give rise to a domain exhaustivity condition that is contextually infelicitous — pairing four kids with three chairs implies that there will be multiple kids sitting on the same chair.

(8) a. Guess which one of the **three** kids will sit on which one of the **four** chairs.
    b. Guess which one of the **four** kids will sit on which one of the **three** chairs.

Contrary to the adopted view, I argue that pair-list multi-*wh* questions are not subject to domain exhaustivity. First, multi-*wh* questions can be felicitously uttered in pair-list contexts where domain exhaustivity is violated. In (9), the sentence repeated from (8b) is felicitous and must be interpreted with a pair-list reading.

(9) (Context: Four kids are playing Musical Chairs and are competing for three chairs.)
    Guess which one of the **four** kids will sit on which one of the **three** chairs.
    $\not\rightsquigarrow$ 'Each of the four kids will sit on one of the three chairs.'

The game rules of Musical Chairs yield two conditions: (i) one of the four kids will not sit on any of the three chairs, and (ii) the rest three kids each will sit on a different chair. Condition (ii) ensures that the embedded multi-*wh* question has a pair-list reading, not a single-pair reading. Condition (i) contradicts the domain exhaustivity inference that each of the kids will sit on one of the chairs. If pair-list readings of multi-*wh* questions were subject to domain exhaustivity, (9) would suffer a presupposition failure, contrary to fact.

Second, unlike their multi-*wh* counterparts, pair-list ∀-questions cannot be felicitously used in contexts where domain exhaustivity is violated. In the context in (10), the quantification domain of the quantificational/*wh*- subject is greatly larger than that of the *wh*-object. The multi-*wh* question (10a) is fully acceptable in this context, but the ∀-question (10b) is not: (10b) presupposes that each candidate will get one of the jobs, contrary to the context.

(10) (Context: **100** candidates are competing for **three** job openings.)
    a. ✓ Guess which candidate will get which job.
    b. # Guess which job will every candidate get.

Likewise, in the Musical Chairs scenario, the multi-*wh* question is felicitous, but the corresponding ∀-question is not.

(11) (Context: Four kids are playing Musical Chairs and are competing for three chairs.)
    a. Guess which one of the four kids will sit on which one of the three chairs. =(9)

[2] One might wonder whether specifying the domain of quantification explicitly can sufficiently remove the confound with domain exhaustivity — could there be additional covert domain restrictions with the *wh*-phrases? In (8) and (9), for example, the confound would remain if the quantification domain of *which one of the four kids* is covertly restricted to a subset of the four kids, excluding the kid who will not sit on a chair. I argue that such covert restrictions are not possible once the quantification domain of a *wh*-phrase has been specified explicitly. As seen in (i), uniqueness is assessed relative to a domain containing all the four contextually relevant kids, as in (ia); if the phrase *which one of the four kids* could range over a subset of the four kids, the uniqueness inference would be as weak as (ib), contrary to fact.

(i) Which one of the four kids cried?
    a. $\rightsquigarrow$ 'Among the four kids, only one of them cried.'
    b. $\not\rightsquigarrow$ 'Among a certain subset of the four kids, only one of them cried.'

b. # Guess which one of the three chairs will each of the four kids sit on.

One might argue that the domain exhaustivity condition of a pair-list multi-*wh* question can be associated with any of the *wh*-phrases, including the *wh*-object. For example, in (9) and (10), it could be the case that domain exhaustivity requires every chair and every job to be taken by a kid and a candidate, respectively. This possibility is ruled out as follows: a multi-*wh* question can be uttered in a pair-list context where neither type of domain exhaustivity is satisfied. For example, sentence (12) is felicitous, and it does not imply a domain exhaustivity inference relative to the boys or to the girls.

(12) (Context: Four boys and four girls will form four boy-girl pairs to perform in a dance competition, but only **two** of the pairs will get into the final round.)
Guess which one of the **four** boys will dance with which one of the **four** girls in the final round.
$\not\rightsquigarrow$ 'Each of the four boys will dance with one of the four girls in the final round.'
$\not\rightsquigarrow$ 'Each of the four girls will dance with one of the four boys in the final round.'

In conclusion, pair-list ∀-questions are subject to domain exhaustivity, whereas pair-list multi-*wh* questions are not. This contrast argues that these two types of pair-list questions should be interpreted differently and composed differently.

2.2. Uniform distribution of QɪQ-readings

The distribution of QɪQ-readings uniformly exhibits a subject–object/adjunct asymmetry (May 1985, 1988; Chierchia 1991, 1993). As seen in (13) and (14), pair-list readings and choice readings are available if the non-*wh* quantifier serves as the subject while the *wh*-phrase serves as the object, but not vice-versa. In (13b), the uniqueness inference triggered by the singular *wh*-subject must take wide scope relative to the ∀-object. As for the ∃-questions in (14), despite (14b) marginally admits a choice reading, (14a) is more preferable if the questioner seeks for a choice answer.[3] The subject–adjunct asymmetry is analogous, as illustrated in (15). Hence, unless there is compelling evidence to suggest otherwise, it is plausible to assume that QɪQ-readings are derived uniformly.

(13) (Context: Ten students made votes for three candidates. Each student voted for only one candidate. The questioner wants to know all of the student-candidate pairs.)
a. Which candidate did every student vote for? (✓Pair-list)
b. # Which student voted for every candidate? (✗Pair-list)
$\rightsquigarrow$ 'Exactly one of the students voted for every candidate.'

(14) (Context: Ten students made votes for three candidates. Each student voted for only one candidate. The questioner is only interested in knowing one of the student-candidate pairs.)
a. Which candidate did one of the students vote for? (✓Choice)
Andy voted for the first candidate.
b. ? Which student voted for one of the candidates? (?Choice)

(15) (Context: Each driver refueled at a nearby station exactly once.)
a. At which station did every driver refuel? (✓Pair-list)
b. # Which driver refueled at every gas station? (✗Pair-list)

[3]The reason why (14b) and (15d) marginally admit choice readings might be that indefinites have more ways to take wide scope than ∀-quantifiers, such as through globally bound choice functions.

c. At which station did [one of the drivers] refuel? (✓Choice)

d. ? Which driver refueled at [one of the nearby stations]? (?Choice)

The idea of treating QɪQ-readings uniformly is further supported by the blocking effect of negative quantifiers. In (16a) where the subject is a conjunction of a ∀-quantifier and an ∃-quantifier, the pair-list reading associated with the ∀-quantifier and the choice reading associated with the ∃-quantifier are both preserved. This question asks the addressee to specify all of the boy-watch-movie pairs and one of the girl-watch-movie pairs. In contrast, since negative quantifiers do not participate in QɪQ-readings (as seen in (4)), coordinating a ∀/∃-quantifier with a negative quantifier blocks the QɪQ-reading. For example, (16b) does not have the reading that requests the addressee to list all the boy-watch-movie pairs and not to list any teacher-watch-movie pairs. (For an explanation based on 'LF Efficiency', see Sect. 6.4.3.)

(16) a. Which movie did [each of the boys and one of the girls] watch? (✓QɪQ)

b. Which movie did [each of the boys and none of the teachers] watch? (✗QɪQ)

c. Which movie did [one of the girls and none of the teachers] watch? (✗QɪQ)

3. Challenges and goals

Section 2 has laid out two goals of this paper: (i) to compose QɪQ-questions uniformly, and (ii) to compose pair-list multi-*wh* questions and pair-list ∀-questions in tandem while explaining their contrast in domain exhaustivity. It is not easy to achieve both goals — a proper solution needs to overcome a few technical challenges and account for a number of semantic effects.[4]

First, for most frameworks of question semantics, the structure in (17) is ill-formed. The generalized quantifier 'Det-boy' takes arguments of type $\langle e, t\rangle$; thus it can only quantify into a t-type expression. However, the contained question *which movie did x watch* is not of type t; it is typically treated as a set of propositions as in Hamblin-Karttunen Semantics, or as a one-place predicate/property as in categorial approaches.

(17) Which movie did Det-boy(s) watch?
*[Det-boy(s) λx_e [which movie did x watch]]

There are two general strategies to solve this type-mismatch problem. One is to extract the domain of quantification of the quantificational subject via a type-shifting operation (Groenendijk and Stokhof 1984; Chierchia 1993; Dayal 1996, 2017; a.o.). For example, Dayal extracts the quantification domain of a ∀-quantifier as extracting the unique minimal witness set of the quantifier. This strategy is feasible in principle but a bit *ad hoc* (see Sect. 4.1.2 and footnote 14).

The other strategy is to create a t-type constituent in the LF that the quantifier can quantify into standardly. For example, in Partition Semantics (Groenendijk and Stokhof 1984), which defines the root denotation of a question as a partition of possible worlds, the formation of a partition involves a t-type node expressing the equivalence of two extensions. Alternatively, Karttunen (1977) and Krifka (2001) recast quantifying-into questions as quantifying-into question-embeddings. The two analyses based on partitions and question-embeddings overcome the type-mismatch problem but bring other problems (reviewed in Appendices A and B). In contrast, my proposal will follow

[4]This paper does not attempt to explain effects that are more likely to be related to syntax in nature, such as superiority effects and constraints on extractions/movements. See Kotek 2014, 2019 and the references therein for detailed discussions.

Fox (2012b) in assuming that the root of a QɪQ-question contains a t-type node that expresses a predication condition (Sect. 4.2 and 6.4).[5]

Second, pair-list readings have a limited distribution in matrix QɪQ-questions: only *each*/*every*-phrases can license pair-list readings for matrix questions. For example, to the ∃2-question in (18), the seeming pair-list answer (18a), which distributes over the two chosen students, is actually an over-informative specification of the cumulative choice answer (18b) (Moltmann and Szabolcsi 1994; Szabolcsi 1997a). Questions with a plural *the*-phrase like (19) are analogous (Srivastav 1991; Krifka 1991; for a different view by Johnston 2019, see Sect. 6.4.5).

(18) Who did two of the students vote for?
  a. Andy voted for Mary, and Billy voted for Jill.
  b. Andy and Billy voted for Mary and Jill. In particular, Andy voted for Mary, and Billy voted for Jill.

(19) Who did the students vote for?

The confound from cumulative answers can be removed by replacing *who* with a singular *wh*-phrase, which triggers a uniqueness presupposition. In the following matrix questions, distributivity above uniqueness is possible only in (20a,b), where the subject is distributive in lexicon. In contrast, for example, the ∃2-question (20d) presupposes that two of the students voted for the same candidate and only this candidate, which contradicts the context. (In (20d-f), '~~EACH~~' means that the reading doesn't involve covert distributivity between the subject and the uniqueness inference triggered by the singular *wh*-object.)

(20) I know that every student voted for a different candidate. Which candidate did ...

a.	... every student vote for?	(*every* $\gg \iota$)
b.	... each student/ each of the students vote for?	(*each* $\gg \iota$)
c.	# ... all/most of the students vote for?	(*all*/*most* $\gg$ ~~EACH~~ $\gg \iota$)
d.	# ... two of the students vote for?	($\exists 2 \gg$ ~~EACH~~ $\gg \iota$)
e.	# ... two or more students vote for?	($\exists 2+ \gg$ ~~EACH~~ $\gg \iota$)
f.	# ... the students vote for?	(*the-NP*$_{\text{PL}} \gg$ ~~EACH~~ $\gg \iota$)

To account for the above distributional constraints of pair-list readings, many works on composing complex questions propose to derive pair-list readings in a way that crashes in questions with a non-universal quantifier (e.g., Dayal 1996 and Fox 2012b; see Sect. 4.) This strategy, however, comes with an expense of failing to account for the choice readings of ∃-questions. In contrast, I argue that a non-interrogative DP can participate in pair-list readings only if it is lexically distributive and is productive in scoping. In my analysis, the above distributional constraints of pair-list readings naturally follow independently observed contrasts in lexical distributivity and scoping between *every*/*each*-phrases and the other quantifiers (for details, see Sect. 6.4.2 and 6.4.4).

Third, there are several semantic effects robustly observed in QɪQ-questions and pair-list multi-*wh* questions. Section 2.1 has discussed two effects, including the uniqueness effect triggered by the singular *wh*-object, as seen in (21a-c), and the domain exhaustivity effect observed only in ∀-questions, as seen in (21a). These effects were not extensively considered until Srivastav 1991/Dayal 1996.

[5]Other than these two general strategies, Inquisitive Semantics also exempts from this type-mismatch problem because it defines declaratives and interrogatives uniformly as a set of sets of propositions (of type $\langle stt, t\rangle$) and generalized quantifiers as functions of type $\langle\langle e, stt\rangle, t\rangle$. To my knowledge, this idea has not been explored extensively. For a possible direction, see Ciardelli and Roelofsen 2018: Sect. 4.3.3.

(21) a. Which movie did every/each boy watch?
$\rightsquigarrow$ 'For every boy x, x watched exactly one movie.'

b. Which boy watched which movie?
$\rightsquigarrow$ 'For every boy x s.t. x watched a movie, x watched exactly one movie.'

c. Which movie did one/two of the boys watch?
$\rightsquigarrow$ 'For some x s.t. x is one/two of the boys, x watched exactly one movie.'

Moreover, embeddings of pair-list questions exhibit 'quantificational (Q-)variability'. As first observed by Berman (1991), question-embeddings modified by a quantificational adverbial (e.g., *mostly, partly, for the most part, in part*) have a Q-variability inference. As illustrated in (22) and (23), in the paraphrase of this inference, the quantification domain of the matrix quantity adverbial *mostly* can be thought of as (a) a set of propositions (Lahiri 1991, 2002; Cremers 2016), (b) a set of sub-questions (Beck and Sharvit 2002), or (c) a set of individuals or pairs (Xiang 2016, 2019a, 2020; Cremers 2018).

(22) Jill mostly knows [which students left].

a. $\rightsquigarrow$ 'Most p: p is a true proposition of the form ⌜*student-x left*⌝, Jill knows p.'

b. $\rightsquigarrow$ 'Most Q: Q is a question of the form ⌜*whether student-x left*⌝, Jill knows Q.'

c. $\rightsquigarrow$ 'Most x: x is an atomic student and x left, Jill knows that x left.'

(23) Jill mostly knows [$_{\text{PAIR-LIST}}$ $\left\{\begin{array}{l}\text{which movie every boy watched}\\ \text{which boy watched which movie}\end{array}\right\}$].

a. $\rightsquigarrow$ 'Most p: p is a true proposition of the form ⌜*boy-x watched movie-y*⌝, Jill knows p.'

b. $\rightsquigarrow$ 'Most Q: Q is a question of the form ⌜*which movie boy-x watched*⌝, Jill knows Q.'

c. $\rightsquigarrow$ 'Most $\langle x, y\rangle$: $\langle x, y\rangle$ is a boy-movie pair and x watched y, Jill knows that x watched y.'

It is commonly claimed that family-of-questions approaches are advantageous in accounting for the Q-variability inference in (23): defining the embedded pair-list question as a family of sub-questions, the Q-variability inference can be defined as in (23b). In contrast, assuming a categorial approach to defining and composing questions, I argue that this inference can be derived as in (23c), which is compatible with a simple functionality approach (for details, see Sect. 7).

4. Two general approaches to composing complex questions

There is a rich literature on composing pair-list multi-*wh* questions and questions with a quantifier. This section reviews two lines of approaches that have tackled both types of questions, including the 'functionality approaches', which assume that these complex questions involve a *wh*-dependency, and the 'family-of-questions approaches', which define each such question as a family of sub-questions.[6]

In this review section, I will focus on two influential accounts by Dayal (1996, 2017) and Fox (2012a,b), which successfully account for the domain exhaustivity and point-wise uniqueness effects in ∀-questions with singular-*wh*. My analysis will take ingredients from these two accounts. For extensive literature reviews, see the appendices of this paper, Xiang 2016: Chap. 5 and 6, Dayal 2017: Chap. 4, and Ciardelli and Roelofsen 2018.

[6]The core assumptions of these two approaches are compatible with each other. For example, Chierchia (1993) assumes a *wh*-dependency while defining a QIQ-question as a family of questions. For more details, see footnote 14.

4.1. Functionality approaches

Wh-questions with a functional reading (called 'functional questions') express a dependency relation between the non-*wh*-subject and the *wh*-object/adjunct. In (24), the fragment answer contains a pronoun interpreted as being bound by the quantificational subject in the question.

(24) Which movie did every-boy$_i$ watch?
His$_i$ favorite superhero movie.

As for pair-list questions, functionality approaches assume that pair-list readings also involve a dependency relation between the ∀/*wh*-subject and the *wh*-object. For example, the pair-list answer (25a) is thought of as the specification of the 'graph' of the function (25b): it pairs elements of the set that the ∀/*wh*-subject ranges over with elements of the set that the *wh*-object ranges over.

(25) Which movie did every boy watch?/ Which boy watched which movie?

a. Andy watched *Ironman*,
Billy watched *Spiderman*,
Clark watched *Hulk*.

b. $f = \begin{bmatrix} a & \rightarrow & i \\ b & \rightarrow & s \\ c & \rightarrow & h \end{bmatrix}$

Functionality approaches were originally proposed for ∀-questions only (Engdahl 1980, 1986; Chierchia 1993). The primary goal was to account for the subject–object/adjunct asymmetry uniformly observed in functional readings and pair-list readings, as illustrated in the following:

(26) Which woman did every boy invite? (✓Individual, ✓Functional, ✓Pair-list)
a. Anna.
b. His mother.
c. Andy invited Mary, Billy invited Susi, Clark invited Jill.

(27) Which woman invited every boy? (✓Individual, ✗Functional, ✗Pair-list)
a. Anna.
b. # His mother. (Intended: 'Every-boy$_i$ was invited by his$_i$ mother.')
c. # Andy invited Mary, Billy invited Susi, Clark invited Jill.

Assuming functionality, one can explain this asymmetry in terms of constraints on dependencies/binding (Chierchia 1993; Jacobson 1994; Williams 1994). For example, Chierchia (1993) argues that weak crossover arises if the object/adjunct binds into the trace of the *wh*-subject.[7]

Further, Dayal (1996, 2017) extends the functionality approach to pair-list multi-*wh* questions. She observes that the corresponding relation expressed by a pair-list answer is a function: the

[7] Chierchia (1993) assumes that the *wh*-trace carries two indices, namely, a functional index *i* bound by the *wh*-phrase and an argument index *j* co-indexed with the non-interrogative quantifier. To bind the *j*-index carried by the *wh*-trace, the non-interrogative quantifier has to be moved to a position that c-commands this *wh*-trace. As in (ib), contrasted by (ia), if the quantifier *every boy* is moved from a position lower than the *wh*-trace, it inevitably moves across a co-indexed expression (viz., the *wh*-trace), causing weak crossover.

(i) a. Which movie did every boy watch?
[which-movie$_i$ ... [every-boy$_j$ ... [t_j watched t_i^j]]] (No crossover)
b. Which boy watched every movie?
*[which-boy$_i$ ... [every-movie$_j$ ... [t_i^j watched t_j]]] (With weak crossover)

correspondence can be one-to-one or many-to-one, but not one-to-many, as witnessed in (28).[8] See also Caponigro and Fălăuş 2020 for an application to multi-*wh* free relatives in Romanian.

(28) Which student talked to which professor? (Dayal 2017: 96)

a. Alice talked to Professor Carl, and Bill talked to Professor Dan.

b. Alice and Bill both talked to Professor Carl.

c. # Alice talked to Professors Carl and Dan.

Assuming functionality, my proposal inherits the advantages of explaining the subject–object/adjunct asymmetry and the unavailability of one-to-many relations in terms of constraints on functionality. Moreover, in Sect. 6, I will show that *wh*-dependencies are independently needed to account for the contrast in domain exhaustivity between multi-*wh* questions and ∀-questions.

4.1.1. *Wh*-dependency in basic functional questions

In the current dominant analysis, *wh*-dependencies in functional questions are derived by assuming a complex *wh*-trace (Groenendijk and Stokhof 1984; Chierchia 1993; a.o.).[9] The tree diagram in (29) illustrates the LF schema to compose a functional ∀-question.[10] In this LF, the *wh*-trace t_i^j carries two indices, namely, an intensional functional index i (of type $\langle s, ee\rangle$) bound by the fronted *wh*-object and an argument index j (of type e) co-indexed with the trace of the quantificational subject. With such indexations, the VP denotes an open sentence expressing a dependency relation between the two arguments of *watched*, and the IP denotes a universal inference over this dependency, read as 'every boy x watched $\boldsymbol{f}_i(x)$'. Details of composition above IP are omitted for now because they vary by the framework of question composition. I will add more details in Sect. 5.

(29) Which movie did every boy watch? (Functional reading)

[8] One might wonder why treating pair-list readings as special functional readings, not vice versa. The reason is that pair-list readings are subject to more constraints than functional readings. As seen in (i), multi-*wh* questions are congruent with fragment answers that are lists of pairs but not with intensional functional answers (Kang 2012; Sharvit and Kang 2017). If pair-list readings were more general than functional readings, we wouldn't expect such a gap.

(i) Which boy watched which movie?

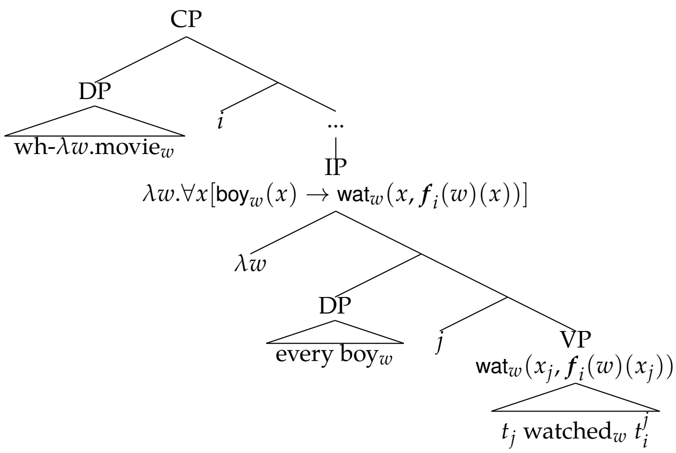
a. # His favorite superhero movie.

b. Andy, *Ironman*, Billy, *Spiderman*, Clark, *Hulk*.

Sharvit and Kang (2017) provide an explanation to why pair-list questions do not admit intensional functional answers. However, the syntax of multi-*wh* questions assumed by Sharvit and Kang is quite different from mine. I leave this issue open.

[9] In contrast to the complex-trace approach, Jacobson (1999, 2014) develops a variable-free approach to functionality which does not make use of indices. In her analysis, functionality is derived by a type-shifting rule, called 'the **z**-rule', which closes off the dependency between the arguments of a predicate. (For example, $\mathbf{z}([\![watched]\!]^w) = \lambda \boldsymbol{f}_{\langle e,e\rangle}\lambda x_e.[\![watched]\!]^w(x, \boldsymbol{f}(x))$.) This approach is especially advantageous in tackling cases where the *wh*-dependent is insitu or is inside an island. For ease of comparing with existing works on composing complex questions, this paper follows the complex-trace approach.

[10] Following Groenendijk and Stokhof (1984), I translate LF representations into the Two-sorted Type Theory (Ty2) of Gallin (1975). Compared with Montague's Intensional Logic, Ty2 is different in that it introduces s (the type of possible worlds) as a basic type (just like e and t), and in that it uses variables and constants of type s which can be thought of as denoting possible worlds. For example, the English common noun *boy* is translated into boy_w in Ty2, where boy is a property of type $\langle s, et\rangle$ and w a world variable of type s. With these assumptions, Ty2 can make direct reference to worlds and allows quantification and abstraction over world variables.



4.1.2. Dayal (1996, 2017) on composing pair-list questions

Dayal (1996, 2017) assumes that the two pair-list questions in (30) uniformly denote a set of conjunctive propositions, and that each of these conjunctive propositions specifies an $\langle e, e\rangle$-type function f from the quantification domain of the $\forall/wh$-subject (i.e., $\mathsf{boy}_@$) to the quantification domain of the *wh*-object (i.e., $\mathsf{mov}_@$).[11] This denotation yields domain exhaustivity since f is defined for every boy.

(30) Which movie did every boy watch?/ Which boy watched which movie?
(The discourse domain has two relevant boys b_1,b_2 and two relevant movies m_1,m_2.)

$$[\![Q_\forall]\!] = [\![Q_{\text{multi-wh}}]\!] = \{\bigcap\{\lambda w.\mathsf{wat}_w(x, f(x)) \mid \mathsf{boy}_@(x)\} \mid f \in [\mathsf{boy}_@ \rightarrow \mathsf{mov}_@]\}$$
$$= \left\{\begin{array}{l} \lambda w.\mathsf{wat}_w(b_1, m_1) \wedge \mathsf{wat}_w(b_2, m_1) \\ \lambda w.\mathsf{wat}_w(b_1, m_1) \wedge \mathsf{wat}_w(b_2, m_2) \\ \lambda w.\mathsf{wat}_w(b_1, m_2) \wedge \mathsf{wat}_w(b_2, m_1) \\ \lambda w.\mathsf{wat}_w(b_1, m_2) \wedge \mathsf{wat}_w(b_2, m_2) \end{array}\right\}$$

Dayal assumes that the two pair-list questions in (30) are composed uniformly as in (31). In this LF, the quantificational/*wh*- subject and the *wh*-object are moved to the specifier of a single functional head $\mathrm{C}^0_{\text{FUNC}}$, and they each are turned into a set of entities via a type-shifting ($\uparrow_{\text{TS}}$) operation. The composition precedes in three steps:

(i) The object-trace carries an extensional functional index i (of type $\langle e, e\rangle$) as well as an argument index j (of type e) that co-refers with the subject-trace. Abstracting these two indices at IP yields the property (32a), which maps an $\langle e, e\rangle$-type function and an individual to a dependency proposition (i.e., an open proposition that expresses a dependency relation between the two arguments of *watched*).

(ii) The functional head $\mathrm{C}^0_{\text{FUNC}}$ introduces domain and range for the function f and creates a 'graph' for f. If q (of type $\langle ee, est\rangle$) is the denotation of IP, the graph of f yielded based on q is the conjunction of the propositions with the form $\ulcorner q(f)(x)\urcorner$ where x is in the domain of f.

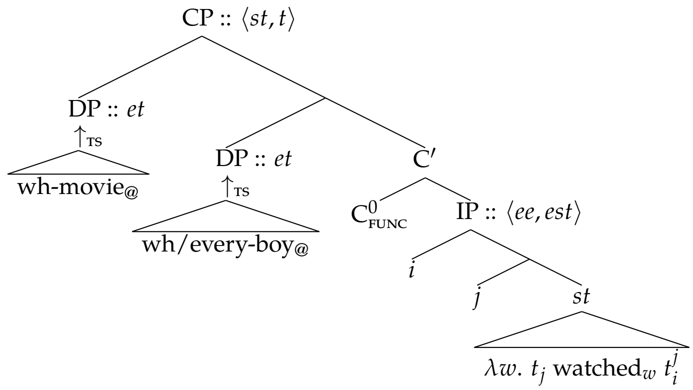
(iii) The sets that the $\forall/wh$-phrases range over are extracted by type-shifting operations.[12] These sets saturate the range and domain arguments introduced by $\mathrm{C}^0_{\text{FUNC}}$.

[11] For simplicity, I assume that the extensions of *wh*-complements are evaluated relative to the actual world '@'.

[12] Dayal (2017) considers two ways to obtain the quantification domain of a *wh*-phrase. One way is to define a *wh*-phrase as an $\exists$-quantifier and extract out its quantification domain via the application of a Be-shifter (Partee 1986). The other way is to define a *wh*-phrase as a set of entities and derive its quantificational meaning via employing an $\exists$-shifter.

With this composition, the CP is interpreted as a set of conjunctive propositions, each of which specifies an $\langle e,e\rangle$-type function that is only defined for the set that the *wh*/∀-subject ranges over.

(31)



(32)
a. $[\![\text{IP}]\!] = \lambda \boldsymbol{f}_{\langle e,e\rangle}\lambda x_e \lambda w.\mathsf{wat}_w(x, \boldsymbol{f}(x))$

b. $[\![\text{C}^0{}_{\text{FUNC}}]\!] = \lambda q_{\langle ee,est\rangle}\lambda D\lambda R\lambda p.\exists \boldsymbol{f} \in [D \to R][p = \bigcap \lambda p'.\exists x \in D[p' = q(\boldsymbol{f})(x)]]$
$= \lambda q_{\langle ee,est\rangle}\lambda D\lambda R.\{\bigcap\{q(\boldsymbol{f})(x) \mid x \in D\} \mid \boldsymbol{f} \in [D \to R]\}$

c. $[\![\text{C}']\!] = \lambda D\lambda R\lambda p.\{\bigcap\{\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}(x)) \mid x \in D\} \mid \boldsymbol{f} \in [D \to R]\}$

d. $[\![\text{CP}]\!] = \{\bigcap\{\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}(x)) \mid x \in \mathsf{boy}_@\} \mid \boldsymbol{f} \in [\mathsf{boy}_@ \to \mathsf{mov}_@]\}$

To account for the uniqueness effects of singular *wh*-phrases, Dayal defines the answerhood (Ans-)operator as in (33), which presupposes the existence of the strongest true answer. The strongest true answer to a question is the true proposition in the Hamblin set of this question that entails all the true propositions in this Hamblin set.

(33) $\textsc{Ans}_{Dayal}(w)(Q) = \exists p[w \in p \in Q \land \forall q[w \in q \in Q \to p \subseteq q]].$
$\iota p[w \in p \in Q \land \forall q[w \in q \in Q \to p \subseteq q]]$

The following presents how the $\textsc{Ans}_{Dayal}$-operator accounts for the uniqueness effects. The ontology of individuals assumes that a singular noun denotes a set of atomic entities, and that a plural noun ranges over both atomic and sum entities (Sharvy 1980; Link 1983). Adopting this ontology, Dayal argues that the Hamblin set of the singular-*wh* question (34a) consists of only propositions naming an atomic boy, and that the Hamblin set of the corresponding plural-*wh* question (34b) includes also propositions naming a sum of boys. In a discourse where two boys Andy and Bill both watched *Hulk*, the true answers to these two questions are as in (34a′) and (34b′), respectively. Crucially, the answer set (34b′) has a strongest member $\lambda w.\mathsf{wat}_w(a \oplus b, h)$, whereas (34a′) doesn't; thus employing $\textsc{Ans}_{Dayal}(w)$ in (34a) yields a presupposition failure. Hence, question (34a) can only be felicitously uttered in worlds where only one of the boys watched *Hulk*, which explains its uniqueness effect.

(34) (Among the considered boys, only Andy and Billy watched *Hulk* in w.)

a. Which boy watched *Hulk*?

a′. $\{\lambda w.\mathsf{wat}_w(a, h), \lambda w.\mathsf{wat}_w(b, h)\}$

b. Which boys watched *Hulk*?

b′. $\{\lambda w.\mathsf{wat}_w(a, h), \lambda w.\mathsf{wat}_w(b, h), \lambda w.\mathsf{wat}_w(a \oplus b, h)\}$

In a pair-list question, if the *wh*-object is singular, the presupposition of $\textsc{Ans}_{Dayal}$ yields point-wise uniqueness. Take (30) for example: if in w_1 the boy b_1 watched only m_1 but the boy b_2 watched both

m_1,m_2, then the top two propositions in the Hamblin set (30) are both true in w_1. Since neither of the true propositions is stronger than the other, applying $\textsc{Ans}_{Dayal}(w_1)$ yields a presupposition failure.

Dayal's analysis also accounts for the domain exhaustivity and point-wise uniqueness effects in pair-list ∀-questions with a singular *wh*-object: domain exhaustivity is hard-wired into the meaning of C^0_{FUNC}; point-wise uniqueness comes from the conjunctive closure encoded within the meaning of C^0_{FUNC} and the presuppositional $\textsc{Ans}_{Dayal}$-operator. This account also manages to keep the semantic type of questions low (i.e., single/double-*wh* questions and ∀-questions are uniformly of type $\langle st, t\rangle$), leaving space to tackle *wh*-constructions that are more complex than pair-list questions (e.g., *wh*-triangles, multi-*wh* echo questions).

However, this analysis faces a number of problems. On the conceptual side, the composition involves a few *ad hoc* or problematic assumptions. First, in the composition, index abstractions are isolated from the moved phrases. Since here the IP involves multiple abstractions, isolating the abstractions from the moved phrases make the binding relations ambiguous. Second, C^0_{FUNC} is structure-specific and the meaning assumed for it is rather complex. It is unclear why a functional head only appears in particular structures and why it should be interpreted as such. For these reasons, Dayal is not fully satisfied with the use of the complex C^0_{FUNC}; she calls this account the 'crazy C^0 approach'. Last, for the composition of ∀-questions, it is syntactically deviant to move a non-interrogative phrase to the specifier of an interrogative CP (Heim 2012).

This analysis also makes several problematic empirical predictions. (Note that these problems are independent from the assumption of functionality.) First, composing pair-list ∀-questions and multi-*wh* questions with the same LF, this account predicts that these questions are semantically equivalent. However, as argued in Sect. 2.1, the two types of questions differ in domain exhaustivity. As seen in (10), repeated below, only the multi-*wh* question can be felicitously used in a pair-list context that violates domain exhaustivity.

(35) (Context: **100** candidates are competing for **three** job openings.)
   a. Guess which candidate will get which job.
   b. # Guess which job will every candidate get.

To account for the contrast in domain exhaustivity, one might assume a twin C^0_{FUNC} that doesn't force domain exhaustivity. With this assumption, however, it would still remain puzzling why this non-exhaustive C^0_{FUNC} cannot appear in pair-list ∀-questions.

Second, this account does not extend to ∃-questions with a choice reading. As seen in Sect. 3, only *every/each*-phrases can license pair-list readings for matrix questions. To avoid over-generating pair-list readings for matrix ∃-questions, Dayal stipulates that the quantification domain of a non-*wh* quantifier must be obtained by extracting the 'unique' minimal witness set of this quantifier.[13] As illustrated in Table 1, among the listed quantifiers, only the ∀-quantifiers have a unique minimal witness set which is not empty. In contrast, the ∃-quantifier has multiple minimal witness sets. The negative quantifier has a unique minimal witness set, but this set is the empty set. With this stipulation, the LF schema (32) for pair-list questions is unavailable for questions with a non-universal quantifier, which leaves choice readings of ∃-questions unexplained.

[13] Live-on sets and witness sets are defined as follows (Barwise and Cooper 1981): for any π of type $\langle et, t\rangle$, π lives on a set B iff $\pi(C) \Leftrightarrow \pi(C \cap B)$ for any set C; if π lives on B, then A is a witness set of π iff $A \subseteq B$ and $\pi(A)$.

Generalized quantifier π	Minimal witness set(s) of π
every/each boy	$\{a, b, c\}$
one of the boys	$\{a\}$, $\{b\}$, $\{c\}$
no boy	$\varnothing$

Table 1: Illustration of minimal witness sets (with three relevant boys *a*,*b*,*c*)

Moreover, without further constraints, this analysis over-predicts pair-list readings for ∃-questions. As re-illustrated in (36a,b), Dayal composes the two pair-list questions uniformly, except that she uses two distinct type-shifting operations (marked as 'TS1' and 'TS2') to extract the set of boys from *which boy* and *every/each boy*. In this analysis, nothing prevents the corresponding ∃-question from being analyzed with the LF (36c), which gives rise to an unwanted pair-list reading. In syntax, if (36b) is well-formed, (36c) should be well-formed as well. In semantics, since *one of the boys* and *which boy* are semantically equivalent, type-shifting operations available for *which boy* should be equally available for *one of the boys*; therefore, (36c) yields the same pair-list reading as in (36a).

(36) All the following LFs yield a pair-list reading:

a. [TS1(which-movie) [TS1(which-boy) [C^0_{FUNC} [$_{\text{IP}}$ ...]]]] — Multi-*wh* question

b. [TS1(which-movie) [TS2(every/each-boy) [C^0_{FUNC} [$_{\text{IP}}$ ...]]]] — ∀-question

c. [TS1(which-movie) [TS1(one-of-the-boys) [C^0_{FUNC} [$_{\text{IP}}$ ...]]]] — ∃-question

Third, as pointed out by Lahiri (2002), defining a pair-list question as a set of conjunctive propositions, this analysis has difficulties in accounting for the Q-variability effects in embeddings of pair-list questions. For example, sentence (37) implies a quantificational inference, which can be paraphrased as if the matrix adverbial *mostly* quantifies over a set of atomic propositions. However, these atomic propositions cannot be retrieved from the question denotation assumed in (30): we cannot retrieve the atomic propositions directly from the conjunction of these propositions.

(37) Jill mostly knows [$_{\text{PAIR-LIST}}$ { which movie every boy watched / which boy watched which movie }].

⇝ 'Most *p*: *p* is a true proposition of the form ⌜*boy-x watched movie-y*⌝, Jill knows *p*.'

To account for the Q-variability effects, in an unpublished work, Dayal (2016) removes the $\bigcap$-closure in the lexicon of C^0_{FUNC} and defines the root of a pair-list question as a family of sets of propositions. The revised account manages to keep the atomic propositions alive, but it sacrifices the advantage of keeping the semantic type of questions low.

4.2. Family-of-questions approaches

Family-of-questions approaches regard a pair-list question as a set/family of sub-questions (Hagstrom 1998; Preuss 2001; Fox 2012a,b; Nicolae 2013; Kotek 2014; Xiang 2016: Chap. 5; Dayal 2016; a.o.). As exemplified in (38), if a simplex question denotes a set of propositions, a family of questions denotes a set of sets of propositions.[14]

[14] The approaches by Groenendijk and Stokhof (1984) and Chierchia (1993) are also family-of-questions approaches. They define a QiQ-question as a family of sub-questions ranging over a minimal witness set (MWS) of the subject quantifier, as in (i). ('$\mathcal{P}_{\text{boy}_@}$' stands for a generalized quantifier ranging over the set of atomic boys. 'MWS($\mathcal{P}_{\text{boy}_@}$, A)' means that A is a minimal witness set of $\mathcal{P}_{\text{boy}_@}$.)

(38) (The discourse domain has two relevant boys b_1,b_2 and two relevant movies m_1,m_2.)
Which movie did every boy watch?/ Which boy watched which movie?

$$
\begin{aligned}
[\![\mathrm{Q}_\forall]\!] = [\![\mathrm{Q}_{\text{multi-}wh}]\!] &= \{[\![\textit{Which movie did x watch?}]\!] \mid x \in \mathsf{boy}_@\} \\
&= \{\{\lambda w.\mathsf{wat}_w(x,y) \mid y \in \mathsf{mov}_@\} \mid x \in \mathsf{boy}_@\} \\
&= \left\{ \begin{array}{l} \{\lambda w.\mathsf{wat}_w(b_1,m_1), \lambda w.\mathsf{wat}_w(b_1,m_2)\} \\ \{\lambda w.\mathsf{wat}_w(b_2,m_1), \lambda w.\mathsf{wat}_w(b_2,m_2)\} \end{array} \right\}
\end{aligned}
$$

The non-flat semantics assumed in (38) easily accounts for the Q-variability inferences of embeddings of pair-list questions. As in (39), such an inference can be defined as if the matrix adverbial *mostly* quantifies over a set of sub-questions of the embedded question.

(39) Jill mostly knows [$_{\text{PAIR-LIST}}$ { which movie every boy watched / which boy watched which movie }].
$\rightsquigarrow$ 'Most Q: Q is a question of the form ⌜*which movie boy-x watched*⌝, Jill knows Q.'

Fox (2012a,b) analyzes the two pair-list questions with different LFs that yield the same root denotation. The LF of a pair-list multi-*wh* question is illustrated in (40). Since *wh*-phrases are treated as indefinites (viz., $[\![\textit{which}]\!] = [\![\textit{some}]\!]$), this LF is read as 'the set of Q s.t. for some boy x, Q is identical to $[\![\textit{Which movie did x watch?}]\!]$', which is simply the set of questions of the form ⌜*Which movie did boy-x watch?*⌝. The composition follows the GB-style Karttunen Semantics (Heim 1995) except that it treats the identity (Id-)operator type-flexible and allows this operator to be iterated.

(40) Which boy watch which movie? (Pair-list reading)
$[_{\text{CP}_2}\ \lambda Q_{\langle st,t\rangle}$ [wh-boy$_@$ λx_e $[_{\text{C}'_2}$[Id Q] $[_{\text{CP}_1}$ λp_{st} [wh-movie$_@$ λy_e $[_{\text{C}'_1}$[Id p] $[_{\text{IP}}$ x watched y]]]]]]]

a. $[\![\text{Id}]\!] = \lambda\alpha_\tau\lambda\beta_\tau.\alpha = \beta$ (τ stands for an arbitrary type)

b. $[\![\text{IP}]\!] = \lambda w.\mathsf{wat}_w(x,y)$

c. $[\![\text{C}'_1]\!] = [\![\text{Id}]\!](p)([\![\text{IP}]\!])$
$= [p = \lambda w.\mathsf{wat}_w(x,y)]$

d. $[\![\text{CP}_1]\!] = \lambda p.\exists y[\mathsf{mov}_@(y) \wedge p = \lambda w.\mathsf{wat}_w(x,y)]$
$= \{\lambda w.\mathsf{wat}_w(x,y) \mid \mathsf{mov}_@(y)\}$

e. $[\![\text{C}'_2]\!] = [\![\text{Id}]\!](Q)([\![\text{CP}_1]\!])$
$= [Q = \{\lambda w.\mathsf{wat}_w(x,y) \mid \mathsf{mov}_@(y)\}]$

(i) $[\![\textit{Which movie did } \mathcal{P}_{\mathsf{boy}_@} \textit{ watch?}]\!]_{\text{QiQ}} = \{[\![\textit{Which member of A watched which movie?}]\!] \mid \textsc{mws}(\mathcal{P}_{\mathsf{boy}_@}, A)\}$

However, the predictions made by these accounts are quite different from the predictions made by the non-flat semantics in (38). For example, Chierchia (1993) defines a sub-question as a set of propositions of the form ⌜*boy-x watched movie-$f(x)$*⌝, as schematized in (ii). The related $\forall/\exists$-questions are thus defined as in (iii).

(ii) $[\![Q_{\mathcal{P}}]\!] = \{\{\lambda w.\mathsf{wat}_w(x, f(x)) \mid x \in A, f \in [A \to \mathsf{boy}_@]\} \mid \textsc{mws}(\mathcal{P}_{\mathsf{boy}_@}, A)\}$

(iii) (The discourse domain has two boys b_1,b_2 and two relevant movies m_1,m_2.)

a. $[\![Q_\forall]\!] = \left\{ \left\{ \begin{array}{l} \lambda w.\mathsf{wat}_w(b_1,m_1), \lambda w.\mathsf{wat}_w(b_2,m_2) \\ \lambda w.\mathsf{wat}_w(b_1,m_2), \lambda w.\mathsf{wat}_w(b_2,m_2) \end{array} \right\} \right\}$

b. $[\![Q_\exists]\!] = \{\{\lambda w.\mathsf{wat}_w(b_1,m_1), \lambda w.\mathsf{wat}_w(b_1,m_2)\}, \{\lambda w.\mathsf{wat}_w(b_2,m_1), \lambda w.\mathsf{wat}_w(b_2,m_2)\}\}$

Chierchia further assumes that answering a family of sub-questions means answering 'one' of the sub-questions (in contrast to Fox's assumption that answering a family of sub-questions means answering 'all' of the sub-questions). Accordingly, since *one of the boys* has multiple minimal witness sets, the QiQ-reading of the $\exists$-question has a choice flavor. Although this account naturally extends to $\exists$-questions, it cannot explain the semantic effects in pair-list $\forall$-questions such as domain exhaustivity and point-wise uniqueness.

f. $[\![\text{CP}_2]\!] = \lambda Q.\exists x[\mathsf{boy}_{@}(x) \wedge Q = \{\lambda w.\mathsf{wat}_w(x,y) \mid \mathsf{mov}_{@}(y)\}]$
$= \{\{\lambda w.\mathsf{wat}_w(x,y) \mid y \in \mathsf{mov}_{@}\} \mid x \in \mathsf{boy}_{@}\}$

The LF of the corresponding pair-list ∀-question is as in (41), read as 'the unique minimal set $\boldsymbol{K}$ s.t. for every boy x: $[\![$*Which movie did x watch?*$]\!]$ is a member of $\boldsymbol{K}$'. The most important operations involved in the formation of this LF are 'quantifying-into-predication' and 'minimization' (*à la* Pafel 1999; Preuss 2001). First, the ∀-subject takes quantifier raising and quantifies into a predication condition, which is yielded by applying a predicative variable $\boldsymbol{K}$ to the open question *Which movie did x watch*. This operation yields a universal predication condition, read as 'for every boy x: $[\![$*Which movie did x watch?*$]\!]$ is a member of $\boldsymbol{K}$'. Next, a (strong) minimization (MIN_S-)operator binds the $\boldsymbol{K}$ variable across the ∀-subject. It applies to the set of $\boldsymbol{K}$ sets that satisfy the universal predication condition and returns the unique minimal $\boldsymbol{K}$ set. This minimal $\boldsymbol{K}$ set is simply the set consisting of exactly all the sub-questions of the form ⌜*Which movie did boy-x watch?*⌝.

(41) Which movie did every boy watch? (Pair-list reading)

$[_{\text{CP}_2}\ \text{MIN}_S\ \lambda \boldsymbol{K}_{\langle stt,t\rangle}$ [every-boy$_{@}$ λx_e [$\boldsymbol{K}$ $[_{\text{CP}_1}\ \lambda p_{st}$ [wh-movie$_{@}$ λy_e [[ID p] $[_{\text{IP}}$ x watched y]]]]]]]

a. $[\![\text{CP}_1]\!] = \{\lambda w.\mathsf{wat}_w(x,y) \mid \mathsf{mov}_{@}(y)\}$ (composition is the same as in (40a-d))

b. $[\![\text{CP}_2]\!] = \text{MIN}_S(\lambda \boldsymbol{K}.[\![\textit{every boy}_{@}]\!](\lambda x.\boldsymbol{K}(\{\lambda w.\mathsf{wat}_w(x,y) \mid \mathsf{mov}_{@}(y)\})))$
$= \text{MIN}_S(\lambda \boldsymbol{K}.\forall x[\mathsf{boy}_{@}(x) \rightarrow \boldsymbol{K}(\{\lambda w.\mathsf{wat}_w(x,y) \mid \mathsf{mov}_{@}(y)\})])$
$= \big\{\{\lambda w.\mathsf{wat}_w(x,y) \mid y \in \mathsf{mov}_{@}\} \mid x \in \mathsf{boy}_{@}\big\}$

(42) $\text{MIN}_S := \lambda \alpha_{\langle \sigma t,t\rangle} : \exists \boldsymbol{K}_{\langle \sigma,t\rangle}[\boldsymbol{K} \in \alpha \wedge \forall \boldsymbol{K}' \in \alpha[\boldsymbol{K} \subseteq \boldsymbol{K}']].\iota \boldsymbol{K}_{\langle \sigma,t\rangle}[\boldsymbol{K} \in \alpha \wedge \forall \boldsymbol{K}' \in \alpha[\boldsymbol{K} \subseteq \boldsymbol{K}']]$
(If α is a set of sets, $\text{MIN}_S(\alpha)$ is the unique minimal set in α which is a subset of every set in α, defined only if this minimal set exists.) (Pafel 1999)

As for the definition of answerhood, Fox (2012a,b) assumes that answering a family of sub-questions amounts to answering all of these sub-questions; in other words, answerhood is applied point-wise and exhaustively. As recursively defined in (43), when applied to a family of sub-questions, the point-wise answerhood-operator imposes ANS_{Dayal} to each sub-question and returns the conjunction of propositions that are the strongest true answer to a sub-question, yielding domain exhaustivity. When the *wh*-object is singular, the presupposition that each of the sub-questions has a strongest true answer also gives rise to point-wise uniqueness.

(43) Point-wise answerhood-operator (Fox 2012a)

$$\text{ANS}_{\text{PW}} := \lambda w \lambda Q. \begin{cases} \text{ANS}_{Dayal}(w)(Q) & \text{if } Q \text{ is of type } \langle st,t\rangle \\ \bigcap\{\text{ANS}_{\text{PW}}(w)(\alpha) \mid \alpha \in Q\} & \text{otherwise} \end{cases}$$

The account of Fox has two advantages over the account of Dayal (1996, 2017). First, as discussed in (39), the non-flat semantics of pair-list questions can easily account for the Q-variability effects in embeddings. Second, the composition is quite neat; it does not use any *ad hoc* type-shifters or any complex operators. In the composition of the pair-list multi-*wh* question, the *wh*-phrases function as ∃-indefinites that quantify into an identity condition. In the composition of the pair-list ∀-question, the ∀-subject standardly composes with a one-place predicate.

However, Fox's analysis has a few empirical problems similarly to Dayal's analysis. First, analyzing pair-list ∀-questions and their multi-*wh* counterparts semantically equivalent, Fox also cannot explain the contrast in domain exhaustivity.[15] Second, Fox's account does not extend to ∃-questions either.

[15] One might propose to reconcile the family-of-questions approach by arguing that pair-list multi-*wh* questions, but not

In the composition of a question with a quantifier, Fox uses the MIN_S-operator to obtain the unique minimal $\boldsymbol{K}$ set that satisfies a quantificational predication condition, which is unavailable if this predication condition is existential. For instance, for the ∃-question (44a), in a discourse with two relevant boys b_1 and b_2, $\boldsymbol{K}$ satisfies the existential predication condition (44b) as long as it is a superset of (44c) or (44d). Among these sets that $\boldsymbol{K}$ may refer to, there isn't one that is a subset of all these sets.

(44) a. Which movie did one of the boys watch?
b. $\exists x[\mathsf{boy}_@(x) \wedge [\![\textit{Which movie did x watch?}]\!] \in \boldsymbol{K}]$
c. $\{[\![\textit{Which movie did } b_1 \textit{ watch?}]\!]\}$
d. $\{[\![\textit{Which movie did } b_2 \textit{ watch?}]\!]\}$

5. The formal theory

I assume a hybrid categorial approach to composing questions, developed in Xiang 2016, 2020. This approach integrates traditional categorial approaches with the GB-style compositional semantics. Compared with frameworks that define questions as sets of propositions (e.g., Hamblin-Karttunen Semantics), categorial approaches define questions as functions over short answers, which allows to derive the Q-variability effects in embeddings of pair-list questions without assuming a non-flat semantics (Sect. 7). However, the core analysis made in Sect. 6 on the composition of the question nucleus is independent from the choice of the framework.

This section lays out only the assumptions that are central to this paper, with some simplifications and modifications. For more details and applications of this framework, see Xiang 2020.

5.1. Defining questions and answers

Following categorial approaches, I define the root denotations of matrix and embedded questions uniformly as functions that map short answers to corresponding propositional answers, as exemplified in (45). After Chierchia and Caponigro (2013), I call such denotations 'topical properties'.[16]

(45) a. $[\![\textit{Which boy came?}]\!] = \lambda x_e : \mathsf{boy}_@(x)\,.\lambda w[\mathsf{came}_w(x)]$
b. $[\![\textit{Which boy came?}]\!]([\![\textit{John}]\!]) = \mathsf{boy}_@(j)\,.\lambda w[\mathsf{came}_w(j)]$

Complete true answers to questions are obtained by the application of the answerhood-operators in (46). Compared with the ANS_{Dayal}-operator (33), the main difference is that the Hamblin set Q is replaced with a topical property $\boldsymbol{P}$, which can supply both short answers and propositional answers.[17] These answerhood-operators account for uniqueness effects in the same way as ANS_{Dayal}.

pair-list ∀-questions, permit covert domain restriction. This possibility has been ruled out by the discussion in Sect. 2.1. First of all, the contrast between the two types of pair-list questions in domain exhaustivity remains even if the domain has been explicitly specified, as seen in (9) and (11). Moreover, as argued in footnote 2, if the quantification domain of a *wh*-phrase has been explicitly specified, it does not take further covert restrictions.

[16] In this paper, λ-terms with presuppositions are represented in the form of $\lambda v_\tau : \beta.\alpha$ (where τ is the semantic type of v, β stands for the additional definedness condition or presupposition, and α stands for the value description). λ-terms without a presupposition are written in the form of $\lambda v_\tau.\alpha$ or $\lambda v_\tau[\alpha]$, whichever is easier to read.

[17] Following Fox (2013), Xiang (2016, 2020) assumes a weaker definition for complete true answers: a true answer to a question is complete as long as it is not asymmetrically entailed by any of the true answers to this question. This answerhood is assumed to account for mention-some readings of questions and free relatives. Since mention-some is not the focus of this paper, for easier comparisons with competing theories in composing complex questions, here I follow Dayal (1996, 2017) and define the complete true answer as the unique strongest true answer. For recent accounts on solving the dilemma between uniqueness and mention-some, see Fox 2018, 2020 and Xiang 2021.

(46) Answerhood-operators (modified from Xiang 2020; to be revised in (60))

a. For the complete true short answer

$$\textsc{Ans}^S(w)(\boldsymbol{P}) = \exists\alpha \in \text{Dom}(\boldsymbol{P})[w \in \boldsymbol{P}(\alpha) \wedge \forall\beta \in \text{Dom}(\boldsymbol{P})[w \in \boldsymbol{P}(\beta) \rightarrow \boldsymbol{P}(\alpha) \subseteq \boldsymbol{P}(\beta)]].$$
$$\iota\alpha \in \text{Dom}(\boldsymbol{P})[w \in \boldsymbol{P}(\alpha) \wedge \forall\beta \in \text{Dom}(\boldsymbol{P})[w \in \boldsymbol{P}(\beta) \rightarrow \boldsymbol{P}(\alpha) \subseteq \boldsymbol{P}(\beta)]]$$

b. For the complete true propositional answer

$$\textsc{Ans}(w)(\boldsymbol{P}) = \boldsymbol{P}(\textsc{Ans}^S(w)(\boldsymbol{P}))$$

5.2. Composing simple *wh*-questions

I define *wh*-phrases as ∃-quantifiers ranging over a polymorphic set. In questions with an extensional reading, the quantification domain of a *wh*-phrase with the form $\ulcorner wh\text{-}\mathrm{A}_w \urcorner$ contains not only elements in the extension of the *wh*-complement $[\![\mathrm{A}]\!]^w$ but also functions from a set of entities to $[\![\mathrm{A}]\!]^w$, as defined in (47a). The semantics of *wh*-phrases in questions with an intensional reading is defined analogously, as schematized in (47b).

(47) a. $[\![wh\text{-}\mathrm{A}_w]\!] = \lambda P.\exists\alpha \in \bigcup \left\{ [\![\mathrm{A}]\!]^w, \{\boldsymbol{f}_{\langle e,e\rangle} \mid \text{Ran}(\boldsymbol{f}) = [\![\mathrm{A}]\!]^w\} \right\} [P(\alpha)]$

b. $[\![wh\text{-}\lambda w.\mathrm{A}_w]\!] = \lambda P.\exists\alpha \in \bigcup \left\{ \begin{array}{l} \{r_{\langle s,e\rangle} \mid \forall w[r_w \in [\![\mathrm{A}]\!]^w]\}, \\ \{\boldsymbol{f}_{\langle s,ee\rangle} \mid \forall w[\text{Ran}(\boldsymbol{f}_w) = [\![\mathrm{A}]\!]^w]\} \end{array} \right\} [P(\alpha)]$

c. For any function $\boldsymbol{f}$ and any set A, $\text{Ran}(\boldsymbol{f}) = A$ iff $\forall x \in \textsc{Dom}(\boldsymbol{f})[\boldsymbol{f}(x) \in A]$.

In the composition of a simplex *wh*-question, the fronted *wh*-phrase is converted into a function domain restrictor via the BeDom-operator (abbreviated as 'bd' in this paper).[18] As defined in (48), if π is an ∃-quantifier, $\textsc{Be}(\pi)$ is the set that π ranges over, and $\textsc{BeDom}(\pi)$ is a function domain restrictor which combines with a function θ and returns the function that is similar to θ but is undefined for items that are not in $\textsc{Be}(\pi)$.

(48) For any π of type $\langle \sigma t, t\rangle$ where σ is an arbitrary type, we have:

a. $\textsc{Be}(\pi) = \lambda x.\pi(\lambda y.y = x)$ (Partee 1986)

b. $\textsc{BeDom}(\pi) = \lambda\theta_\tau.\iota P_\tau \left[\begin{array}{l} [\text{Dom}(P) = \text{Dom}(\theta) \cap \textsc{Be}(\pi)] \\ \quad \wedge \forall\alpha \in \text{Dom}(P)[P(\alpha) = \theta(\alpha)] \end{array} \right]$ (Xiang 2016, 2020)

As exemplified in (49), the fronted '$^{\text{BD}}$(wh-boy$_@$)' applies to the simple 'came'-function defined for all entities and returns a more restrictive 'came'-function only defined for atomic boys.

(49) Which boy came?

$[_{\text{CP}}\ ^{\text{BD}}(\text{wh-boy}_@)\ [_\gamma\ i\ [_{\text{IP}}\ \lambda w.\ t_i\ \text{came}_w\ ]]]$

a. $[\![\gamma]\!] = \lambda x_e \lambda w.\mathsf{came}_w(x)$

b. $[\![\text{CP}]\!] = \lambda x_e : \mathsf{boy}_@(x).\lambda w[\mathsf{came}_w(x)]$

The following illustrates the derivations for individual and functional readings of *wh*-questions with a quantifier. An individual reading arises if the *wh*-phrase binds an individual trace, as in (50a); a functional reading arises if the *wh*-phrase binds an (intensional) functional trace, as in (50b).

[18] Crucially, $\textsc{BeDom}(\pi)$ is type-flexible — it can combine with any function of a $\langle \sigma, \ldots\rangle$ type where σ is the type of an element in $\textsc{Be}(\pi)$. Type-flexibility makes it possible to compose a question regardless of whether the *wh*-phrase binds an individual or functional variable, and regardless of how many *wh*-phrases there are in this question. This assumption overcomes difficulties with traditional categorial approaches in composing multi-*wh* questions with single-pair readings.

(50) Which movie did every boy watch?

a. Individual reading:
'Which movie y is s.t. every boy watched y?'
$[_{\text{CP}}$ $^{\text{BD}}$(wh-movie$_@$) $[_\gamma$ i $[_{\text{IP}}$ λw. every-boy$_w$ watched$_w$ t_i]]]

i. $[\![\gamma]\!] = \lambda y_e \lambda w. \forall x[\mathsf{boy}_w(x) \rightarrow \mathsf{wat}_w(x, y)]$

ii. $[\![\text{CP}]\!] = \lambda y_e : \mathsf{mov}_@(y). \lambda w[\forall x[\mathsf{boy}_w(x) \rightarrow \mathsf{wat}_w(x, y)]]$

b. (Intensional) functional reading:
'Which Skolem function f to atomic movies is s.t. for every boy x, x watched $f(x)$?'
$[_{\text{CP}}$ $^{\text{BD}}$(wh-λw.movie$_w$) $[_\gamma$ i $[_{\text{IP}}$ λw. every-boy$_w$ j $[_{\text{VP}}$ t_j watched$_w$ t_i^j]]]]

i. $[\![\gamma]\!] = \lambda f_{\langle s,ee \rangle} \lambda w. \forall x[\mathsf{boy}_w(x) \rightarrow \mathsf{wat}_w(x, f_w(x))]$

ii. $[\![\text{CP}]\!] = \lambda f_{\langle s,ee \rangle} : \forall w'[\text{Ran}(f_{w'}) = \mathsf{mov}_{w'}]. \lambda w[\forall x[\mathsf{boy}_w(x) \rightarrow \mathsf{wat}_w(x, f_w(x))]]$

6. Proposal

In line with functionality approaches, I analyze pair-list readings of multi-*wh* questions and QiQ-readings of questions with a quantifier as extensional functional readings. For both types of questions, I assume that the composition involves a quantificational binding-in operation applied into a 'dependency sentence'. A dependency sentence is an open sentence with the form $\ulcorner x$ P $f(x) \urcorner$ which expresses a functional dependency relation between the two arguments of the two-place predicate P. In particular, in the composition of a pair-list multi-*wh* question, the *wh*-subject existentially quantifies into an identity condition (*à la* Karttunen Semantics). In contrast, in the composition of a QiQ-question, the quantificational subject quantifies into a predication condition (*à la* Fox 2012b). The LF schema is as follows, repeated from (6):

(51) LF schema for composing complex questions

a. Which boy watched which movie? (Pair-list reading)
... [which-movie$_j$ ... which-boy$_i$ $[_{\text{IDENT}}$ ... [t_i watched t_j^i]]]

b. Which movie did Det-boy(s) watch? (QiQ-reading)
... [which-movie$_j$ ... Det-boy(s)$_i$ $[_{\text{PRED}}$ ... [t_i watched t_j^i]]]

The subtle distinctions between IDENT and PRED give rise to a contrast in domain exhaustivity between multi-*wh* questions and ∀-questions. The LF schema for QiQ-questions automatically explains why ∀-questions and ∃-questions have pair-list readings and choice readings, respectively, and why NO-questions do not have QiQ-readings. What's more, with known contrasts among non-interrogative quantifiers in lexical distributivity and scoping, this analysis also explains why counting quantifiers do not participate in QiQ-readings.

In what follows, I will first provide the root denotation of each type of complex questions (Sect. 6.1) and revisit the definition of answerhood (Sect. 6.2). Next, I will show how to derive each of these root denotations compositionally (Sect. 6.3 and 6.4). Finally, section 6.5 will summarize the proposal.

6.1. Question denotations

I assume that pair-list readings and QiQ-readings of complex *wh*-questions are extensional functional readings. When having a pair-list/QiQ reading, a question denotes a topical property that maps an $\langle e, e \rangle$-type function f to the conjunction of a set of propositions that describes the graph of f.

Illustrations of those topical properties are given in (52) and (53) in tandem. The (a)-denotations are represented in a way isomorphic to the structures of composition (for details of composition, see Sect. 6.3 and 6.4). The (b)-denotations are semantically equivalent to their (a)-counterparts but are represented in a way more convenient for comparison.

(52) $[\![\textit{Which boy watched which movie?}]\!]_{\text{pair-list}}$
$\Leftrightarrow \lambda \boldsymbol{f}_{\langle e,e\rangle} : \text{Ran}(\boldsymbol{f}) = \mathsf{mov}_@ . \bar{\bigcap}\{p \mid \exists\text{-}\mathsf{boy}_@(\lambda x.p = \lambda w.\mathsf{wat}_w(x, \boldsymbol{f}(x)))\}$ (a)
$\Leftrightarrow \lambda \boldsymbol{f}_{\langle e,e\rangle} : \text{Ran}(\boldsymbol{f}) = \mathsf{mov}_@ . \bar{\bigcap}\{\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}(x)) \mid \mathsf{boy}_@(x)\}$ (b)

(53) $[\![\textit{Which movie did Det-boy(s) watch?}]\!]_{\text{QiQ}}$
$\Leftrightarrow \lambda \boldsymbol{f}_{\langle e,e\rangle} : \text{Ran}(\boldsymbol{f}) = \mathsf{mov}_@ . \bar{\bigcap} f^{\text{MIN}}_{\text{CH}}(\{\boldsymbol{K} \mid \mathsf{Det\text{-}boy}_@(\lambda x.\boldsymbol{K}(\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}(x))))\})$ (a)
$\Leftrightarrow \lambda \boldsymbol{f}_{\langle e,e\rangle} : \text{Ran}(\boldsymbol{f}) = \mathsf{mov}_@ \wedge \mathsf{Det\text{-}boy}_@(\text{Dom}(\boldsymbol{f}))\ .$
$\bar{\bigcap} f^{\text{MIN}}_{\text{CH}}(\{\boldsymbol{K} \mid \mathsf{Det\text{-}boy}_@(\lambda x.\boldsymbol{K}(\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}(x))))\})$ (b)

The denotations above introduce two new operators, namely, $\bar{\bigcap}$ and $f^{\text{MIN}}_{\text{CH}}$. The $\bar{\bigcap}$-operator is like the regular intersection operator except that it ignores the undefined elements.[19]

(54) $\bar{\bigcap} := \lambda A_{\langle \tau t,t\rangle} . \bigcap\{a \mid a \in A \wedge a \text{ is defined}\}$

The $f^{\text{MIN}}_{\text{CH}}$-operator combines a weak minimization-operator $\textsc{min}_W$ with a choice-function variable f_{CH}, which gets existentially bound at a global site.[20] The $\textsc{min}_W$-operator is weaker than Pafel-Fox's $\textsc{min}_S$-operator: for any set α, a member x of α is a minimal member of α as long as no member of α is a proper subset/subpart of x, not necessarily that x is a subset/subpart of every member of α.[21] The choice between $\textsc{min}_S$ and $f^{\text{MIN}}_{\text{CH}}$ makes no difference in $\forall$-questions, but only the latter works for $\exists$-questions (Sect. 6.4.2).

(55) $f^{\text{MIN}}_{\text{CH}} := \lambda \alpha_{\langle \sigma,t\rangle} . f_{\text{CH}}(\textsc{min}_W(\alpha))$

(56) a. $\textsc{min}_W := \lambda \alpha_{\langle \sigma,t\rangle} . \{x_\sigma \mid x \in \alpha \wedge \neg\exists y \in \alpha[y < x]\}$
b. $\textsc{min}_S := \lambda \alpha_{\langle \sigma,t\rangle} . \iota x_\sigma[x \in \alpha \wedge \forall y \in \alpha[y \geq x]]$ (generalized from (42))
['$<$' stands for the proper subset relation if α is a set of sets and the proper subpart relation if α is a set of non-sets; '$\geq$' is analogous.]

Notice a contrast between (52b) and (53b): (52b) only has a restriction only the range of the input functions, while (53b) also has restriction on the domain of these functions. More concretely, in (52), the topical property of the multi-*wh* question takes any function that maps entities to atomic

[19]In specific, if a set A contains an undefined member, applying the $\bar{\bigcap}$-operator to A does not pass up the undefinedness of this member to the returned intersection, as in (ib), contrasted by (ia). However, as in (ic), if a definedness condition is applied to the set A, not to a member of A, $\bar{\bigcap} A$ inherits this condition.

(i) Assume that $\boldsymbol{f}$ is defined for two atomic boys a and b, but not to boy c. Then:
a. $\bigcap\{\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}(x)) \mid \mathsf{boy}_@(x)\}$ is undefined;
b. $\bar{\bigcap}\{\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}(x)) \mid \mathsf{boy}_@(x)\} = \lambda w.\mathsf{wat}_w(a, \boldsymbol{f}(a)) \wedge \mathsf{wat}_w(b, \boldsymbol{f}(b))$;
c. $\bar{\bigcap}[\forall x \in \{a,b,c\}[x \in \text{Dom}(\boldsymbol{f})].\{\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}(x)) \mid \mathsf{boy}_@(x)\}]$ is undefined.

[20]For readers who are familiar with Boolean Semantics, the $f^{\text{MIN}}_{\text{CH}}$-operator is roughly the same as the collectivity raising operator in Winter 2001.

[21]The following illustrates the contrast between the $f^{\text{MIN}}_{\text{CH}}$-operator and the $\textsc{min}_S$-operator:

(i) Let a and b be two distinct entities, $A = \{\varnothing, \{a\}, \{b\}\}$, and $B = \{\{a\}, \{b\}\}$. Then:
a. $\textsc{min}_S(A) = f^{\text{MIN}}_{\text{CH}}(A) = \varnothing$;
b. $\textsc{min}_S(B)$ is undefined; while $f^{\text{MIN}}_{\text{CH}}(B)$ has two possible values: $\{a\}$ and $\{b\}$.

movies as its input and the graph description of this function as its output. In contrast, in (53) the topical property of the QiQ-question is more restrictive — it is only defined for functions that map Det-boy(s) to atomic movies. The additional domain restriction in (53b) (i.e., $\textsf{Det-boy}_{@}(\text{Dom}(\boldsymbol{f}))$) arises as a definedness condition of the value description in (53a): the quantificational predication condition $\textsf{Det-boy}_{@}(\lambda x.\boldsymbol{K}(\lambda w.\textsf{wat}_w(x, \boldsymbol{f}(x))))$, read as 'for Det-boy(s) x, the proposition $\ulcorner x$ *watched* $\boldsymbol{f}(x)\urcorner$ is a member of $\boldsymbol{K}$', is defined only if the function $\boldsymbol{f}$ is defined for Det-boy(s).

For a concrete illustration of (53), consider the QɪQ-denotation of a ∀-question:

(57) ⟦*Which movie did every/each boy watch?*⟧

$$\Leftrightarrow \lambda \boldsymbol{f}_{\langle e,e\rangle}: \underbrace{\text{Ran}(\boldsymbol{f}) = \textsf{mov}_{@}}_{\text{from } wh\text{-obj}} . \underbrace{\bar{\bigcap} f_{\text{CH}}^{\text{MIN}}(\{\boldsymbol{K} \mid \forall\text{-}\textsf{boy}_{@}(\lambda x.\boldsymbol{K}(\lambda w.\textsf{wat}_w(x, \boldsymbol{f}(x))))\})}_{\text{(i) from question nucleus}} \quad \text{(a)}$$

$$\Leftrightarrow \lambda \boldsymbol{f}_{\langle e,e\rangle}: \underbrace{\text{Ran}(\boldsymbol{f}) = \textsf{mov}_{@}}_{\text{from } wh\text{-obj}} . \underbrace{\bar{\bigcap} [\forall\text{-}\textsf{boy}_{@}(\text{Dom}(\boldsymbol{f})).\{\lambda w.\textsf{wat}_w(x, \boldsymbol{f}(x)) \mid \textsf{boy}_{@}(x)\}]}_{=\text{(i)}} \quad \text{(a')}$$

$$\Leftrightarrow \lambda \boldsymbol{f}_{\langle e,e\rangle}: \underbrace{\text{Ran}(\boldsymbol{f}) = \textsf{mov}_{@}}_{\text{from } wh\text{-obj}} \wedge \underbrace{\forall\text{-}\textsf{boy}_{@}(\text{Dom}(\boldsymbol{f}))}_{\text{definedness cond of (i)}} .\bar{\bigcap}\{\lambda w.\textsf{wat}_w(x, \boldsymbol{f}(x)) \mid \textsf{boy}_{@}(x)\} \quad \text{(b)}$$

In (57a), the input $\boldsymbol{f}$ can be any $\langle e, e\rangle$-type function from entities to atomic movies. For each such $\boldsymbol{f}$, the output is the conjunction of the set that consists of exactly all the propositions of the form $\ulcorner$*boy-x watched* $\boldsymbol{f}(x)\urcorner$. Crucially, as represented in (57a'), the output inference carries a definedness condition: the universal predication condition w.r.t. the open sentence 'boy-x watched $\boldsymbol{f}(x)$' is defined only if $\boldsymbol{f}$ is defined for every boy. This condition projects to the entire topical property as in (57b), yielding domain exhaustivity. In short, the ∀-question (57) is semantically equivalent to the multi-*wh* counterpart (52), except (57) presupposes domain exhaustivity.

At this point, it is clear why I pursue a functionality approach instead of a family-of-questions approach: assuming a *wh*-dependency, I attribute the domain exhaustivity effect to a definedness condition arising from an operation applied within the question nucleus (viz., the universal quantificational binding-in operation w.r.t. a dependency proposition). In this approach, the contrast in domain exhaustivity between multi-*wh* questions and ∀-questions can be explained in terms of the differences of their nucleus. In family-of-questions approaches, however, domain exhaustivity is attributed to an operation applied outside the question nucleus (e.g., the point-wise answerhood-operator of Fox 2012a,b); such accounts cannot capture the semantic contrast between ∀-questions and multi-*wh* questions in terms of their structural differences.

6.2. Redefining answerhood

As pointed out by Floris Roelofsen (pers. comm.), the answerhood-operator assumed in (46a) overgenerates possible short answers for pair-list questions. For example, the topical property of *Which boy watched which movie* is defined for any $\langle e, e\rangle$-type functions that map entities to atomic movies, not just those consisting of only boy-movie pairs. To solve this problem, I define the answerhood in (58a) for possible short answers. The added constraint, namely, that every subset of α yields a propositional answer that is possibly true, rules out functions that allow inputs that are non-boys.

(58) Answerhood for possible answers

a. For short answers

$$\mathbb{A}^S(\boldsymbol{P}) = \begin{cases} \text{Dom}(\boldsymbol{P}) & \text{if } \boldsymbol{P} \in D_{\langle e,\tau\rangle} \\ \{\alpha \mid \alpha \in \text{Dom}(\boldsymbol{P}) \wedge \forall \beta \subseteq \alpha[\exists w \in \boldsymbol{P}(\beta)]\} & \text{if } \boldsymbol{P} \in D_{\langle\langle e,e\rangle,\tau\rangle} \end{cases}$$

b. For propositional answers
$\mathbb{A}(\boldsymbol{P}) = \{\boldsymbol{P}(\alpha) \mid \alpha \in \mathbb{A}^S(\boldsymbol{P})\}$

Next, let's consider the answerhood for complete true answers. For pair-list questions like (59) with a number-neutral *wh*-subject and a semi-distributive predicate, the same pair-list propositional answer can be derived based on distinct possible short answers, as those listed in (59a-c).

(59) Which boy or boys watched which movie?
(Context: The boys b_1,b_2 both watched the movie m_1, and the boy b_3 watched movie m_2.)
a. $[b_1 \to m_1, b_2 \to m_1, b_3 \to m_2]$, $[b_1 \oplus b_2 \to m_1, b_3 \to m_2]$
b. $[b_1 \to m_1, b_1 \oplus b_2 \to m_1, b_3 \to m_2]$, $[b_2 \to m_1, b_1 \oplus b_2 \to m_1, b_3 \to m_2]$
c. $[b_1 \to m_1, b_2 \to m_1, b_1 \oplus b_2 \to m_1, b_3 \to m_2]$

Given this multi-to-one mapping from short answers to propositional answers, I redefine the answerhood for complete true answers as follows — the complete true short answer is the maximal short answer that yields the strongest true propositional answer:[22]

(60) Answerhood for complete true answers (final)
a. For short answers
$\text{Ans}^S(w)(\boldsymbol{P}) = \exists\alpha \in \mathbb{X}(w)(\boldsymbol{P}).\text{Max}[\mathbb{X}(w)(\boldsymbol{P})]$, where
$\mathbb{X}(w)(\boldsymbol{P}) = \{\alpha \mid \alpha \in \mathbb{A}(\boldsymbol{P}) \wedge w \in \boldsymbol{P}(\alpha) \wedge \forall\beta \in \mathbb{A}(\boldsymbol{P})[w \in \boldsymbol{P}(\beta) \to \boldsymbol{P}(\alpha) \subseteq \boldsymbol{P}(\beta)]\}$
b. For propositional answers
$\text{Ans}(w)(\boldsymbol{P}) = \boldsymbol{P}(\text{Ans}^S(w)(\boldsymbol{P}))$

6.3. Composing pair-list multi-*wh* questions

Figure 1 illustrates the composition of a pair-list multi-*wh* question. As marked in the tree diagram, this composition precedes in four steps.

(i) *Derive a functional dependency.* The argument index carried by the complex functional trace of the *wh*-object is co-indexed with the subject trace, yielding a dependency proposition which expresses a dependency relation between the two arguments of *watched*.

(ii) *Quantificational binding into an identity condition.* Employing an identity (Id-)operator yields an identity relation between a covert propositional variable p and the dependency proposition denoted by IP. At Node ①, the *wh*-subject, interpreted as an $\exists$-quantifier, binds the argument index inside the IP across the Id-operator, yielding an existential identity condition w.r.t. the dependency proposition.

(iii) *Derive a function graph description.* Abstracting p returns the set of propositions with the form $\ulcorner$*boy-x watched* $\boldsymbol{f}_i(x)\urcorner$. Conjoining this set yields the graph description of the function $\boldsymbol{f}_i$. Here the $\bar{\bigcap}$-closure can be perceived as a 'function graph creator' in the sense of Dayal 2017. Note that, as defined in (54), the $\bar{\bigcap}$-operator ignores the undefined members in the set it applies to; therefore, the yielded conjunction is defined even if $\boldsymbol{f}_i$ is undefined for some of the boys. In other words, the application of the $\bar{\bigcap}$-operator does not force domain exhaustivity.

(iv) *Create a topical property.* Abstracting the functional index yields a property (of type $\langle ee, st\rangle$) that maps each $\langle e, e\rangle$-type function to the graph description of this function. Further, the fronted

[22] For the purpose of defining answerhood, it doesn't make a big difference whether we assume that the complete true short answer is the maximal element in $\mathbb{X}(w)(\boldsymbol{P})$ as in (59c), a minimal element as those in (59a), or any other element.

wh-object '$^{\text{BD}}$(wh-movie$_@$)' restricts the domain of this property — it requests the range of each input function to be a set of atomic movies.

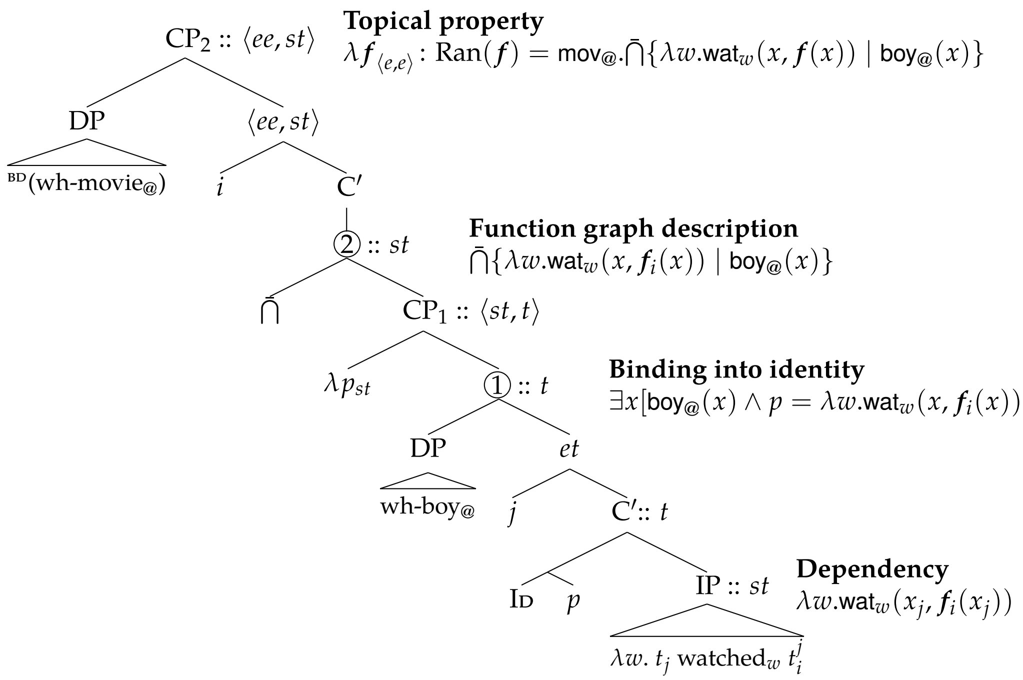


Figure 1: Composition of the pair-list multi-*wh* question *Which boy watched which movie?*

(61) Steps (i) & (ii): Bind into an identity condition w.r.t. a dependency proposition

a. $[\![\text{IP}]\!] = \lambda w.\textsf{wat}_w(x_j, \boldsymbol{f}_i(x_j))$

b. $[\![\text{I}_{\text{D}}]\!] = \lambda \alpha_\tau \lambda \beta_\tau . \alpha = \beta$

c. $[\![\text{C}']\!] = [\![\text{I}_{\text{D}}]\!](p)([\![\text{IP}]\!])$
$\quad = [p = \lambda w.\textsf{wat}_w(x_j, \boldsymbol{f}_i(x_j))]$

d. $[\![\textit{wh-boy}_@]\!] = \lambda P_{\langle e,t\rangle} . \exists x[\textsf{boy}_@(x) \wedge P(x)]$

e. $[\![①]\!] = [\![\textit{wh-boy}_@]\!]([\![\text{C}']\!])$
$\quad = \exists x[\textsf{boy}_@(x) \wedge p = \lambda w.\textsf{wat}_w(x, \boldsymbol{f}_i(x))]$

(62) Step (iii): Create a function graph description

a. $[\![\text{CP}_1]\!] = \lambda p.\exists x[\textsf{boy}_@(x) \wedge p = \lambda w.\textsf{wat}_w(x, \boldsymbol{f}_i(x))]$
$\quad = \{\lambda w.\textsf{wat}_w(x, \boldsymbol{f}_i(x)) \mid \textsf{boy}_@(x)\}$

b. $[\![②]\!] = \bar{\bigcap}\{\lambda w.\textsf{wat}_w(x, \boldsymbol{f}_i(x)) \mid \textsf{boy}_@(x)\}$

(63) Step (iv): Create a topical property
$[\![\text{CP}_2]\!] = \lambda \boldsymbol{f}_{\langle e,e\rangle} : \text{Ran}(\boldsymbol{f}) = \textsf{mov}_@ . \bar{\bigcap}\{\lambda w.\textsf{wat}_w(x, \boldsymbol{f}(x)) \mid \textsf{boy}_@(x)\}$

Eagle-eyed readers might notice that here the *wh*-object is moved over the fronted *wh*-subject, which violates the generalization of 'tucking-in' (Richards 1997). Although violations of tucking-in are sometimes permitted for D-linked *wh*-phrases, it is certainly problematic to say that pair-list readings are only available in constructions that violate tucking-in. However, this problem does not come from the specific assumptions on composing pair-list multi-*wh* questions; it is a consequence of requiring covert/overt *wh*-fronting in question composition. This problem can be avoided if we assume a framework of composition that allows *wh*-insitu. For example, in Variable-free Semantics (Jacobson 2014), abstractions can be passed up by type-shifting operations. Integrating my core proposal on composing pair-list questions into such frameworks allows to create the wanted topical property without fronting the object *wh*-phrase.[23]

6.4. Composing QɪQ-questions

QɪQ-questions with the form *Which movie did Det-boy(s) watch?* are composed uniformly with the LF schema in Figure 2. The steps of composition are in parallel with those for the pair-list multi-*wh* question *Which boy watched which movie?*. The following subsections will explain how this composition schema works for each type of QiQ-questions.

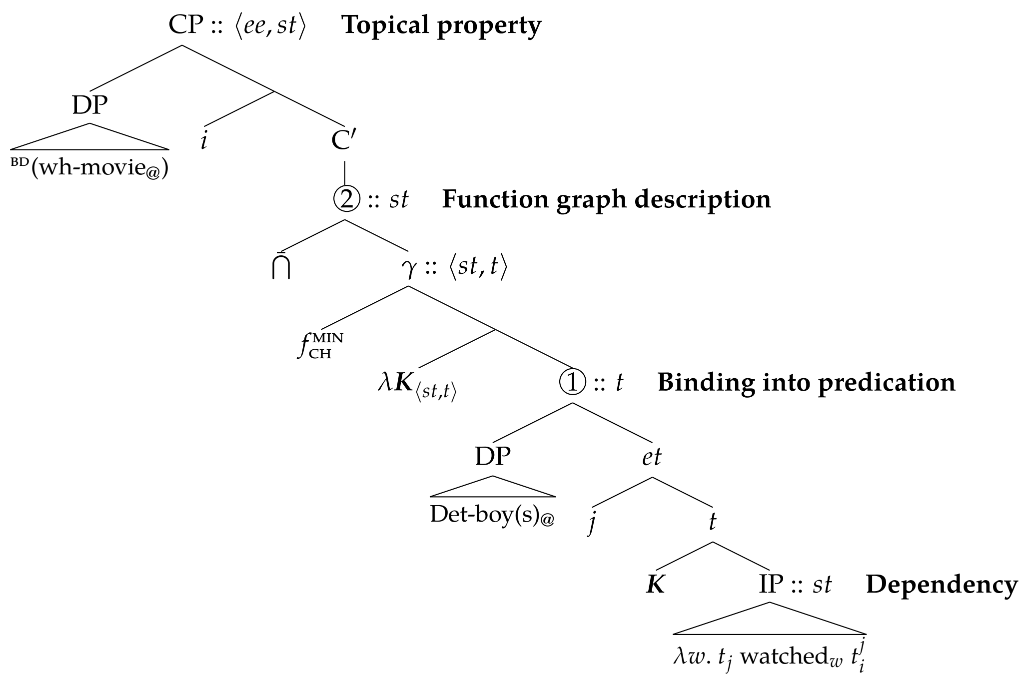


Figure 2: Composition of the QɪQ-question *Which movie did Det-boy(s) watch?*

In denotation (64b), the condition on the range of f (i.e., that f maps to atomic movies) comes from the fronted *wh*-object. All the other conditions, including the condition on the domain of f (i.e., that f is defined for Det-boy(s)) and the output proposition which describes the graph of the input function, come from the question nucleus (viz., Node ②).

(64) $[\![\textit{Which movie did Det-boy(s) watch?}]\!]_{\text{QiQ}}$ (repeated from (53))

$$\Leftrightarrow \lambda f_{\langle e,e\rangle} : \underbrace{\text{Ran}(f) = \textsf{mov}_@}_{\text{from } wh\text{-object}} . \underbrace{\bar{\bigcap} f^{\text{MIN}}_{\text{CH}}(\{K \mid \textsf{Det-boy}_@(\lambda x.K(\lambda w.\textsf{wat}_w(x, f(x))))\})}_{\text{(i) from nucleus}} \quad \text{(a)}$$

$$\Leftrightarrow \lambda f_{\langle e,e\rangle} : \text{Ran}(f) = \textsf{mov}_@ \wedge \overbrace{\textsf{Det-boy}_@(\text{Dom}(f))}^{\text{definedness cond of (i)}} . \bar{\bigcap} f^{\text{MIN}}_{\text{CH}}(\{K \mid \textsf{Det-boy}_@(\lambda x.K(\lambda w.\textsf{wat}_w(x, f(x))))\}) \quad \text{(b)}$$

[23]For an attempt of composing complex questions with a variable-free grammar, see Xiang 2019b.

Recall that *wh*-questions with a quantificational subject admit both functional readings and QiQ-readings. Here let's compare the derivations of these two readings. In both readings, the question involves a *wh*-dependency, derived by letting the subject bind into the complex functional trace of the *wh*-object. Despite this similarity, the composition of a pair-list reading makes use of another two operations, i.e., quantifying-into-predication and minimization. These operations are similar to what Fox (2012b) assumes for composing ∀-questions (see (41)), but they depart from Fox's implementation in two respects, yielding desirable consequences in accounting for the domain exhaustivity effects in ∀-questions and the choice readings of ∃-questions. First, in the presented analysis, the predication operation is applied to a dependency proposition (not to a question). Binding into the dependency proposition is crucial for the derivation of domain exhaustivity in ∀-questions (Sect. 6.4.1). Second, the $f_{\text{CH}}^{\text{MIN}}$-operator is weaker than the MIN_S-operator that Fox adopts from Pafel (1999): $f_{\text{CH}}^{\text{MIN}}$ doesn't require the existence of a unique minimal member (Sect. 6.1). Replacing MIN_S with $f_{\text{CH}}^{\text{MIN}}$ makes the analysis feasible in tackling choice ∃-questions (Sect. 6.4.2).

6.4.1. Composing ∀-questions

This section presents the composition of pair-list ∀-questions. The primary goals are to derive the pair-list readings and to account for the domain exhaustivity effects. The LF is given in Figure 3. In parallel to the composition of pair-list multi-*wh* questions (Sect. 6.3), I divide the composition into four steps:

(i) *Derive a functional dependency.* The IP denotes a dependency proposition, composed in the same way as the IP in the pair-list multi-*wh* question.

(ii) *Quantificational binding into a predication condition.* A null predication-operator $\boldsymbol{K}$ is applied to IP, yielding a predication condition that the dependency proposition denoted by the IP is a member of $\boldsymbol{K}$. Next, the subject *every/each-boy* quantifies into this predication condition and binds the argument index j, yielding a universal predication condition as stated in (65b).

(iii) *Create a function graph description.* Abstracting the predicative variable $\boldsymbol{K}$ returns the set of $\boldsymbol{K}$ sets that satisfy the universal predication condition, as in (66a). These are the sets that contain all the propositions with the form ⌜*boy-x watched* $\boldsymbol{f}_i(x)$⌝. Next, applying the minimizer $f_{\text{CH}}^{\text{MIN}}$ returns a minimal $\boldsymbol{K}$ set that satisfies the universal quantification predication condition, as in (66b). Here there is only one such minimal $\boldsymbol{K}$ set, i.e., $\{\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}_i(x)) \mid \mathsf{boy}_@(x)\}$. Finally, applying the intersection-operator $\bar{\bigcap}$ to this set returns the graph description of $\boldsymbol{f}_i$, as in (66c).

(iv) *Create a topical property.* The fronted '$^{\text{BD}}$(wh-movie$_@$)' binds the functional index j and restricts the range of any input $\boldsymbol{f}$ to the set of atomic movies. The possible inputs of this topical property are therefore functions that map each boy to an atomic movie, and the outputs are conjunctive propositions that describe the graph of each input function.

Step (ii) of this composition — quantificational binding into predication — is the heart of the analysis. First, it carries forward the advantage of Fox's analysis that the quantificational subject standardly combines with a one-place predicate of type $\langle e, t\rangle$. In contrast to earlier accounts (e.g., Groenendijk and Stokhof 1984; Chierchia 1993; Dayal 1996, 2017), there is no need to assume a type-shifting operation or make use of a minimal witness set. What's more, since here the ∀-subject binds into the functional *wh*-trace, it yields a definedness condition that the function $\boldsymbol{f}_i$ is defined for every boy. This definedness condition projects to CP, yielding domain exhaustivity.[24]

[24]Note that the $\bar{\bigcap}$-operator does not remove this definedness condition: as specified in the denotation of the γ-node, this definedness condition is applied to the minimal $\boldsymbol{K}$ set as a whole, not to the members of $\boldsymbol{K}$. See (i) for relevant illustrations.

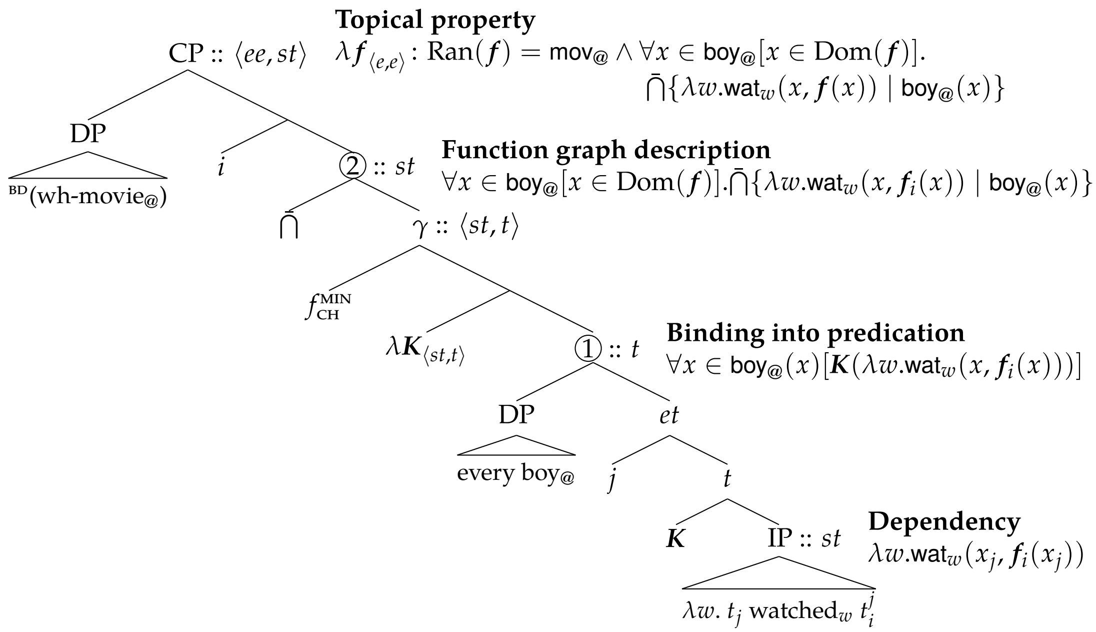


Figure 3: Composition of the ∀-question *Which movie did every boy watch?*

(65) Steps (i) & (ii): Bind into the predication condition w.r.t. a dependency proposition

a. $[\![\text{IP}]\!] = \lambda w.\mathsf{wat}_w(x_j, \boldsymbol{f}_i(x_j))$ (equivalent to (61))

b. $[\![①]\!] = [\![\textit{every boy}_@]\!](\lambda x.\boldsymbol{K}(\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}_i(x))))$
$= \forall x \in \mathsf{boy}_@[\boldsymbol{K}(\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}_i(x)))]$ (defined only if $\forall x \in \mathsf{boy}_@[x \in \mathrm{Dom}(\boldsymbol{f}_i)]$)
(For every boy x, the proposition 'x *watched* $\boldsymbol{f}_i(x)$' is a member of $\boldsymbol{K}$.)

(66) Step (iii): Create a function graph description

a. $[\![\lambda \boldsymbol{K}.①]\!] = \lambda \boldsymbol{K}.\forall x \in \mathsf{boy}_@[\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}_i(x)) \in \boldsymbol{K}]$
$= \lambda \boldsymbol{K}\colon \forall x \in \mathsf{boy}_@[x \in \mathrm{Dom}(\boldsymbol{f}_i)].\{\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}_i(x)) \mid \mathsf{boy}_@(x)\} \subseteq \boldsymbol{K}$

b. $[\![\gamma]\!] = f^{\text{MIN}}_{\text{CH}}([\![\lambda \boldsymbol{K}.①]\!])$
$= \forall x \in \mathsf{boy}_@[x \in \mathrm{Dom}(\boldsymbol{f}_i)].\{\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}_i(x)) \mid \mathsf{boy}_@(x)\}$

c. $[\![②]\!] = \forall x \in \mathsf{boy}_@[x \in \mathrm{Dom}(\boldsymbol{f}_i)].\bar{\bigcap}\{\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}_i(x)) \mid \mathsf{boy}_@(x)\}$

(67) Step (iv): Create a topical property
$[\![\text{CP}]\!] = \lambda \boldsymbol{f}_{\langle e,e\rangle}\colon \mathrm{Ran}(\boldsymbol{f}) = \mathsf{mov}_@ \wedge \forall x \in \mathsf{boy}_@[x \in \mathrm{Dom}(\boldsymbol{f})].\bar{\bigcap}\{\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}(x)) \mid \mathsf{boy}_@(x)\}$

The explanation of domain exhaustivity crucially relies on the presence of a ∀-quantifier: domain exhaustivity comes from the binding relation between a ∀-quantifier and the argument index of the functional *wh*-trace. Nicely, this analysis does not over-predict domain exhaustivity for a pair-list multi-*wh* question: in a multi-*wh* question, the argument variable of the functional trace of the *wh*-object is 'existentially' bound by the *wh*-subject. For comparison, the family-of-questions approach of Fox (2012a,b) attributes domain exhaustivity to an operation outside the question nucleus, namely, the point-wise answerhood-operator. Since the selection of answerhood is independent from the root structure/meaning of a question, the family-of-questions approach cannot explain the contrast in domain exhaustivity between ∀-questions and multi-*wh* questions.

In the rest sections, I will describe the characteristics of the QiQ-reading of each type of questions in terms of three parameters, including:

[±D-EXH]:	whether the reading is subject to domain exhaustivity;
[±PL]:	whether the reading is a pair-list reading;
[±CH]:	whether the reading has a 'choice' flavor.

The QIQ-reading of a ∀-question is [+D-EXH,+PL,−CH]. It presupposes domain exhaustivity because the universal predication condition (from Node ①) is defined only if the input function $\boldsymbol{f}$ is defined for 'every' boy. It expects a pair-list answer because the yielded eligible minimal proposition set $\boldsymbol{K}$ (from Node γ) that satisfies the universal predication condition is a 'non-singleton' set ranging over multiple boys. It does not have a choice flavor because there is 'only one' such eligible minimal $\boldsymbol{K}$ set.

6.4.2. Composing ∃-questions

Choice readings of ∃-questions are derived in the same way as pair-list readings of ∀-questions. At Node ①, the ∃-subject binds into the complex functional trace of the *wh*-object across the null predicate $\boldsymbol{K}$, yielding an existential predication condition w.r.t. a dependency proposition. At Node γ, applying the $f_{\text{CH}}^{\text{MIN}}$-operator returns one of the minimal $\boldsymbol{K}$ sets that satisfy the existential predication condition. Crucially, unlike the case of the ∀-question, here there are 'multiple' minimal $\boldsymbol{K}$ sets that satisfy the quantificational predication condition, each of which is a singleton set consisting of a proposition with the form ⌜*boy-x watched* $\boldsymbol{f}_i(x)$⌝. Each such minimal $\boldsymbol{K}$ set supplies a possible topical property, which therefore gives rise to a choice flavor.

(68) Which movie did one of the boys watch?

[$_{\text{CP}}$ $^{\text{BD}}$(wh-movie$_@$) i [$_②$ $\bar{\bigcap}$ [$_\gamma$ $f_{\text{CH}}^{\text{MIN}}$ $\lambda \boldsymbol{K}$ [$_①$ one-of-the-boys$_@$ j [$\boldsymbol{K}$ ($\lambda w.x_j$-watched$_w$-$\boldsymbol{f}_i(x_j)$)]]]]]

a. $[\![①]\!] = \exists x \in \mathsf{boy}_@[\boldsymbol{K}(\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}_i(x)))]$

b. $[\![\gamma]\!] = f_{\text{CH}}^{\text{MIN}}([\![\lambda \boldsymbol{K}.①]\!])$
$= f_{\text{CH}}(\{\{\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}_i(x))\} \mid x \in \mathsf{boy}_@\})$
$= \{\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}_i(x))\}$, where x is the chosen boy

c. $[\![②]\!] = \lambda w.\mathsf{wat}_w(x, \boldsymbol{f}_i(x))$, where x is the chosen boy

d. $[\![\text{CP}]\!] = \lambda \boldsymbol{f}_{\langle e,e \rangle} : \text{Ran}(\boldsymbol{f}) = \mathsf{mov}_@ .\lambda w[\mathsf{wat}_w(x, \boldsymbol{f}(x))]$, where x is the chosen boy

Note that this approach does not assume a choice-function analysis of indefinites. In (68), the $f_{\text{CH}}^{\text{MIN}}$-operator, which contains a choice-function variable f_{CH}, applies to a family of singleton sets of propositions, not to a set of boys. The subject *one of the boys* is treated standardly as an existential generalized quantifier. Therefore, more precisely, Node γ should be read as 'the chosen singleton set of propositions with the form ⌜{*boy-x watched* $\boldsymbol{f}_i(x)$}⌝'. I further assume that the choice-function variable f_{CH} is existentially bound at a global site. The full paraphrase of the LF (68) is as follows:

(69) 'For some choice function f_{CH}, what is the $\langle e, e \rangle$-function $\boldsymbol{f}$ to atomic movies s.t. the conjunction of the singleton set {*boy-x watched* $\boldsymbol{f}(x)$} chosen by f_{CH} is true?'

The QIQ-reading of an ∃-question yielded with the above analysis is [−D-EXH,−PL,+CH]. This reading is not subject to domain exhaustivity because the existential predication condition (68a) only requires the input function $\boldsymbol{f}$ to be defined for 'at least one' of the boys.[25] Possible answers to this question are single-pairs, not pair-lists, because the minimal $\boldsymbol{K}$ sets satisfying the existential

[25] In (68c,d), there is no need to write out the domain condition that $\boldsymbol{f}$ is defined for at least one boy, because this condition is entailed by the definedness condition of the output proposition: for any x, the proposition $\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}(x))$ is defined only if $\boldsymbol{f}$ is defined for x.

predication condition are all 'singleton' sets, as seen in (68b). The yielded QɪQ-reading has a choice flavor, because there can be 'multiple' minimal $\boldsymbol{K}$ sets that satisfy the existential predication condition.

The above discussion is for the ∃1-quantifier *one of the boys*. The rest of this section considers other indefinites with the form 'Num-(of-the-)NP'. Recall that pair-list readings are not available in matrix ∃-question; for example, the ∃2-question (70c) cannot be interpreted with distributivity between quantification and uniqueness.

(70) I know that every student voted for a different candidate. Which candidate did ...

a. ... every/each student vote for? $(\textit{every/each} \gg \iota)$

b. ... one of the students vote for? $(\exists 1 \gg \iota)$

c. # ... two of the students vote for? $(\exists 2 \gg$ ~~EACH~~ $\gg \iota)$

To avoid over-generating pair-list readings, pioneering works derive pair-list readings in ways that would crash in questions with a non-universal quantifier. In Dayal's analysis, the derivation of pair-list readings crashes because ∃-quantifiers have multiple minimal witness sets. In Fox 's analysis, the derivation crashes because we cannot find the unique minimal set among the sets of sub-questions that satisfy an existential predication condition. Obviously, this strategy comes with an expense of failing to account for the choice readings of ∃-questions.

I assume that the determiner of the numeral-modified indefinite *two of the boys* is not ∃2 but rather the simple ∃; in other words, the cardinal numeral *two* is part of the restrictor of the determiner. With this assumption, the quantifier *two of the boys* ranges over the set of entities that are pluralities of two boys; in other words, it denotes a set of sets that contain at least 'one' of such plural entities.

(71) a. $\exists 2 := \lambda P_{\langle e,t\rangle} \lambda Q_{\langle e,t\rangle}.|P \cap Q| = 2$

b. $\exists := \lambda P_{\langle e,t\rangle} \lambda Q_{\langle e,t\rangle}.P \cap Q \neq \varnothing$

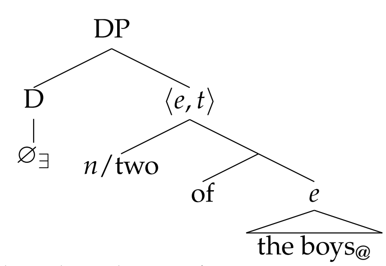
This assumption is supported by the contrast between (72a) and (72b): unlike the distributive ∀-quantifiers *every/each boy*, the indefinite *two (of the) boys* can grammatically combine with a collective predicate such as *formed a team*. This contrast argues that the indefinite *two (of the) boys* in lexicon is not distributive; it cannot be analyzed as an existential distribution over two atomic boys.

(72) a. Every/Each boy joined/*formed a team.

b. Two (of the) boys joined/formed a team.

The composition of *two of the boys* precedes as in (73). First, *of* combines with a plural entity denoted by the *the*-phrase and returns a set of subparts of this entity. Next, the numeral *two*, as a predicate restrictor, combines with a set of entities and returns a subset consisting of only the entities that have exactly two atomic (Aᴛ) subparts, as in (73c,d). Finally, a covert existential determiner $\varnothing_\exists$ combines with this set-denoting NumP and returns an existential generalized quantifier (Link 1987).

(73) *two of the boys*



Assume that the discourse domain has three boys a,b,c:

a. $[\![\textit{the boys}_@]\!] = a \oplus b \oplus c$

b. $[\![\textit{of}]\!] = \lambda x_e.\{y \mid y \leq x\}$

c. $[\![\textit{two}]\!] = \lambda Q_{\langle e,t\rangle}.\{x \mid |\textsc{At}(x)| = 2 \wedge Q(x)\}$

d. $[\![\textit{two of the boys}_@]\!] = \{a \oplus b, b \oplus c, a \oplus c\}$

e. $[\![\varnothing_\exists \textit{ two of the boys}_@]\!] = \lambda P_{\langle e,t\rangle}.\exists x[|\textsc{At}(x)| = 2 \wedge \mathsf{boys}_@(x) \wedge P(x)]$
$= \lambda P_{\langle e,t\rangle}.\exists x \in \{a \oplus b, b \oplus c, a \oplus c\}[P(x)]$

Let's return to the composition of the matrix ∃2-question. The same as in (68b), here the minimal $\boldsymbol{K}$ sets yielded by the application of the $f_{\text{CH}}^{\text{MIN}}$-operator are all 'singleton' sets. Each of these sets consists of one single proposition with the form $\ulcorner x \textit{ watched } \boldsymbol{f}(x) \urcorner$ where x is the plurality of two boys, as in (74b). Hence, the derived QIQ-reading is [−PL], just like in the ∃1-question.

(74) Which movie did two of the boys watch? (QIQ-reading)

$[_{\text{CP}}{}^{\text{BD}}(\text{wh-movie}_@)\ i\ [_{②} \bar{\bigcap}\ [_\gamma\ f_{\text{CH}}^{\text{MIN}}\ \lambda \boldsymbol{K}\ [_{①}\ \text{two-of-the-boys}_@\ j\ [\ \boldsymbol{K}\ (\lambda w.x_j\text{-watched}_w\text{-}\boldsymbol{f}_i(x_j))\ ]]]]]$

a. $[\![①]\!] = \exists x \in \mathsf{2\text{-}boys}_@[\boldsymbol{K}(\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}_i(x)))]$
('$\mathsf{2\text{-}boys}_@$' abbreviates the set of entities that are pluralities of two boys in @.)

b. $[\![\gamma]\!] = \{\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}_i(x))\}$, where x is the chosen two boys

c. $[\![②]\!] = \lambda w.\mathsf{wat}_w(x, \boldsymbol{f}_i(x))$, where x is the chosen two boys

d. $[\![\text{CP}]\!] = \lambda \boldsymbol{f}_{\langle e,e\rangle} : \text{Ran}(\boldsymbol{f}) = \mathsf{mov}_@.\lambda w[\mathsf{wat}_w(x, \boldsymbol{f}(x))]$, where x is the chosen two boys

In contrast to matrix ∃-questions, extensional embeddings of ∃-questions sometimes admit pair-list readings (Szabolcsi 1997a; Beghelli 1997; Appendix B). For example, the embedding sentence (75) is felicitous even if each boy watched a different movie. I assume that this sentence has the LF in (75a) and the meaning in (75b). In this LF, the indefinite moves over the embedding verb *know*. Its trace in the matrix clause is associated with a covert distributor EACH, which yields the 'EACH ≫ ι' reading.

(75) Susi knows [which movie two of the boys watched]. (∃2 ≫ EACH ≫ ι)

a. [[$\varnothing_\exists$ two-of-the-boys$_@$] λx_e [[x EACH] λy_e [Susi knows which movie y watched]]]

b. $\exists x[x \in \mathsf{2\text{-}boys}_@ \wedge \forall y \in \textsc{At}(x)[[\![\textit{Susi knows which movie y watched}]\!]]]$

This analysis is supported by the contrast between (75) and (76): adding overt *each* to the embedded question makes the pair-list reading unavailable.[26] If the matrix trace of the indefinite is associated with covert EACH, the local trace y would be atomic, which cannot be associated with overt *each*.

(76) Susi knows [which movie two of the boys **each** watched]. (∃2 ≫ ~~EACH~~ ≫ ι)

[[$\varnothing_\exists$ two-of-the-boys$_@$] λx_e [[x EACH] λy_e [Susi knows which movie y (#**each**) watched]]]

In matrix ∃-questions, however, pair-list readings cannot be licensed by VP-EACH/*each*. In sentence (77), the contribution of *each* to meaning is that the two chosen boys watched the (same) movie separately, not that they watched possibly a different movie. The presented analysis straightforwardly explains this fact: to derive a pair-list QIQ-reading as in the ∀-question, the quantificational predication condition has to be distributive. Such distributivity arises only if one of the following conditions is met: (i) the quantifier itself is lexically distributive, or (ii) a separate distributor appears between the quantifier and the null predication-operator $\boldsymbol{K}$. Condition (i) is satisfied in questions

[26] I thank an anonymous reviewer of *L&P* for bringing up this data to my attention.

with an *each/every*-subject. Condition (ii) does not apply to English, because VP-EACH/*each* can only appear within an IP as in (77a), not as high as in (77b).

(77) Which movie did two of the boys **each** watch? ($\exists 2 \gg$ ~~EACH~~ $\gg \iota$)

a. ... [two-of-the-boys$_{@}$ λx_e [$\boldsymbol{K}$ [$_{\text{IP}}$ λw [[x **each**] λy_e [y watched$_w$ $\boldsymbol{f}(y)$]]]]]

b. * ... [two-of-the-boys$_{@}$ λx_e [[x **each**] λy_e [$\boldsymbol{K}$ [$_{\text{IP}}$ λw [y watched$_w$ $\boldsymbol{f}(y)$]]]]]

6.4.3. Questions with a negative quantifier

Negative quantifiers do not participate in QIQ-readings. For example, the NO-question in (78) can be responded to by specifying a single movie or a Skolem function to atomic movies, but not by silence.

(78) Which movie did no boy/ none of the boys watch? (✓Individual, ✓Functional, ✗QIQ)
Hulk./ The movie that his grandpa recommended./ #[Silence]

The proposed analysis easily explains why QIQ-reading are not available in NO-questions. The minimal set that contains none of the propositions with the form ⌜*boy-x watched* $\boldsymbol{f}_i(x)$⌝ is the empty set, whose conjunction is undefined. Hence, composing a NO-question with the proposed LF schema for QIQ-readings yields a deviant topical property which maps any input to undefinedness.

(79) Which movie did none of the boys watch? (✗QiQ-reading)
[$^{\text{BD}}$(wh-movie$_{@}$) i [$_{②}$ $\bar{\bigcap}$ [$_{\gamma}$ $f_{\text{CH}}^{\text{MIN}}$ $\lambda \boldsymbol{K}$ [$_{①}$ none-of-the-boys$_{@}$ j [$\boldsymbol{K}$ $(\lambda w.x_j$-watched$_w$-$\boldsymbol{f}_i(x_j))$]]]]]

a. $[\![\textit{none of the boy}_{@}]\!] = \lambda P_{\langle e,t\rangle}.\neg\exists x[\mathsf{boys}_{@}(x) \wedge P(x)]$

b. $[\![①]\!] = \neg\exists x \in \mathsf{boys}_{@}[\boldsymbol{K}(\lambda w.\mathsf{wat}_w(x, \boldsymbol{f}_i(x)))]$

c. $[\![\gamma]\!] = f_{\text{CH}}^{\text{MIN}}([\![\lambda \boldsymbol{K}.①]\!]) = \varnothing$

d. $[\![②]\!]$ is undefined

Relatedly, recall that QIQ-readings are unavailable if the quantificational subject is a GQ-coordination involving a negative conjunct, as seen in (80a,b). Such questions, if analyzed with the LF schema for QIQ-readings, give rise to meanings equivalent to the QIQ-readings of the questions in (81a,b) which do not contain a negative conjunct. For example, for (80a), the minimal set that contains every proposition with the form ⌜*boy-x watched* $\boldsymbol{f}(x)$⌝ and no proposition with the form ⌜*teacher-x watched* $\boldsymbol{f}(x)$⌝ is simply the set of propositions with the form ⌜*boy-x watched* $\boldsymbol{f}(x)$⌝.

(80) a. Which movie did [each of the boys and none of the teachers] watch? (✗QIQ)

b. Which movie did [one of the girls and none of the teachers] watch? (✗QIQ)

(81) a Which movie did [each of the boys] watch? (✓QIQ)

b Which movie did [one of the girls] watch? (✓QIQ)

Given the semantic equivalence between (80a,b) and (81a,b), I explain the blocking effect of negative quantifiers in terms of the Efficiency Constraint. According to this constraint, the QIQ-structures of (80a,b) are ill-formed because of the existence of the simplifications in (81a,b).[27]

(82) Efficiency (*a la* Meyer 2013)

a. LF α is ill-formed if there is an LF β s.t. β is a simplification of α.

b. β is a simplification of α iff (i) $[\![\alpha]\!] = [\![\beta]\!]$, and (ii) β can be derived from α by replacing nodes in α with their subconstituents.

[27] I thank an anonymous reviewer of *L&P* for pointing out the direction of this explanation.

6.4.4. Questions with a counting quantifier

Decreasing quantifiers (e.g., *at most two boys*, *less than three boys*) do not license QIQ-readings. In (83), the boy(s)-movie-pair answer (83b) is not a choice answer; instead, it is an individual answer, where uniqueness scopes above the quantifier.

(83) Which movie did at most two/ less than three boys watch?
'For at most two/ less than three boys x, [tell me] which movie did x watch?'

a. *Hulk*. (Intended: '*Hulk* is the only movie watched by at most two/ less than three boys. The other movies were watched by more boys.')

b. Andy and Billy watched *Hulk*.

i. ✓ Individual reading: '*Hulk* is the **only** movie watched by at most two/ less than three boys, who are Andy and Billy. The other movies were watched by more boys.'

ii. ✗ Choice reading: 'Andy and Billy are two boys who both watched **only** *Hulk*.'

It is quite appealing to extend the analysis proposed for negative quantifiers to these decreasing quantifiers. Following Hackl (2000), Xiang (2019a) decomposes a decreasing quantifier into a negative determiner *no* and a set-denoting restrictor, as in (84). With this decompositional analysis, the unavailability of QIQ-readings in (83) can be explained in the same way as in the NO-question (79).

(84) a. $[\![\textit{at most two boys}_@]\!] = \lambda P_{\langle e,t\rangle}.\neg\exists x[|\text{AT}(x)| > 2 \wedge \textsf{boys}_@(x) \wedge P(x)]$

b. $[\![\textit{less than three boys}_@]\!] = \lambda P_{\langle e,t\rangle}.\neg\exists x[|\text{AT}(x)| \geq 3 \wedge \textsf{boys}_@(x) \wedge P(x)]$

However, despite having a non-decreasing subject, sentence (85) doesn't admit a QIQ/choice-reading either. Like in (83), here the uniqueness inference triggered by the singular *wh*-object must scope above the quantificational subject. This fact argues that the unavailability of QIQ-readings in (83) and (85) has nothing to do with the monotonicity pattern of the quantificational subject.

(85) Which movie did at least/ exactly two boys watch? (✓Individual, ✓Functional, ✗QIQ)
'For at least/ exactly two boys x, [tell me] which movie did x watch?

In contrast to Xiang 2019a, this paper attributes the unavailability of QIQ-readings in (83) and (85) to a syntactic constraint that counting quantifiers are scopally unproductive (Szabolcsi 1997b; Beghelli and Stowell 1997). Beghelli and Stowell (1997) classify non-interrogative quantifiers into the following categories and argue that they have different landing sites. In particular, counting quantifiers have very local scope (take scope essentially in situ) and resist specific interpretations.

(86) Types of non-interrogative quantifiers (Beghelli and Stowell 1997)

a. Negative quantifiers: *no*-NP.

b. Universal-distributive quantifiers: *every/each*-NP

c. Grouping quantifiers: indefinites like *a*/*some*/*several*-NP, bare-numeral quantifiers (e.g., *one student*, *three students*), and *the*-phrases.

d. Counting quantifiers: decreasing quantifiers headed with determiners like *few*, *fewer than five*, and *at most six* and generally cardinality expressions with a modified numerals (e.g., *more than five*, *between six and nine*).

To derive a QIQ-reading, the quantifier must escape IP and move across a null predication-operator ***K***. Counting quantifiers cannot have such global scope and hence do not participate in QIQ-readings.

6.4.5. Questions with a non-quantificational subject

For questions with a non-quantificational subject, the difference between their individual reading and the reading generated from the LF schema for QIQ-readings is trivial. For example, the QIQ-answer to (87), if available, is the conjunction of the minimal set containing the proposition '*The boys watched f(the-boys)*', which is simply this proposition itself. Although it is hard to tell whether QIQ-readings are truly available in these questions, composing these questions with the proposed LF schema for QIQ-readings does not over-generate any unwanted meanings.

(87) Which movie did the boys watch?

a. $[_{\text{CP}}{}^{\text{BD}}(\text{wh-movie}_{@})\ i\ [_{②}\ \bar{\bigcap}\ [_{\gamma}\ f^{\text{MIN}}_{\text{CH}}\ \lambda \boldsymbol{K}\ [_{①}\ \textsc{lift}(\text{the-boys}_{@})\ j\ [\ \boldsymbol{K}\ (\lambda w.x_j\text{-watched}_w\text{-}\boldsymbol{f}_i(x_j))\ ]]]]]$

b. $\lambda \boldsymbol{f}_{\langle e,e\rangle} : \text{Ran}(\boldsymbol{f}) = \mathsf{mov}_{@}.\lambda w[\mathsf{wat}_w(\mathsf{the\text{-}boys}_{@}, \boldsymbol{f}(\mathsf{the\text{-}boys}_{@}))]$

Since non-quantificational expressions are not distributive in lexicon, the readings derived from the LF schema of QIQ-readings are not pair-list. The same as the matrix ∃2-question in (77), distributivity from VP-*each* is applied locally within IP as in (88a), not as high as in (88b,c). Hence, the QIQ-readings of questions with a non-quantificational subject are [+D-EXH,−PL,−CH].[28]

(88) Which movie did the boys **each** watch?

a. ... [LIFT(the-boys$_{@}$) λx_e [$\boldsymbol{K}$ [$_{\text{IP}}$ λw [[x **each**] λy_e [y watched$_w$ $\boldsymbol{f}(y)$]]]]]

b. * ... [LIFT(the-boys$_{@}$) λx_e [[x **each**] λy_e [$\boldsymbol{K}$ [$_{\text{IP}}$ λw [y watched$_w$ $\boldsymbol{f}(y)$]]]]]

c. * ... [[the-boys$_{@}$ **each**] λy_e [$\boldsymbol{K}$ [$_{\text{IP}}$ λw [y watched$_w$ $\boldsymbol{f}(y)$]]]]

Strikingly, Johnston (2019) observes cases like (89a) where it appears that the definite plural *the players* licenses a pair-list reading. He further find out that such pair-list readings exhibit a subject–object asymmetry, just like what is observed in ∀-questions.[29] To account for these observations, Johnston assumes that the definite plural carries a covert DP-internal EACH, which turns this definite plural into a universal distributive quantifier.

(89) (Context: In a basketball team, each of the five players chose a jersey, numbered from 1 to 5.)

a. Which numbers did the players pick?
Ann picked 1, Ben picked 2, Chris picked 3, Dan picked 4, Emma picked 5.

b. Which players picked the numbers?
#Ann picked 1, Ben picked 2, Chris picked 3, Dan picked 4, Emma picked 5.

In what follows, however, I argue that the seeming pair-list reading in (89a) is not a QIQ-reading; instead, it is a (non-QIQ) functional reading involving '*respective* distributivity'. First of all, to see why it is not a QIQ-reading, compare the following questions in the same pair-list context:

(90) (Context: In a basketball team, each of the five players chose a jersey, numbered from 1 to 5.)

a. Which {#numbers, number} did each of the players pick?

b. Which {numbers, #number} did the players pick?

In the ∀-question (90b), the *wh*-object must be singular because each player picked only one number; in (90a), the *wh*-object must be plural because multiple numbers were picked collectively. This

[28]Here domain exhaustivity is trivially satisfied. For example, the set that the Montagovian individual 'LIFT(the-boys)' ranges over is a singleton set containing only the plural entity denoted by *the boys*.

[29]I thank Bernhard Schwarz (pers. comm.) for bringing this issue to my attention.

contrast argues that (90a) and (90b) have different question nucleuses; if these questions have the same nucleus, they would allow for the same *wh*-phrases.

Why does (89a) admit a pair-list answer? The discussion above has excluded the possibility of applying a DP-internal EACH to the definite plural: if it were available, *the players* would function the same as *each of the players*, which leaves the contrast in (90) unexplained. The licensing of pair-list cannot be ascribed to a (covert) VP-*each* either: adding overt *each* to the question makes it infelicitous, since it implies that each player picked multiple numbers, contrary to context.

(91) (Context: In a basketball team, each of the five players chose a jersey, numbered from 1 to 5.)
Which numbers did the players (??each) pick?

I argue that the seeming pair-list reading of (89a) is a functional reading with *respective* distributivity. The question–answer pair is paraphrased as follows:

(92) 'Which numbers did the players pick, respectively?'
'The players Ann,Ben,Chris,Dan,Emma picked the numbers 1-to-5, respectively.'

Formally, *respective* distributivity is derived via the application of a covert operator Resp_g (Gawron and Kehler 2004; Chaves 2012; Law 2019): Resp_g combines with two pluralities (i.e., a plural predicate P and a plural individual x), breaks them into parts, pairs the parts using a pragmatically available sequencing function g, and performs a pair-wise evaluation facilitated by g.

(93) $\text{Resp}_g := \lambda P \lambda x. \forall i[1 \leq i \leq |g| \rightarrow [g(P)(i)](g(x)(i))]$
(The i-th part of the property P holds for the i-th part of the individual x.)

Question (89a) is composed as follows. The same as in any functional reading, the subject *the players* binds into the complex functional trace of the *wh*-object, yielding a *wh*-dependency. However, unlike other functional readings, here a Resp_g-operator is applied between the predicate $\textit{pick}_w\text{-}\boldsymbol{f}_i(x_j)$ and the subject-trace, yielding *respective* distributivity.

(94) $[_{\text{CP}}\ {}^{\text{BD}}(\text{wh-numbers}_{@})\ i\ \lambda w.\ [_{\text{IP}}\ \text{the-players}_{@}\ j\ [_{\text{VP}}\ x_j\ \text{Resp}_g\ \text{picked}_w\text{-}\boldsymbol{f}_i(x_j)\ ]]]$

Since *respective* distributivity involves a dependency between the two arguments of Resp_g, the subject–object asymmetry in (89) can be explained in terms of constraints on dependency. Moreover, this analysis explains why the *wh*-object must be plural: Resp_g requires the predicate $\text{pick}_w(\boldsymbol{f}_i(x_j))$ to be plural, which further requires the range of $\boldsymbol{f}_i$ to be a set of plurals.

6.5. Interim summary

To sum up the core analysis, I argued that pair-list readings of multi-*wh* questions and QIQ-readings of questions with a quantificational subject are extensional functional readings. As schematized in (95) and described in (96), the composition of these questions precedes in four steps.

(95) a. Which boy watched which movie? (Pair-list reading)
$[_{\text{D}}\ {}^{\text{BD}}(\text{wh-movie}_{@})\ i\ [_{\text{C}}\ \bar{\bigcap}\ \lambda p_{\langle s,t\rangle}\ [_{\text{B}}\ \text{wh-boy}_{@}\ j\ [\ [\text{ID}\ p]\ [_{\text{A}}\ \lambda w.\ x_j\ \text{watched}_w\ \boldsymbol{f}_i(x_j)\ ]]]]$
b. Which movie did Det-boy watch? (QIQ-reading)
$[_{\text{D}}\ {}^{\text{BD}}(\text{wh-movie}_{@})\ i\ [_{\text{C}}\ \bar{\bigcap}\ f^{\text{MIN}}_{\text{CH}}\ \lambda \boldsymbol{K}_{\langle st,t\rangle}\ [_{\text{B}}\ \text{Det-boy}_{@}\ j\ [\ \boldsymbol{K}\ [_{\text{A}}\ \lambda w.\ x_j\ \text{watched}_w\ \boldsymbol{f}_i(x_j)\ ]]]]$

(96) (A) Indexations with the two traces yield a *wh*-dependency;

(B) the quantificational/*wh*- subject binds into the *wh*-dependency sentence across an identity/predication operator.

(C) conjoining a set of propositions with the dependency form (A) yields a graph description.

(D) the fronted *wh*-object restricts the range of the input functions.

Table 2 compares the nucleus denotations of five complex questions. In all these questions, the asserted component of the nucleus denotation is the conjunction of a set of propositions representing the graph of the input function f. In the four questions with a non-*wh*-subject, the quantificational predication condition yielded at (B) yields a definedness condition that restricts the domain of f.

Subject-type	Domain condition of f	Graph description of f	D-EXH	PL	CH
which boy		$\bar{\bigcap}\{\lambda w.\mathsf{W}_w(x, f(x)) \mid \mathsf{B}_@(x)\}$	−	+	−
every/each boy	$\forall x \in \mathsf{B}_@[x \in \mathrm{Dom}(f)]$	$\bar{\bigcap}\{\lambda w.\mathsf{W}_w(x, f(x)) \mid \mathsf{B}_@(x)\}$	+	+	−
n of the boys	$\exists x \in n\text{-}\mathsf{Bs}_@[x \in \mathrm{Dom}(f)]$	$\bar{\bigcap}\{\lambda w.\mathsf{W}_w(x, f(x))\}$ where $x \in n\text{-}\mathsf{Bs}_@$	−	−	+
LIFT(*the boys*)	$\text{the-}\mathsf{Bs}_@ \in \mathrm{Dom}(f)$	$\bar{\bigcap}\{\lambda w.\mathsf{W}_w(x, f(x))\}$ where $x = \text{the-}\mathsf{Bs}_@$	+	−	−
none of the boy	$\neg\exists x \in \mathsf{B}_@[x \in \mathrm{Dom}(f)]$	$\bar{\bigcap}\varnothing$	−	−	−

Table 2: Comparing the denotation of the question nucleus

In questions with a quantifier, the QIQ-effect is derived by extracting one of the minimal proposition sets that satisfy the quantificational predication condition yielded at (B). This analysis explains the contrasts between ∀-questions and ∃-questions w.r.t. the following parameters:

- [±D-EXH]: As in a ∀-question, the yielded QIQ-reading presupposes domain exhaustivity if the quantificational predication condition yielded at (B) is subject to a definedness condition that the input f is defined for 'every' element in the quantification domain of the subject.
- [±PL]: As in a ∀-question, with other conditions being equal, the yielded QIQ-reading admits pair-list answers only if there is a 'non-singleton' set of propositions that minimally satisfies the quantificational predication condition yielded at (B). To derive such a non-singleton minimal set, the quantificational subject must be lexically distributive.
- [±CH]: As in an ∃-question, with other conditions being equal, the yielded QIQ-reading has a choice flavor if there are 'multiple' minimal proposition sets that satisfy the quantificational predication condition yielded at (B).

I further why in many cases QIQ-readings are unavailable. In NO-questions, QIQ-readings are semantically deviant because the only minimal proposition set that satisfies a negative quantificational predication condition is the empty set. In questions with a counting quantifier, the LF schema for QIQ-readings is infeasible because counting quantifiers are unproductive in scoping.

I finally discussed another source of pair-list readings in questions with a plural definite subject: although plural definites are not distributive in lexicon, pair-list readings might arise through a locally applied *respective* distributor.

7. Quantificational variability effects

As seen in Sect. 4.1.2, defining pair-list questions as sets of conjunctive propositions, the analysis of Dayal (1996, 2017) cannot account for the Q-variability effects in the embeddings of pair-list questions.

Dayal defines simplex and pair-list questions uniformly as sets of propositions. For embeddings of simplex questions, the most natural way for Dayal to define the Q-variability inference is to let the matrix adverbial quantify over a set of atomic propositions, as exemplified in (97).

(97) Jill mostly knows [which students left].
$\leadsto$ 'Most p: p is a true proposition with the form $\ulcorner$*student-x left*$\urcorner$, Jill knows p.'

This proposition-based definition, however, is infeasible for embeddings of pair-list questions if a pair-list question denotes a set of conjunctive propositions (Lahiri 2002). For example, in a scenario where the three relevant boys b_1,b_2,b_3 watched and only watched the movies m_1,m_2,m_3, respectively, the strongest true propositional answer to the embedded pair-list question in (98) is $\lambda w.\mathsf{wat}_w(m_1, b_1) \wedge \mathsf{wat}_w(m_2, b_2) \wedge \mathsf{wat}_w(m_3, b_3)$, and the Q-variability inference is true if Jill knows at least two of the three atomic conjuncts, as in (98a); however, these conjuncts cannot be semantically retrieved out of their conjunction. In contrast, family-of-questions approaches such as Fox 2012a,b can derive this inference by letting the matrix adverb quantify over a set of sub-questions, as paraphrased in (98b).

(98) Jill mostly knows [$_{\text{PAIR-LIST}}$ $\left\{\begin{array}{l}\text{which movie every boy watched}\\ \text{which boy watched which movie}\end{array}\right\}$].

a. $\leadsto$ 'Most p: p is a true proposition of the form $\ulcorner$*boy-x watched movie-y*$\urcorner$, Jill knows p.'

b. $\leadsto$ 'Most Q: Q is a question the form $\ulcorner$*which movie boy-x watched*$\urcorner$, Jill knows Q.'

c. $\leadsto$ 'Most $\langle x, y\rangle$: $\langle x, y\rangle$ is a boy-movie pair and x watched y, Jill knows that x watched y.'

Although this paper does not pursue a family-of-questions approach, the assumed categorial approach to question composition unlocks the option (98c), where the quantification domain of *mostly* is a set of atomic functions. In my proposal, a pair-list question denotes a topical property that maps each input $\langle e, e\rangle$-type function to a conjunctive proposition. From this topical property, we can extract the function that yields the strongest true answer to this question and define the quantification domain of *mostly* as a set of atomic subparts of this function. For example in (100), the strongest true answer is the function (100a), and its atomic subparts are those in (100b).

(99) a. A function f is atomic iff $\bigoplus \mathrm{Dom}(f)$ is atomic.

b. $\textsc{At}(f) = \{f' \mid f' \subseteq f \text{ and } f' \text{ is atomic}\}$

(100) Which boy watched which movie?/ Which movie did every boy watch?
(The discourse domain includes three boys b_1,b_2,b_3 and three movies m_1,m_2,m_3. In a world w, b_1 watched only m_1, b_2 watched only m_2, and b_3 watched only m_3.)

a. $\textsc{Ans}^S(w)([\![\mathrm{Q}]\!]) = \begin{bmatrix} b_1 \to m_1 \\ b_2 \to m_2 \\ b_3 \to m_3 \end{bmatrix}$

b. $\textsc{At}(\textsc{Ans}^S(w)([\![\mathrm{Q}]\!])) = \left\{\begin{array}{l} [b_1 \to m_1] \\ [b_2 \to m_2] \\ [b_3 \to m_3] \end{array}\right\}$

Xiang 2020 provides two ways to define a Q-variability inference based on short answers. Ignoring the complications needed for accounting for mention-some readings, I schematize these two definitions as in (101a,b).[30] (For a compositional derivation, see Cremers 2018.) In both definitions, the quantification domain of the matrix adverbial *mostly* is a set of atomic entities or a set of atomic $\langle e, e\rangle$-type functions.

[30] Xiang 2020 considers also mention-some readings of questions, where a question can have multiple complete true answers. Once mention-some reading is concerned, $\textsc{Ans}^S(w)([\![\mathrm{Q}]\!])$ needs to be defined as a set of entities/functions, not one single entity/function.

(101) The Q-variability inference of 'Jill mostly knows Q':

a. $\lambda w.\text{MOST}\ x[x \in \text{AT}(\text{ANS}^S(w)(\llbracket Q \rrbracket))][\mathsf{know}_w(j, \llbracket Q \rrbracket(x)]$
(For most x s.t. x is an atomic subpart of the strongest true short answer to Q, Jill knows the inference $\llbracket Q \rrbracket(x)$.)

b. $\lambda w.\text{MOST}\ x[x \in \text{AT}(\text{ANS}^S(w)(\llbracket Q \rrbracket))][\mathsf{know}_w(j, \lambda w'.x \leq \text{ANS}^S(w')(\llbracket Q \rrbracket))]$
(For most x s.t. x is an atomic subpart of the strongest true short answer to Q, Jill knows that x is a subpart of the strongest true short answer to Q.)

In (101a), the scope of the adverbial *mostly* says that Jill knows an atomic proposition, which is derived by applying the topical property of the embedded question to an entity or an $\langle e, e \rangle$-type function x, where x is an atomic subpart of the strongest true answer to the embedded question. This definition works for embeddings of multi-*wh* questions, but not for embeddings of ∀-question: the topical property of the pair-list ∀-question *which movie every boy watched* is only defined for $\langle e, e \rangle$-type functions that are defined for every boy, not for atomic functions such as $[b_1 \rightarrow m_1]$.

Alternatively, in (101b), the scope of *mostly* says that Jill knows a sub-divisive inference, which is semantically equivalent to that Jill correctly identifies most of the boy-watched-movie pairs. This definitions works also for ∀-questions. In the context described in (100), this sub-divisive inference is true iff in every world w' s.t. w' is compatible with Jill's belief, the strongest true short answer to the embedded ∀-question in w' is among the seven functions in Figure 4. This figure illustrates a partition of possible worlds grouped based on which movie each of the three boys watched. The world w described in (100) belongs to the middle cell. In the other cells, correspondences conflicting with w are colored in gray. It is straightforward to see that the union of the seven cells is equivalent to the following proposition: 'for most of the pairs $\langle b, m \rangle$ in $\{\langle b_1, m_1 \rangle, \langle b_2, m_2 \rangle, \langle b_3, m_3 \rangle\}$, b watched m.' Knowing this inference means correctly identifying most of the three boy-watched-movie pairs.

	$\begin{bmatrix} b_1 \rightarrow m_2 \\ b_2 \rightarrow m_2 \\ b_3 \rightarrow m_3 \end{bmatrix}$	$\begin{bmatrix} b_1 \rightarrow m_3 \\ b_2 \rightarrow m_2 \\ b_3 \rightarrow m_3 \end{bmatrix}$
$\begin{bmatrix} b_1 \rightarrow m_1 \\ b_2 \rightarrow m_1 \\ b_3 \rightarrow m_3 \end{bmatrix}$	$\begin{bmatrix} b_1 \rightarrow m_1 \\ b_2 \rightarrow m_2 \\ b_3 \rightarrow m_3 \end{bmatrix}$	$\begin{bmatrix} b_1 \rightarrow m_1 \\ b_2 \rightarrow m_3 \\ b_3 \rightarrow m_3 \end{bmatrix}$
$\begin{bmatrix} b_1 \rightarrow m_1 \\ b_2 \rightarrow m_2 \\ b_3 \rightarrow m_1 \end{bmatrix}$	$\begin{bmatrix} b_1 \rightarrow m_1 \\ b_2 \rightarrow m_2 \\ b_3 \rightarrow m_2 \end{bmatrix}$	

Figure 4: Illustration of the sub-divisive inference in the quantification scope of (101b)

8. Conclusions

In this paper started with a novel observation that pair-list ∀-questions and their multi-*wh* counterparts are semantically different — only the ∀-questions are subject to domain exhaustivity. Given this contrast, I argued that the structure of composition of a pair-list ∀-question must be distinct from that of its multi-*wh* counterpart. Furthermore, drawing on the uniform syntactic constraints on distributing QɪQ-readings, I concluded that QɪQ-readings of matrix questions should be derived uniformly.

Influential accounts such as Dayal 1996, 2017 and Fox 2012a,b are not aware of the contrast in domain exhaustivity between ∀-questions and multi-*wh* questions. These accounts treat pair-list questions uniformly and compose these questions either with the same LF or with different LFs that yield the same root denotation. In addition, to explain why only subject *every/each*-phrases license pair-list readings, these accounts derive pair-list readings in a way that crashes in questions with a non-universal quantifier. In consequence, they over predict domain exhaustivity effects for multi-*wh* questions and fail to account for the choice readings of ∃-questions.

This paper presented a novel analysis of composing complex questions. This analysis had three main ingredients. First, in line with functionality approaches, I proposed that QɪQ-questions and pair-list multi-*wh* questions both involve *wh*-dependencies — the quantificational/*wh*- subject binds the argument index of the functional trace of the *wh*-object. In particular, in a pair-list multi-*wh* question, the *wh*-subject quantifies into an identity condition w.r.t. a proposition expressing this dependency; in a QɪQ-question, the quantificational subject quantifies into a predication condition w.r.t. this dependency. The subtle differences between the two quantificational binding-into-dependency operations are responsible of the contrast in domain exhaustivity between multi-*wh* questions and ∀-questions. Second, for questions with a quantifier, inspired by Fox (2012b), I assumed that the seeming QɪQ-effect is derived by extracting one of the minimal set of propositions that satisfy the quantificational predication condition. This analysis naturally predicts which questions admit QɪQ-readings; it also predicts whether the QɪQ-reading of a question is subject to domain exhaustivity, admits pair-list answers, and has a choice flavor. Finally, assuming a categorial approach, the presented analysis also overcame the difficulty with the functionality analysis of Dayal 1996 in accounting for the Q-variability effects.

Appendix A. The partition-based approach

Section 3 has mentioned that the following LF, repeated from (17), suffers type-mismatch for most frameworks of question semantics:

(102) Which movie did Det-boy watch?
*[Det-boy λx_e [Which movie did x watch]]

Partition Semantics exempts from this type-mismatch problem. Groenendijk and Stokhof (1984: Chap. 3) first analyze the pair-list ∀-question (103) as a partition of possible worlds grouped in terms of which boy watched which movie. In the derivation of this denotation, the quantifier *every boy* quantifies into an identity condition (of type t), which says that x watched the same movies in w and in w'.

(103) Which movie did every boy watch?
$\lambda w \lambda w'.\forall x[\mathsf{boy}_{@}(x) \rightarrow \{y \mid \mathsf{mov}_{@}(y) \wedge \mathsf{wat}_{w}(x,y)\} = \{y \mid \mathsf{mov}_{@}(y) \wedge \mathsf{wat}_{w'}(x,y)\}]$
(w and w' are in the same partition cell iff for every boy x, x watched the same movies in w and in w'.)

However, Groenendijk and Stokhof themselves are not satisfied with this account since it does not extend to questions with a non-universal quantifier. For example, the predicted meaning for the corresponding ∃-question (104) is not a partition (see also Krifka 2001). Thus, Groenendijk and Stokhof ultimately pursues another family-of-questions approach using witness sets (footnote 14).

(104) Which movie did one of the boys watch?
$\lambda w \lambda w'.\exists x[\mathsf{boy}_{@}(x) \wedge \{y \mid \mathsf{mov}_{@}(y) \wedge \mathsf{wat}_{w}(x,y)\} = \{y \mid \mathsf{mov}_{@}(y) \wedge \mathsf{wat}_{w'}(x,y)\}]$

(w and w' are in the same partition cell iff for one of the boys x, x watched the same movies in w and in w'.)

For a concrete illustration, consider a discourse with two boys a,b and two movies m_1,m_2. The four worlds vary by which boy watched which movie. w_1,w_2,w_3 are grouped in one cell C_1: a watched the same movie in w_1 and w_2 (and b watched the same movie in w_1 and w_3). Likewise, w_2,w_3,w_4 all belong to the cell C_2: b watched the same movie in w_2 and w_4. In addition, C_1 and C_2 are distinct cells because neither boy watched the same movie in w_1 and w_4. The world grouping in Figure 5 is clearly not a partition: C_1 overlaps with C_2 — they both contain w_2 and w_3. Moreover, from this world grouping, we cannot identify which movie any of the boys watched. For example, if w_1 is the actual world, then C_1 is the cell which the actual world belongs to; however, based on C_1, we cannot decide on whether a watched m_1 (as in w_1 and w_2) or he watched m_2 (as in w_3).

	C_1 :		C_2 :	
	$w_1 : \{\langle a, m_1\rangle, \langle b, m_2\rangle\}$ (shaded)		$w_1 : \{\langle a, m_1\rangle, \langle b, m_2\rangle\}$	
	$w_2 : \{\langle a, m_1\rangle, \langle b, m_1\rangle\}$ (shaded)		$w_2 : \{\langle a, m_1\rangle, \langle b, m_1\rangle\}$ (shaded)	......
	$w_3 : \{\langle a, m_2\rangle, \langle b, m_2\rangle\}$ (shaded)		$w_3 : \{\langle a, m_2\rangle, \langle b, m_2\rangle\}$ (shaded)	
	$w_4 : \{\langle a, m_2\rangle, \langle b, m_1\rangle\}$		$w_4 : \{\langle a, m_2\rangle, \langle b, m_1\rangle\}$ (shaded)	

Figure 5: World grouping yielded by (104)

In addition, this analysis inherits the theory-internal problems with Partition Semantics. For instance, since Partition Semantics cannot explain the uniqueness effects of singular-*wh* questions (Xiang 2020), the partition-based account cannot explain the point-wise uniqueness effects in pair-list ∀-questions.

Appendix B. The question-embedding approach

Another intuitive and framework-independent way to solve the type-mismatch problem in quantifying-into questions is to reduce matrix questions into question-embeddings (Karttunen 1977; Krifka 2001). The LF assumed by Karttunen (1977) is given in (105). Basically, whatever the embedded question denotes, the question-embedding is a t-type expression which can be quantified into.

(105) Which movie did Det-boy(s) watch?
[Det-boy(s) λx_e [I-ASK-YOU [Which movie did x watch]]]

This analysis crucially requires the quantifier in the embedded question to scope over the intensional embedding predicate ASK. In the following, however, drawing on the limited distribution of pair-list readings in matrix questions and intensional question-embeddings, I argue that this scoping pattern is not available.[31]

As discussed in Sect. 3 and explained in Sect. 6.4, only *every*/*each*-phrases can license pair-list readings for matrix questions. As for question-embeddings, Szabolcsi (1997a) observes a contrast

[31] Krifka (2001) assumes the structure in (i) where the quantifier scopes over a speech act operator QUEST. This analysis exempts from the over-generation problem since Krifka assumes that speech acts cannot be disjoined. However, it also leaves the choice readings of ∃-questions unexplained.

(i) Which movie did every boy watch?
[every-boy λx_e [QUEST [Which movie did x watch]]]

between intensional complements and extensional complements. In particular, in embeddings with an extensional predicate (e.g., *know*, *find out*), numeral-modified indefinites such as *two of the boys* may also license a pair-list reading. For example, in a pair-list context where each boy watched a different movie, (106a,b) can be uttered felicitously and interpreted with the following scopal pattern: '$\exists 2 \gg$ EACH $\gg$ V $\gg \iota$' where 'V' stands for the embedding predicate. As argued in Sect. 6.4.2, this reading can be derived from the LF in (107) (see also (75)): the indefinite takes wide scope relative to the embedding predicate, and its trace in the matrix clause is associated with a covert EACH.[32]

(106) Susi knew that each boy watched a different movie. In addition, ...
  a. Susi **knew** which movie each/two of the boys watched.
  b. Susi **found out** which movie each/two of the boys watched.

(107) Susi V-ed which movie two of the boys watched.
  [[two-of-the-boys λx [[x EACH] λy [Susi V-ed which movie y watched]]]

However, embeddings with an intensional predicate (e.g., *ask*, *wonder*) behave the same as matrix questions — only *every*/*each*-phrases may license pair-list readings in these embeddings. For example, in (108a,b) the uniqueness inference triggered by *which movie* must be interpreted between the embedding predicate and the quantifier: ASK $\gg \iota \gg \exists 2$.

(108) Susi knew that every boy watched a different movie. ...
  a. Susi **wondered** which movie each/#two of the boys watched.
  b. Susi **asked me** which movie each/#two of the boys watched.

The lack of pair-list readings argues that the LF (107) is not available for (108a,b). Szabolcsi (1997a) argues that intensional predicates create weak islands, which prevent the quantifiers in the embedded questions from taking wide scope. If this explanation is on the right track, the embedding structure (105), which requires the quantifier in the embedded question to scope over ASK, should be infeasible.

Acknowledgement [To be added ...]

[32] Rather than assuming covert movement of the quantifier, Szabolcsi (1997a) derives the wide scope reading by type-lifting the interrogative complements of extensional predicates. Combining the type-lifted question-denotation (i) with an embedding predicate P yields a wide scope reading of the quantifier π relative to P. Further, Szabolcsi argues that *wonder*-type predicates cannot select for lifted questions, and hence that quantifiers in intensional complements cannot take wide scope.

(i) Complement of *find out*-type predicates: $\lambda P.\pi(\lambda x.P(\text{which } y[x \text{ watched } y]))$

References

Barwise, Jon, and Robin Cooper. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4(2):159–219.

Beck, Sigrid, and Yael Sharvit. 2002. Pluralities of questions. *Journal of Semantics* 19:105–157.

Beghelli, Filippo. 1997. The syntax of distributivity and pair-list readings. In *Ways of scope taking*, ed. Anna Szabolcsi, 349–408. Dordrecht: Kluwer.

Beghelli, Filippo, and Tim Stowell. 1997. Distributivity and negation: The syntax of *each* and *every*. In *Ways of scope taking*, ed. Anna Szabolcsi, 71–107. Dordrecht: Kluwer.

Berman, Stephen Robert. 1991. On the semantics and logical form of *wh*-clauses. Ph.D. dissertation, University of Massachusetts at Amherst.

Caponigro, Ivano, and Anamaria Fălăuş. 2020. Unveiling multiple *wh*- free relative clauses and their functional *wh*-words. *Semantics and Pragmatics*. DOI:http://dx.doi.org/10.3765/sp.13.2.

Chaves, Rui P. 2012. Conjunction, cumulation and respectively readings. *Journal of Linguistics* 48(2): 297–344.

Chierchia, Gennaro. 1991. Functional WH and weak crossover. In *Proceedings of the West Coast Conference on Formal Linguistics 10*, ed. Dawn Bates, 75–90.

Chierchia, Gennaro. 1993. Questions with quantifiers. *Natural Language Semantics* 1:181–234.

Chierchia, Gennaro, and Ivano Caponigro. 2013. Questions on questions and free relatives. Presentation at the 18th Sinn und Bedeutung, University of the Basque Country (UPV/EHU) in Vitoria-Gasteiz, September 2013.

Ciardelli, Ivano, and Floris Roelofsen. 2018. An inquisitive perspective on modals and quantifiers. *Annual Review of Linguistics* 4:129–149.

Cremers, Alexandre. 2016. On the semantics of embedded questions. Ph.D. dissertation, École normale supérieure, Paris.

Cremers, Alexandre. 2018. Plurality effects in an exhaustification-based theory of embedded questions. *Natural Language Semantics* 26:193–251.

Dayal, Veneeta. 1996. *Locality in WH quantification: Questions and relative clauses in Hindi*. Dordrecht: Kluwer.

Dayal, Veneeta. 2002. Single-pair versus multiple-pair answers: Wh-in-situ and scope. *Linguistic Inquiry* 33:512–520.

Dayal, Veneeta. 2016. List answers through higher order questions. Colloquium talk at MIT, February 2016.

Dayal, Veneeta. 2017. *Questions*. Oxford: Oxford University Press.

Engdahl, Elisabet. 1980. The syntax and semantics of questions in Swedish. Ph.D. dissertation, University of Massachusetts at Amherst.

Engdahl, Elisabet. 1986. *Constituent questions*. Dordrecht: Reidel.

Fox, Danny. 2012a. Multiple wh-questions: Uniqueness, pair-list and second order questions. Class notes for MIT seminars.

Fox, Danny. 2012b. Pair-list with universal quantifiers. Class notes for MIT seminars.

Fox, Danny. 2013. Mention-some readings of questions. Class notes for MIT seminars.

Fox, Danny. 2018. Partition by exhaustification: Comments on Dayal 1996. In *Proceedings of Sinn und Bedeutung 22*, ed. Uli Sauerland and Stephanie Solt, 403–434. Berlin: Leibniz-Centre General Linguistics. 403–434. https://semanticsarchive.net/Archive/jQ5MTcyZ/.

Fox, Danny. 2020. Partition by exhaustification: towards a solution to Gentile and Schwarz's puzzle. Manuscript, MIT. https://semanticsarchive.net/Archive/TljZGNjZ/.

Gallin, Daniel. 1975. *Intensional and higher-order modal logic*. Amsterdam: North-Holland Publishing Company.

Gawron, Jean Mark, and Andrew Kehler. 2004. The semantics of respective readings, conjunction, and filler-gap dependencies. *Linguistics and Philosophy* 27:169–207.

Groenendijk, Jeroen, and Martin Stokhof. 1984. On the semantics of questions and the pragmatics of answers. *Varieties of formal semantics* 3:143–170.

Hackl, Martin. 2000. Comparative quantifiers. Ph.D. dissertation, MIT.

Hagstrom, Paul Alan. 1998. Decomposing questions. Ph.D. dissertation, MIT.

Heim, Irene. 1995. Notes on questions. Class notes for MIT seminars.

Heim, Irene. 2012. Notes on questions. Class notes for MIT semantics.

Jacobson, Pauline. 1994. Binding connectivity in copular sentences. In *Proceedings of Semantics and Linguistic Theory 4*, ed. M. Harvey and L. Santelmann, 161–178. https://doi.org/10.3765/salt.v4i0.2456.

Jacobson, Pauline. 1999. Towards a variable-free semantics. *Linguistics and Philosophy* 22:117–185.

Jacobson, Pauline. 2014. *Compositional semantics: An introduction to the syntax/semantics interface*. Oxford: Oxford University Press.

Johnston, William. 2019. Pair-list questions from dependent plurals. In *Proceedings of the 22nd Amsterdam Colloquium*, 201–210. https://archive.illc.uva.nl/AC/AC2019/uploaded_files/inlineitem/Johnston_Pair-list_questions_from_dependent_plurals.pdf.

Kang, Jungmin. 2012. Why short-form functional reading answers are not possible in multiple *wh*-questions. In *University of Pennsylvania Working Papers in Linguistics*, volume 18. https://repository.upenn.edu/pwpl/vol18/iss1/15.

Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1:3–44.

Kotek, Hadas. 2014. Composing questions. Ph.D. dissertation, MIT.

Kotek, Hadas. 2019. *Composing questions*. Cambridge, MA: MIT Press.

Krifka, Manfred. 1991. Definite noun-phrases are not quantifiers. *Linguistic Inquiry* 23:156–163.

Krifka, Manfred. 2001. Quantifying into question acts. *Natural Language Semantics* 9:1–40.

Lahiri, Utpal. 1991. Embedded interrogatives and predicates that embed them. Ph.D. dissertation, MIT.

Lahiri, Utpal. 2002. *Questions and answers in embedded contexts*. Oxford: Oxford University Press.

Law, Hoi-Ki. 2019. Constraints on distributivity. Ph.D. dissertation, Rutgers University.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In *Meaning, use, and interpretation of language*, ed. Christoph Schwarze Rainer Bäuerle and Arnim von Stechow, 302–323. De Gruyter.

Link, Godehard. 1987. Generalized quantifiers and plurals. In *Generalized quantifiers*, ed. P. Gärdenfors, 151–180. Dordrecht: Reidel.

May, Robert. 1985. *Logical form: Its structure and derivation*, volume 12. Cambridge, MA: MIT press.

May, Robert. 1988. Ambiguities of quantification and *wh*: A reply to Williams. *Linguistic Inquiry* 19:118–135.

Meyer, Marie-Christine. 2013. Ignorance and grammar. Ph.D. dissertation, MIT.

Moltmann, Friederike, and Anna Szabolcsi. 1994. Scope interactions with pair-list quantifiers. In *Proceedings of the North East Linguistic Society 24*, ed. Mercé González, volume 2, 381–395. GLSA Publications, University of Massachusetts, Amherst.

Nicolae, Andreea Cristina. 2013. *Any* questions? Polarity as a window into the structure of questions. Ph.D. dissertation, Harvard University.

Pafel, Jürgen. 1999. Interrogative quantifiers within scope. *Linguistics and Philosophy* 22:255–310.

Partee, Barbara H. 1986. Noun phrase interpretation and type-shifting principles. In *Studies in discourse representation theory and the theory of generalized quantifiers*, ed. J. Groenendijk, D. de Jongh, and M. Stokhof, 357–381. Foris, Dordrecht.

Preuss, Susanne Marie-Luise. 2001. Issues in the semantics of questions with quantifiers. Ph.D. dissertation, Rutgers University.

Richards, Norvin Waldemar. 1997. What moves where when in which languages? Ph.D. dissertation, MIT.

Sharvit, Yael, and Jungmin Kang. 2017. Fragment functional answers. In *Proceedings of Semantics and Linguistic Theory 26*, ed. Mary Moroney, Carol-Rose Little, Jacob Collard, and Dan Burgdorf. 1099–1118. https://doi.org/10.3765/salt.v26i0.3792.

Sharvy, Richard. 1980. A more general theory of definite descriptions. *The Philosophical Review* 89:607–624.

Srivastav, Veneeta. 1991. WH dependencies in Hindi and the theory of grammar. Ph.D. dissertation, Cornell University.

Szabolcsi, Anna. 1997a. Quantifiers in pair-list readings. In *Ways of scope taking*, ed. Anna Szabolcsi, 1–27. Dordrecht: Kluwer.

Szabolcsi, Anna. 1997b. Strategies for scope taking. In *Ways of scope taking*, ed. Anna Szabolcsi, 109–154. Dordrecht: Kluwer.

Williams, Edwin Samuel. 1994. *Thematic structure in syntax*. 23. Cambridge, MA: MIT Press.

Winter, Yoad. 2001. *Flexibility principles in boolean semantics*. Cambridge, MA: MIT Press.

Xiang, Yimei. 2016. Interpreting questions with non-exhaustive answers. Ph.D. dissertation, Harvard University.

Xiang, Yimei. 2019a. Getting quantifying-into questions uniformly: Functionality, domain exhaustivity, and quantificational variability. In *Proceedings of Semantics and Linguistic Theory 29*, ed. Katherine Blake, Forrest Davis, Kaelyn Lamp, and Joseph Rhyne, 160–179. https://doi.org/10.3765/salt.v29i0.4610.

Xiang, Yimei. 2019b. A variable-free approach to *wh*-dependency. Presentation at the Workshop on Complex *wh*-questions, Nantes, November 2019.

Xiang, Yimei. 2020. A hybrid categorial approach to question composition. *Linguistics and Philosophy*. Online-first articles. https://doi.org/10.1007/s10988-020-09294-8.

Xiang, Yimei. 2021. Relativized exhaustivity: uniqueness and mention-some. Manuscript, Rutgers University. https://ling.auf.net/lingbuzz/005322.