Two types of higher-order readings of WH-questions[*]

Yimei Xiang

Rutgers University, New Brunswick, New Jersey, USA
yimei.xiang@rutgers.edu

Abstract

Evidence from questions with modals and collective predicates suggests that sometimes WH-questions must be interpreted with higher-order readings. In such readings, questions expect answers naming generalized quantifiers rather than entities. This paper investigates the distribution and the derivation of those readings. First, I argue that the generalized quantifiers that can serve as semantic answers to questions are subject to two constraints — Positiveness and Homogeneity. Next, I present two ways to account for a distributional constraint with higher-order readings, that is, questions with a singular-marked or numeral-modified WH-complement admit elided disjunctive answers but conjunctive answers.

1 Introduction

WHPs are commonly treated as functions (e.g., existential ($\exists$-)quantifiers or domain restrictors) over first-order (FO-)predicates. In this view, the domain for quantification or abstraction is the set denoted by the extension of the WH-complement. The following uses categorial approaches to illustrate this idea (framework chosen just for ease of presentation): in the question *which boy left*, the WHP *which boy* combines with a FO-predicate $\lambda x_e.\mathsf{left}(x)$ and restricts the domain to the set of atomic boys. The yielded question denotation is a FO-predicate $\lambda x_e : \mathsf{boy}(x).\mathsf{left}(x)$. It can serve as a function for an entity-denoting answer as in (1a), and an argument of a generalized quantifier (GQ)-denoting answer as in (1b).[1] I henceforth call this functional question denotation a "Q-function" and its domain a "Q-domain".

(1) a. $[\![Q]\!]([\![Andy]\!]) = (\lambda x_e : \mathsf{boy}(x).\mathsf{left}(x))(a) = \mathsf{prof}(a).\mathsf{left}(a)$

b. $[\![Andy\ or\ Bill]\!]([\![Q]\!]) = (a^{\Uparrow} \cup b^{\Uparrow})(\lambda x_e : \mathsf{boy}(x).\mathsf{left}(x)) = \mathsf{boy}(a) \wedge \mathsf{boy}(b).\mathsf{left}(a) \vee \mathsf{left}(b)$

In some cases, however, a WH-question can only be completely addressed by a GQ. This fact shows that WH-questions have also higher-order (HO-)readings, in which the yielded Q-functions take GQs as arguments. More strikingly, there seems to be two types of HO-readings, which admit different HO-answers and have different distributions. In particular, questions in which the WHP is number-unmarked (*who*, *what*) or the WH-complement is bare plural (*which children*) admit both conjunctions and disjunctions, while questions in which the *wh*-complement is SG-marked (*which child*) or NUM-modified (*which two children*) admit only disjunctions. I call these two readings 'conjunction-admitting' and 'disjunction-only', respectively.

The rest of this paper is organized as follows. Section 2 provides evidence for the existence of HO-readings. Section 3 discusses the constraints as to which HO-meanings can be included

[1]Here '$a^{\Uparrow} \cup b^{\Uparrow}$' denotes the disjunction of two Montagovian individuals. Disjunctions are standardly treated as set **unions**. Since entities are not sets, they must be Montague-lifted into GQs before being disjoined: for any entity α, the corresponding Montagovian entity is $\alpha^{\Uparrow}$ such that $\alpha^{\Uparrow} = \lambda m_{\langle e,t \rangle}.m(\alpha)$. This idea follows a more general schema in Partee & Rooth 1983. The conjunctive *and* is usually treated as ambiguous between as denoting an **intersection** operator '$\cap$' or a **summation** operator '$\oplus$' (Link 1983; Hoeksema 1988; a.o).

in a Q-domain. Sections 4 and 5 explain the distributional constraints of the two HO-readings and derive them compositionally. Section 6 concludes.

2   Evidence for higher-order readings

Treating questions as FO-functions yield two predictions with respect to GQ-denoting answers. First, the specified GQs must be interpreted with a wide scope relative to the scopal elements in the question nucleus. Second, for any propositional answer p specifying a GQ π, p and π are not in the answer space (viz., Hamblin set) and the Q-domain of this question; instead, the derivation of p involves additional operations to the propositions in the answer space. This section presents counterexamples to both predictions, showing that FO-reading is insufficient.

Evidence for disjunctions and $\exists$-quantifiers   Spector (2007, 2008) observes that elided disjunctions can completely address WH-questions with a necessity modal ("$\Box$-questions" henceforth). As seen in (2b), interpreted with a narrow scope, the disjunction can be the complete specification of John's reading obligations — there is no specific book that John has to read, he just has to choose between the French novels and the Russian novels. The same complete answer reading is available with $\exists$-quantifiers, as in (3).

(2)   "Which books does John have to read?" "The French novels or the Russian novels."
    a. 'John has to read F or R, I do not know which exactly.'   (Partial: $or \gg \Box$)
    b. 'John has to read F or R, and the choice is up to him.'   (Complete: $\Box \gg or$)

(3)   "Which books does John have to read?"
    "At least/ More than/ Exactly two books by Balzac."   ($^{\text{OK}}\exists \gg \Box$, $^{\text{OK}}\Box \gg \exists$)

Spector derives the complete answer reading of (2b) by interpreting the disjunctive answer as a GQ $\mathsf{f}^{\Uparrow} \cup \mathsf{r}^{\Uparrow}$ (see definitions in footnote 1) and assuming semantic reconstruction in the nucleus (*a la* Cresti 1995). Adapting this analysis to the categorial approach, I assume the LF and Q-function as in (4a-b). In the LF, WH-movement leaves a HO-trace π (of type $\langle et, t \rangle$) under the $\Box$-modal *have to*. The yielded Q-function is defined for GQs that range over a set of books.[2] (This domain condition is to be revised.) Combining with this HO-function, a GQ-denoting answer is interpreted at whichever scopal position that π takes, as seen in (4c).

(4)   a. [Which-books $\lambda\pi$ [have-to [$\pi_{\langle et,t\rangle}$ λx [John read x_e]]]]
    b. $[\![Q]\!] = \lambda\pi_{\langle et,t\rangle} : \text{SMLO}(\pi) \subseteq {}^{*}\mathsf{book}.\Box[\lambda w.\pi(\lambda x.\mathsf{read}_w(\mathsf{j}, x))]$ (To be revised)
    c. $[\![Q]\!](\mathsf{f}^{\Uparrow} \cup \mathsf{r}^{\Uparrow}) = \{\mathsf{f}, \mathsf{r}\} \subseteq {}^{*}\mathsf{book}.\Box[\lambda w.\mathsf{read}_w(\mathsf{j}, \mathsf{f}) \vee \mathsf{read}_w(\mathsf{j}, \mathsf{r})]$

$\Box$-questions provide evidence for ruling in disjunctions and $\exists$-quantifiers because they may yield Q-functions that are non-reducible relative to disjunctions and $\exists$-quantifiers, as in (5a). This diagnostic, however, does not extend to conjunctions — $\lambda\pi.\mathsf{J}$ has to read π is reducible relative to conjunctions, as seen in (5b).

(5)   A function θ is reducible relative to a GQ π iff $\theta \circ \pi = \pi(\lambda x_e.\theta \circ x^{\Uparrow})$ where '$\circ$' stands for the operation for combining θ with a GQ. For example:
    a. $[\lambda\pi.\mathsf{J}$ has to read $\pi](\mathsf{f}^{\Uparrow} \cup \mathsf{r}^{\Uparrow}) \neq \mathsf{J}$ has to read $\mathsf{f} \vee \mathsf{J}$ has to read r
    b. $[\lambda\pi.\mathsf{J}$ has to read $\pi](\mathsf{f}^{\Uparrow} \cap \mathsf{r}^{\Uparrow}) = \mathsf{J}$ has to read $\mathsf{f} \wedge \mathsf{J}$ has to read r

---

[2]Here 'SMLO(π)' stands for the smallest **live-on** set of π. For any π (of type $\langle \tau t, t \rangle$) and A (of type $\langle \tau, t \rangle$), we have π lives on A iff $\forall B[\pi(B) \leftrightarrow \pi(A \cap B)]$ (Barwise & Cooper 1981), and π **ranges over** A iff A is the smallest live-on set of π (Szabolcsi 1997). In most cases (except for exceptives such as *every/no student but John*), the smallest live-on set of a GQ of the form 'Det-NP' is the set denoted by the extension of the complement NP.

2

Evidence for conjunctions and ∀-quantifiers   The following presents the arguments from Xiang (2016: ch 1) for ruling in conjunctions, which draw on the lack of uniqueness effects in questions with a stubbornly collective predicate.

First, for stubborn collectivity, observe in (6) that the quantized collective predicates *formed a/one team* admits only a collective reading, in contrast to its plural counterpart *formed teams* and other collective predicates like *lifted the piano* which admit also cumulative/ non-atomic distributive readings. Next, for absence of uniqueness, compare (7a-b) in the same context. The declarative-embedding sentence (7a) suffers a presupposition failure — the factive *know* embeds a collective declarative that is false in the given context. In contrast, the corresponding question-embedding sentence (7b) is felicitous, and it implies the inference (7c) that John knows precisely the component members of all the teams formed by the considered kids.

(6) (*w*: *The four kids formed two teams in total: $a + b$ formed one, and $c + d$ formed one.*)

     a. # The kids formed a/one team.

     b.   The kids formed teams.

(7) a. # John knows [that **the kids** formed a/one team].

     b.   John knows [**which kids** formed a/one team].

     c.   → *John knows that $a + b$ formed a team <u>and</u> $c + d$ formed a team.*

The conjunctive inference in (7c) is quite surprising — where does the conjunctive closure come from? Clearly, no matter how we analyze distributivity/collectivity, this closure cannot come from the predicate *formed a/one team* or anywhere within the question nucleus, otherwise (6a) would admit a distributive/covered reading. I propose that this conjunctive closure comes from the WHP: the WHP quantifies over a set of HO-meanings including the conjunction $(a \oplus b)^{\Uparrow} \cap (c \oplus d)^{\Uparrow}$ (i.e., the intersection of Montagovian plurals). The LF and Q-function are as follows:

(8) Which kids formed a team?

     a. [which-kids $\lambda\pi$ [$_{\text{IP}}$ $\pi_{\langle et,t \rangle}$ λx [$_{\text{VP}}$ x_e formed a team]]]

     b. $[\![Q]\!] = \lambda\pi_{\langle et,t \rangle} : \text{SMLO}(\pi) \subseteq {}^{*}\text{kid}.\lambda w[\pi(\lambda x.\text{f-a-t}_w(x))]$ (To be revised)

     c. $[\![Q]\!]((a \oplus b)^{\Uparrow} \cap (c \oplus d)^{\Uparrow}) = \{a \oplus b, c \oplus d\} \subseteq {}^{*}\text{kid}.\lambda w[\text{f-a-t}_w(a \oplus b) \wedge \text{f-a-t}_w(c \oplus d)]$

One might suggest to ascribe this conjunctive closure to operations outside the question root. The answerhood-operator of Heim (1994), for example, returns the conjunction of the true propositions in the answer space. This idea, however, cannot account for the contrast in (9): (9b) is infelicitous because the NUM-modified WHP yields a false uniqueness inference.

(9) (Context is the same as in (6).)

     a.   John knows [**which kids** formed a/one team].

     b. # John knows [**which two kids** formed a/one team].
        → *Among the relevant kids, exactly two of them formed one single team.*

Uniqueness effects in questions are standardly explained by "Dayal's (1996) presupposition" — a question is defined only if it has a strongest true answer. In the given context, the strongest true answer to the embedded question in (7b) is the conjunctive proposition (8c), derived based on the Boolean conjunction $(a \oplus b)^{\Uparrow} \cap (c \oplus d)^{\Uparrow}$. To account for the uniqueness effect in (9), one just needs to assume that the WHP *which two kids* (in contrast to *which kids*) does not quantify over HO-meanings and thus does not support HO-readings.

Evidence for GQ-coordinations   The diagnostics based on non-reducibility and stubborn collectivity rule in GQs such as dis/con-junctions and $\exists/\forall$-quantifiers. Combinations of these two diagnostics also rule in the coordinations of these GQs. For example, the question (10a) involves both a necessity modal *have to* and a stubbornly collective predicate *present a paper together*. The answer should be interpreted as a Boolean disjunction over Boolean conjunctions of plurals (viz., $((a_1 \oplus b_1)^{\Uparrow} \cap (c_1 \oplus d_1)^{\Uparrow}) \cup ((a_2 \oplus b_2)^{\Uparrow} \cap (c_2 \oplus d_2)^{\Uparrow}))$ and be read with the following scopal pattern: $\square \gg$ *or* $\gg$ *and* $\gg$ *a paper*. To obtain this scopal pattern, the question nucleus should involve a HO-WH-trace in between the modal and the collective predicate, as in (10c).

(10)   a. "Which students have to present a paper together this week?"

      b. "The two junior linguists ($a_1 \oplus b_1$) and the two senior linguists ($c_1 \oplus d_1$), OR, the two junior philosophers ($a_2 \oplus b_2$) and the two senior philosophers ($c_2 \oplus d_2$)."

      c. [which-students $\lambda\pi$ [have to [$\pi_{\langle et,t \rangle}$ λx [x_e present a paper together this week]]]]

3   Constraints of higher-order Q-domain

Section 2 has provided evidence for ruling in a variety of GQs and GQ-coordinations into the Q-domain. This section, in contrast, discusses GQs that should be excluded from the Q-domain.

Consider the truth conditions of the Q-embedding sentence (11b) under the context (11a). Strikingly, (11b) implies that Sue knows John's reading obligation (i), but not that she knows (ii); Sue can be ignorant about whether John should read any books by Betty. Given this contrast, Spector (2008) proposes that the GQs involved in a Q-domain must be increasing: 'x knows Q' implies that x knows the complete/strongest true answer of Q; therefore, that Sue can be ignorant about (ii) excludes the decreasing GQ *no book by Betty* and the non-monotonic GQ-coordination *at least two books by Anne and no book by Betty* from the Q-domain.

(11)   a. Context: John's reading obligations include the following:
        (i) he must read at least two books by Anne; (ii) he must read no book by Betty.

      b. Sue knows which books John must read.

The following example extends Spector's diagnostic to a broader range of GQs. Among the list of GQs, (a-b) are increasing, (c-d) are decreasing, and (e) is non-monotonic. Intuitively, the question-embedding sentence (13) implies that Sue knows about not only (a-b) but also (e).

(12)   John is playing a board game. This game requires that him to play ...

      a. {at least three, more than two} red spades;

      b. every black spade except the smallest one in his hand;

      c. {at most three, less than four} black diamonds;

      d. no red diamond except largest one in his hand;

      e. exactly two hearts;

(13)   Sue knows which cards John must play.

Spector's increasing-ness constraint incorrectly rules out the non-monotonic GQ *exactly two hearts* and fails to predict that (13) implies that Sue knows the requirement (e). In contrast, I propose that whether a HO-meaning should be ruled in into a Q-domain is determined by its **positiveness** — only positive GQs and their Boolean coordinations can be in a Q-domain. A GQ being positive means that it ensures existence. For example, *at least two books* and *exactly two books* entail *some books*, while *no book* and *less than two books* do not entail *some book(s)*. A

formal definition of positiveness is as follows, where some is the basic $\exists$-quantifier, and $\text{SMLO}(\pi)$ is the smallest live-on set of π (see definitions in footnote 2):

(14)   For any π of type $\langle et, t \rangle$, π is **positive** iff $\pi \subseteq \text{some}(\text{SMLO}(\pi))$.

Table 1 compares monotonicity and positiveness for GQs that range over books. Observe that increasing GQs are positive, while decreasing ($\downarrow_{\text{MON}}$) and non-monotonic (N.M.) GQs are not.

GQ π	$\text{SMLO}(\pi)$	Increasing?	Positive?
at least two books	books	Yes	Yes
every book except a	book $- \{a\}$	Yes	Yes
at most two books	books	No ($\downarrow_{\text{MON}}$)	No
no book except a	book $- \{a\}$	No ($\downarrow_{\text{MON}}$)	No
exactly two books	books	No (N.M.)	Yes
two to ten books	books	No (N.M.)	Yes

Table 1: Increasing-ness/monotonicity versus positiveness

Table 1 considers only simplex GQs of the form 'Det+NP'. Benjamin Spector (p.c.) points out that positiveness does not exclude the unwanted non-monotonic GQ-coordinations such as *every article and no book*. Representing this GQ-conjunction π as $\{E \mid A \subseteq E \land B \cap E = \emptyset\}$, we have $\text{SMLO}(\pi) = A \cup B$,[3] and $\pi \subseteq \text{some}(A \cup B)$. Benjamin Spector further points out that a GQ π can be semantically decomposed into the conjunction $\pi^+ \cap \pi^-$, where π^+ is the strongest increasing GQ entailed π, and π^- is the strongest decreasing GQ entailed by π. In particular, π is increasing iff $\pi^- = D_{\langle e,t \rangle}$, and π is decreasing iff $\pi^+ = D_{\langle e,t \rangle}$.

(15)   For any $\pi_{\langle et,t \rangle}$: $\pi^+ =_{\text{df}} \{P \mid \exists P' \subseteq P[\pi(P')]\}$, and $\pi^- =_{\text{df}} \{P \mid \exists P' \supseteq P[\pi(P')]\}$.

To rule out non-monotonic GQ-coordinations like *every article and no book*, we just need to exclude the non-monotonic πs where the retrieved π^+ and π^- range over different sets. For example, the simplex GQ *exactly two books* is not excluded since it is formed by conjoining two GQs ranging over the same set books (i.e., *at least two books* and *no more than two books*). I call such GQs "homogenous", defined as in (16). Table 2 compares homogeneity and positiveness. Observe that *some A and/or no B* is positive but not homogenous.

(16)   For any π of type $\langle et, t \rangle$, π is **homogenous** iff π is monotonic or $\text{SMLO}(\pi^+) = \text{SMLO}(\pi^-)$.

GQ π	$\text{SMLO}(\pi)$	$\text{SMLO}(\pi^+)$	$\text{SMLO}(\pi^-)$	Positive?	Homogenous?
at least two B	B	B	D_e	Yes	Yes
exactly two B	B	B	B	Yes	Yes
every A and no B	$A \cup B$	A	B	Yes	No

Table 2: Positiveness versus homogeneity

In conclusion, the Q-domain yielded by the phrase '*wh*-A' in a HO-reading, if any, is the set consisting of the positive homogenous GQs ranging over a subset of A. I write this set as $^{\text{H}}A$.

---

[3] The following shows why $A \cup B$ is the smallest live-on set of π: for any a, replacing E with $E \cap (A \cup B - \{a\})$ in the set description makes no change to the defined set of E iff $a \notin A$ and $a \notin B$.

(i) $\{E \mid [A \subseteq (E \cap ((A \cup B) - \{a\}))] \land [B \cap (E \cap ((A \cup B) - \{a\})) = \emptyset]\}$
$= \{E \mid [A \subseteq E \land A \subseteq (A \cup B - \{a\})] \land [(B \cap ((A \cup B) - \{a\})) \cap E = \emptyset]\}$
$= \{E \mid [A \subseteq E \land A \subseteq (A \cup B) \land a \notin A] \land [(B - \{a\}) \cap E = \emptyset]\}$ $(= \pi \text{ iff } a \notin A \cup B)$

(17) $^{\text{H}}A = \{\pi_{\langle et,t \rangle} \mid \text{SMLO}(\pi) \subseteq A \wedge \pi \subseteq \mathsf{some}(\text{SMLO}(\pi)) \wedge \pi \text{ is homogenous}\}$

It is yet unclear where the positiveness and homogeneity constraints come from. They could be in the lexicon of a type-shifting operator, presupposed by the HO-WH-trace, or constraints on semantic reconstruction. For now, I treat 'H' as a syntactically presented operator that asserts positiveness and homogeneity.

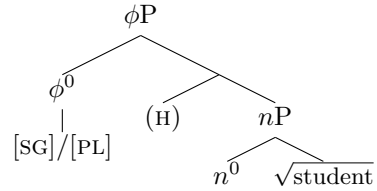# 4   Distributing the 'conjunction-admitting' reading

As argued in section 2, uniqueness effects in SG-marked and NUM-modified questions show that the Q-domain of these question do not contain Boolean conjunctions and further that these questions do not have HO-readings.[4] Interestingly, unlike NUM-modifiers, PP-modifiers such as *in a group of two* do not trigger uniqueness and do not block HO-readings. This contrast suggests that the availability of HO-reading is sensitive to the internal structure of the *wh*-complement.

(18)   (*w*: $a + b$ presented a paper together, and $c + d$ presented a paper together.)

    a.   # I know which **two students** presented a paper together.

    b.   I know which **students in a group of two** presented a paper together.

To account for the above distributional constraints, I propose that the H-shifter is applied locally to the root nP within the *wh*-complement. In the following, I argue that the application of H is blocked in SG-marked nouns and NUM-modified nouns due to conflicts in meaning and types.

First, I assume the structure in (19) for a singular/plural bare noun. The denotation of the root nP is a semi-lattice structure (Harbour 2014). The number feature [SG]/[PL] is evaluated at ϕ^0. Following Sauerland (2003), I interpret [SG] as an atomicity modifier while treating [PL] semantically vacuous. This analysis explains why SG-marked WHPs do not support HO-readings: applying H to nP returns a set of GQs, which are all non-atomic and are conflicting with the atomicity requirement of [SG]. In contrast, H-shifter can be freely used in simple PL-marked and number-neutral *wh*Ps because in those cases the ϕ^0 does not presuppose atomicity.

(19)   *student* (consider three students *abc*)

    a. $[\![\text{SG}]\!] = \lambda P_{\langle e,t \rangle} \lambda x_e.\text{ATOM}(x) \wedge P(x)$

    b. $[\![n\text{P}]\!] = \{a, b, c, a \oplus b, a \oplus c, b \oplus c, a \oplus b \oplus c\}$

    c. $[\]\!] = \{a, b, c\}$
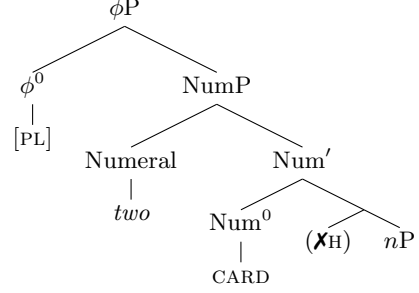
    d. $[\)]\!] = \emptyset$

Next, for NUM-modified NPs, I place cardinal NUM-modifiers at [Spec, NumP] and assume that Num^0 is occupied by a cardinality predicate CARD (Scontras 2014). As defined in (20a), CARD takes a predicate P and a numeral n and returns the set of individuals in P each of which is constituted of exactly-n atoms. These assumptions automatically explain why the H-shifter cannot be used in a NUM-modified NP: the CARD-predicate at Num^0 requires to check the cardinality of the elements in the set it combines with and hence it cannot combine with a set of GQs. In contrast to NUM-modifiers, PP-modifiers are adjoined to the entire NP/ϕP. Hence, the H-shifter can be used in a PP-modified NP without causing a type-mismatch.

---

[4]Note that a SG-marked phrase can be number-neutral in semantics (e.g., Spanish *quién* 'who.SG'). For details see Maldonado (2017), Elliot et. al. (2018), and Alonso-Ovalle & Rouillard (2019).

(20) *two students*

  a. $[\![\text{CARD}]\!] = \lambda P \lambda n \lambda x. P(x) \wedge |x| = n$

  b. Without the H-shifter
      $[\![\text{Num}']\!] = \lambda n \lambda x.{}^*\text{stdt}(x) \wedge |x| = n$
      $[\![\text{NumP}]\!] = \lambda x.{}^*\text{stdt}(x) \wedge |x| = 2$
      $\qquad = \{a \oplus b, b \oplus c, a \oplus c\}$

  c. With the H-shifter:
      Num$'$ suffers type-mismatch

(tree diagram: ϕP branching into ϕ^0 [PL] and NumP; NumP branches into Numeral (two) and Num$'$; Num$'$ branches into Num^0 (CARD), ($\boldsymbol{X}$H), and nP)

5 Deriving the 'disjunction-only' reading

Surprisingly, in responding to □-questions with a SG-marked or NUM-modified WHP, narrow scope disjunctions are not as bad as conjunctions, as seen in (21).

(21) "Which textbook should we read for this class?"
      "*Heim&Kratzer* or/#and *Meaning&Grammar*, (the choice is up to you.)"   (□ ≫ *or*)

A similar contrast is observed with questions with a possibility modal ("◇-questions" henceforth). As in (22), multi-choice answers to ◇-questions can be equivalently expressed with an elided conjunction or an elided disjunction. In particular, the disjunction is read as free choice.

(22) "What can I use (as a textbook) for this class?"
      "*Heim&Kratzer* or/and *Meaning&Grammar*, (the choice is up to you.)"

For SG-marked ◇-questions, it was commonly thought that these questions are subject to uniqueness and hence cannot have multi-choice readings. Hirsch & Schwarz (2019) finds out that, however, multi-choice readings are available as long as the uniqueness inference evoked by the SG-marked WHP takes narrow scope, as seen in (23).

(23) "Which letter could be missing in *fo_m*?"
      "The missing letter could be *a*, and the missing letter could be *r*."

In Hirsch & Schwarz's example (23), the multi-choice answer can be expressed as a conjunction over two full sentences. Interestingly, as seen in (24), if elided, multi-choice answers to SG-marked ◇-questions must be disjunctions. NUM-modified ◇-questions have the same behavior.[5]

(24) a. "Which textbook can I use for this class?" "*H&K* or/ #and *M&G*."
      b. "Which letter could be missing in *fo_m*?" "Letter *a* or/ #and letter *r*."

Here arise two puzzles: first, why these questions admit and only admit disjunctions? second, why such a 'disjunction-only' HO-reading is available despite that the WHP is SG-marked or

---

[5]Gentile & Schwarz (2018) make a similar observation with *how many*-questions. *How many*-questions presuppose uniqueness: (i-a) cannot be felicitously responded by a multi-choice answer expressed by a conjunction of two cardinal numerals. Given that the predicate *solved this problem together* is stubbornly collective, Gentile & Schwarz conjecture that the Q-domain of this question does not include Boolean conjunctions over numerals. Further, Gentile & Schwarz observe that ◇-modals can obviate violations of uniqueness in *how many*-questions: (i-b) admits multi-choice answers like "Two is ok and three is ok" and does not presuppose uniqueness. In analogy to (24), I add to (i-b) that the multi-choice answer cannot be expressed by an elided conjunction.

(i) a. "How many students solved this problem together?" "#Two and three."
     b. "How many students are allowed to solve this problem together?" "Two or/#and three."

NUM-modified? The following presents two analyses to these puzzles. One approach treats the 'disjunction-only' reading the very same reading as the regular HO-reading but gives a weaker semantics to singular and numeral-modified nouns. In this approach, the distributional difference between conjunctive and disjunctive answers comes from that atomicity and cardinality restrictors remove Boolean conjunctions but not disjunctions. The other approach assumes that the derivation of this reading involves reconstructing the WH-complement to the question nucleus and interpreting uniqueness locally. In this approach, conjunctive answers are unacceptable because applying conjunction directly over uniqueness yields a contradiction.

Solution 1: A uniform approach This approach treats the 'disjunction-only' reading the very same reading as the 'conjunction-admitting' reading.[6] To derive these two readings uniformly, all we need is to allow some of the Boolean disjunctions to be atomic and/or cardinal, just like entities. In the following definitions, the (a)-condition on minimal witness sets ensures the atomic/cardinal GQ to be a disjunction, an $\exists$-quantifier, or a Montagovian individual.[7]

(25)   A GQ π is atomic iff (a) the minimal witness sets of π are all singleton sets, and (b) every member in the smallest live-on set of π is atomic.

(26)   A GQ π has the cardinality n iff (a) the minimal witness sets of π are all singleton sets, and (b) every member in the smallest live-on set of π has the cardinality n.

With the above assumptions, I re-define the [SG] feature and the CARD predicate polymorphically as in (27) and (28). 'MWS(A, x)' is read as 'A is a minimal witness set of x'.

(27) $[\![\text{SG}]\!] = \lambda P \lambda x. \begin{cases} P(x) \wedge \text{ATOM}(x) & \text{if } P \subseteq D_e \\ P(x) \wedge \forall A[\text{MWS}(A, x) \to |A| = 1] \wedge \forall y \in \text{SMLO}(x)[\text{ATOM}(y)] & \text{if } P \subseteq D_{\langle et, t \rangle} \end{cases}$

(28) $[\![\text{CARD}]\!] = \lambda P \lambda n \lambda x. \begin{cases} P(x) \wedge |x| = n & \text{if } P \subseteq D_e \\ P(x) \wedge \forall A[\text{MWS}(A, x) \to |A| = 1] \wedge \forall y \in \text{SMLO}(x)[|y| = n] & \text{if } P \subseteq D_{\langle et, t \rangle} \end{cases}$

With the new definitions, the H-shifter can be used in singular and NUM-modified nouns. With three students abc, *student* and *two students* are interpreted as in (a) for FO-readings and as in (b) for HO-readings. The NP structures are the same as in (19)/(20).

(29)   *student*:
   a.   for FO: $[\]]\!] = \{a, b, c\}$
   b.   for HO: $[\)]]\!] = \{\bigsqcup A \mid A \subseteq \{x^{\Uparrow} \mid x \in \{a, b, c\}\}$

(30)   *two students*: we have $[\]]\!] = [\![\text{NumP}]\!]$, where ...
   a.   for FO: $[\![[_{\text{NumP}} \, two \, \text{CARD} \, (n\text{P})]]\!] = \{a \oplus b, b \oplus c, a \oplus c\}$
   b.   for HO: $[\![[_{\text{NumP}} \, two \, \text{CARD} \, \text{H}(n\text{P})]]\!] = \{\bigsqcup A \mid A \subseteq \{x^{\Uparrow} \mid x \in \{a \oplus b, b \oplus c, a \oplus c\}\}$

In sum, in the uniform approach, the disjunction-only reading (for a SG-marked question) and the conjunction-admitting reading (for a number-neutral or PL-marked question) are derived uniformly as follows:

(31)   Which book(s) does John have to read?
   [which-$^{\text{H}}$book(s) $\lambda \pi$ [have-to [$\pi_{\langle et, t \rangle}$ λx [John read x_e]]]]

---

[6]I thank Manuel Križ for pointing out the direction of this analysis.

[7]Witness sets are defined in terms of the living-on property as follows: if a GQ π lives on a set B, then A is a witness set of π iff $A \subseteq B$ and $\pi(A)$. For example, given a discourse domain including three students abc, the $\forall$-quantifier *every student* has a unique minimal witness set $\{a, b, c\}$, while the singular $\exists$-quantifier *some student* has three minimal witness sets $\{a\}$, $\{b\}$, and $\{c\}$, each of which consists of one atomic student.

Solution 2: A reconstruction approach This approach makes use of syntactic reconstruction. As in (32a-b), the *wh*-complement *book* is reconstructed back to the nucleus. This reconstruction has two consequences. First, it leaves a semantically unmarked variable D as the restrictor of the whP, which can be type-lifted by the h-shifter without causing a type-mismatch or a conflict w.r.t. atomicity. Thus, a HO-reading arises if the h-shifter is applied to the D variable and if *wh*-phrase binds a HO-trace, as in (32b). Second, uniqueness is evaluated at the scopal position where the reconstructed noun adjoins to. This prediction explains why a question in this reading rejects conjunctive answers: combining a conjunction with a predicate of uniqueness yields a contradiction, as seen in (33).

(32) "Which book does John have to read?" "Book A or/#and Book B."

    a. [which$_D$ λx [have-to [x is the book John read]]] $(\Box \gg \iota)$
       $[\![Q_{\text{FO}}]\!] = \lambda x_e : x \in D.\Box\lambda w[x = \iota y[\mathsf{book}_w(y) \wedge \mathsf{read}_w(y)]]$

    b. [which$_{\text{H}D}$ $\lambda\pi$ [have-to [$\pi_{\langle et,t\rangle}$ $\lambda x.$ x_e is the book John read]]] $(\Box \gg \pi \gg \iota)$
       $[\![Q_{\text{HO}}]\!] = \lambda\pi_{\langle et,t\rangle} : \pi \in {}^{\text{H}}D.\Box\lambda w[\pi(\lambda x_e.x = \iota y[\mathsf{book}_w(y) \wedge \mathsf{read}_w(j,y)])]$

(33) $[\![Q_{\text{HO}}]\!](a^{\Uparrow} \cap b^{\Uparrow}) = \Box\lambda w[a = \iota y[\mathsf{book}_w(y) \wedge \mathsf{read}_w(j,y)] \wedge [b = \iota y[\mathsf{book}_w(y) \wedge \mathsf{read}_w(j,y)]]$
    (#It has to be the case that the book that John read is Book A and is Book B.)

This analysis also extends to $\Diamond$-questions. In Xiang (2016: ch. 2), I argue that the mention-all (MA-)reading expecting a conjunctive answer and the reading expecting a disjunctive answer are derived with distinct LF structures, as shown in (34) [with simplification]. In particular, conjunctive MA arises if the HO-WH-trace π scopes over the $\Diamond$-modal, making the conjunctive MA answer the strongest true answer. In contrast, disjunctive MA arises if the trace π (regardless of its scope) is associated with a free choice licensing operator DOU ($\approx$ Mandarin FC-licensing particle *dou*, see details in Xiang (To appear)). Roughly, applying DOU to a disjunctive answer yields a free choice interpretation, making the disjunctive MA answer the strongest true answer.

(34) "What can we use for this class?" "Book A or/and Book B."

    a. Conjunctive MA $(\pi \gg \Diamond)$
       [what $\lambda\pi_{\langle et,t\rangle}$ [π λx_e [can [we use x]]]] $[\![Q]\!] = \lambda\pi_{\langle et,t\rangle}.\pi(\lambda x_e.\Diamond f(x))$
    b. Disjunctive MA $(\text{DOU} \gg \Diamond \gg \pi)$
       [what $\lambda\pi_{\langle et,t\rangle}$ [DOU [can π λx_e [we use x]]]] $[\![Q]\!] = \lambda\pi_{\langle et,t\rangle}.\text{DOU}\Diamond\pi(\lambda x_e.f(x))$

Next, consider the SG-marked $\Diamond$-question in (35). Reconstructing the singular noun *book* and letting the trace π take scope above the $\Diamond$-modal yield the following scopal pattern: $\pi \gg \iota \gg \Diamond$. As shown in (35a), unless A and B are the same book, combining the derived Q-function with the Boolean conjunction $a^{\Uparrow} \cap b^{\Uparrow}$ yields a contradiction. In contrast, as seen in (35b), if π takes a narrow scope relative to the $\Diamond$-modal and is associated with DOU across the modal, the derived free choice inference is not contradictory and expresses the desired MA-answer.

(35) "Which book can we use for this class?" "Book A or/#and Book B."

    a. If $\pi \gg \iota \gg \Diamond$:
       $[\![Q]\!] = \lambda\pi_{\langle et,t\rangle} : \pi \in {}^{\text{H}}D.\lambda w[\pi(\lambda x_e.x = \iota y[\mathsf{book}_w(y) \wedge \Diamond_w f(y)])]$
       $[\![Q]\!](a^{\Uparrow} \cap b^{\Uparrow}) = \lambda w[a = \iota y[\mathsf{book}_w(y) \wedge \Diamond_w f(y)] \wedge [b = \iota y[\mathsf{book}_w(y) \wedge \Diamond_w f(y)]]$
       (#a is the unique book that we can use, and b is the unique book that we can use.)

    b. If $\text{DOU} \gg \Diamond \gg \pi \gg \iota$ (viz., FC $\gg \Diamond \gg \iota$):
       $[\![Q]\!] = \lambda\pi_{\langle et,t\rangle} : \pi \in {}^{\text{H}}D.\text{DOU}\Diamond[\lambda w.\pi(\lambda x_e.x = \iota y[\mathsf{book}_w(y) \wedge f_w(y)])]$
       $[\![Q]\!](a^{\Uparrow} \cup b^{\Uparrow}) = \text{DOU}\Diamond[\lambda w.(a^{\Uparrow} \cup b^{\Uparrow})(\lambda x_e.x = \iota y[\mathsf{book}_w(y) \wedge f_w(y)])]$
       $= \Diamond\lambda w[a = \iota y[\mathsf{book}_w(y) \wedge f_w(y)]] \cap \Diamond\lambda w[b = \iota y[\mathsf{book}_w(y) \wedge f_w(y)]]$
       (a can be the unique book that we use, and b can be the unique book that we use.)

6   Conclusions

This paper investigates the higher-order (HO)-readings of *wh*-questions. First, drawing on evidence from questions with necessity modals or collective predicates, I showed that sometimes a *wh*-question can only be completely addressed by a GQ and must be interpreted with a HO-reading. Next, I argued that the GQs that can serve as complete answers of questions are subject to two constraints — positiveness and homogeneity. Incorporating these constraints into the meaning of a H-shifter, I proposed that HO-reading arises if the H-shifter converts the *wh*-restrictor into a set of HO-meanings and if the *wh*-phrase binds a HO-trace. Accordingly, HO-readings are unavailable if the application of the H-shifter is blocked, either by the atomicity constraint of [SG] in singular nouns, or by the cardinality constraint in NUM-modified nouns.

Further, I observed that SG-marked and NUM-modified questions admit only disjunctive answers. I provided two explanations to this distribution. In the uniform account, these questions admit disjunctions because disjunctions (but not conjunctions) may satisfy the atomicity/cardinality requirement. In the reconstruction account, the WH-complement is reconstructed, which gives rise to local uniqueness and yields contradictions for conjunctive answers.

References

[1] L. Alonso-Ovalle and V. Rouillard. Number inflection, spanish bare interrogatives, and higher-order quantification. In *Proceedings of NELS 49*, 2019.

[2] J. Barwise and R. Cooper. Generalized quantifiers and natural language. *Linguistics and Philosophy*, 4(2):159–219, 1981.

[3] D. Cresti. Extraction and reconstruction. *Natural Language Semantics*, 3(1):79–122, 1995.

[4] P. D. Elliott, A. C. Nicolae, and U. Sauerland. Who and what do *who* and *what* range over cross-linguistically, 2018.

[5] F. Gentile and B. Schwarz. A uniqueness puzzle: *how many*-questions and non-distributive predication. In *Proceedings of Sinn und Bedeutung 21*, 2018.

[6] J. Groenendijk and M. Stokhof. Type-shifting rules and the semantics of interrogatives. In *Properties, types and meaning*, pages 21–68. Springer, 1989.

[7] D. Harbour. Paucity, abundance, and the theory of number. *Language*, 90(1):185–229, 2014.

[8] A. Hirsch and B. Schwarz. Singular which, mention-some, and variable scope uniqueness. Poster presented at Semantics and Linguistic Theory 29, 2019.

[9] G. Link. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In C. S. Rainer Bäuerle and A. von Stechow, editors, *Meaning, use, and interpretation of language*, pages 302–323. De Gruyter, 1983.

[10] M. Maldonado. Plural marking and d-linking in Spanish interrogatives, 2017.

[11] U. Sauerland. A new semantics for number. In *Proceedings of SALT 13*, 2003.

[12] G. Scontras. *The semantics of measurement*. PhD thesis, Harvard University, 2014.

[13] B. Spector. Modalized questions and exhaustivity. In *Proceedings of SALT 17*, 2007.

[14] B. Spector. An unnoticed reading for wh-questions: Elided answers and weak islands. *Linguistic Inquiry*, 39(4):677–686, 2008.

[15] A. Szabolcsi. Background notions in lattice theory and generalized quantifiers. In A. Szabolcsi, editor, *Ways of Scope Taking*, pages 1–27. Springer Netherlands, Dordrecht, 1997.

[16] Y. Xiang. *Interpreting questions with non-exhaustive answers*. PhD thesis, Harvard University Cambridge, Massachusetts, 2016.

[17] Y. Xiang. Function alternations of the Mandarin particle *dou*: Distributor, free choice licensor, and 'even'. *Journal of Semantics*, To appear.