Cognitive Psychology 144 (2023) 101584

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

Keeping quantifier meaning in mind: Connecting semantics, cognition, and pragmatics

Tyler Knowlton [a,*], John Trueswell [b], Anna Papafragou [c]

[a] MindCORE, University of Pennsylvania, 3740 Hamilton Walk, Philadelphia, PA 19104, USA
[b] Department of Psychology, University of Pennsylvania, 425 South University Ave, Philadelphia, PA 19104, USA
[c] Department of Linguistics, University of Pennsylvania, 3401 Walnut Street, Philadelphia, PA 19104, USA

ARTICLE INFO

Keywords:
Meaning
Pragmatics
Quantification
Psychosemantics

ABSTRACT

A complete theory of the meaning of linguistic expressions needs to explain their semantic properties, their links to non-linguistic cognition, and their use in communication. Even though in principle interconnected, these areas are generally not pursued in tandem. We present a novel take on the semantics-cognition-pragmatics interface. We propose that formal semantic differences in expressions' meanings lead those meanings to activate distinct cognitive systems, which in turn have downstream effects on when speakers prefer to use those expressions. As a case study, we focus on the quantifiers "each" and "every", which can be used to talk about the same state of the world, but have been argued to differ in meaning. In particular, we adopt a mentalistic proposal about these quantifiers on which "each" has a purely individualistic meaning that interfaces with the psychological system for representing object-files, whereas "every" has a meaning that implicates a group and interfaces with the psychological system for representing ensembles. In seven experiments, we demonstrate that this account correctly predicts both known and newly-observed constraints on how "each" and "every" are pragmatically used. More generally, this integrated approach to semantics, cognition, and pragmatics suggests that canonical patterns of language use can be affected in predictable ways by fine-grained differences in semantic meanings and the cognitive systems to which those meanings connect.

1. Introduction

What are linguistic meanings? An influential tradition in the study of semantics maintains that natural language expressions are names for things in speakers' environments (e.g., Davidson, 1967; Montague, 1973; Lewis, 1975; Heim & Kratzer, 1998). On this view, the meaning of a common noun like "frog" is essentially taken to be the set of all frogs. Likewise, the meaning of a predicate like "is green" is taken to be the set of all green things. Logical words like "every" are then taken to name relations between two independent sets (Barwise & Cooper, 1981). So the expression "every frog is green" is a label for a possible state of the world: the state that obtains when *the frogs are a subset of the green things*. That is, the meaning of the expression consists of the conditions under which it is true. This traditional view of meaning is often described as 'mind-external', because the semantic theory is based on facts pertaining to the external world rather than mental representations of it.

Situating meanings in the mind-external world in this way has afforded theorists a high degree of precision in stating hypotheses about meanings. As noted above, if the meaning of the expression "every frog is green" consists of the conditions under which that

* Corresponding author.
E-mail address: tzknowlt@upenn.edu (T. Knowlton).

https://doi.org/10.1016/j.cogpsych.2023.101584
Received 10 August 2022; Received in revised form 20 June 2023; Accepted 26 June 2023
Available online 3 July 2023

sentence is true of the world, then the meaning can be reasonably regimented in precise (e.g., set-theoretic) terms, such as '$\textsc{true}$ if and only if {x: Frog(x)}⊆{x: Green(x)}'. But this focus on the mind-external world comes at the cost of putting semantics at odds with the rest of linguistics, which is often taken to be a branch of cognitive science. Indeed, in contrast to the stated goals of leading theories in natural language syntax (e.g., Chomsky, 1964; 1965; Boeckx & Hornstein, 2010), much of the foundational work in the study of linguistic semantics has taken a decidedly 'anti-psychologistic' viewpoint (Pietroski, 2017; see Partee, 1979, 1995b, 2011, 2018 for helpful endeavors to wrestle with this disconnect). The following question exemplifies the issue: when a speaker of English understands the expression "every frog is green", do they literally represent and relate two sets, the set of frogs and the set of green things? The mind-external view, on its own, makes no predictions here. The meaning of the expression connects to the world directly; how people might mentally represent that connection is taken to be outside of the semantic theorist's job description.

In contrast, mentalistic views of meaning take answering the question of representation to be a primary goal for a semantic theory. As such, these views embrace psychology and treat linguistic meanings as tools for connecting expressions to non-linguistic thought. One version of mentalistic semantics – which we adopt here – aims for an integrative approach: retain the formal precision of mind-external theories of meaning but treat particular formal specifications as hypothesized psychological objects (e.g., Jackendoff, 1983; Papafragou, 2000; Carston, 2002; Pietroski, Lidz, Hunter, & Halberda, 2009; Lidz, Pietroski, Halberda, & Hunter, 2011; Pietroski, 2018; Wellwood, 2020; Knowlton et al., 2021).[1] On this sort of view, the meanings of lexical items like "frog" and "green" are instructions to access the relevant concepts. Combining those concepts with the concept named by "every" yields a complex concept with a particular formal structure. This complex concept can be thought of as a 'Language of Thought' representation (Fodor, 1975) or a psychologized version of the linguists' notion of a 'logical form'; a 'psycho-logical form' (Hornstein, 1984; Knowlton, 2021; Soames, 1987). The resulting representation is a thought, which may (but need not) be truth-evaluable (Pietroski, 2005). In this way, the semantic theory can exhibit the virtues of the mind-external approach (e.g., formal precision; a focus on how the meaning of the whole expression is composed out of the meaning of its parts; a connection to truth-conditions) while still being a genuine psychological theory.

Abstracting away from many of the details, then, the resulting picture is two vastly different approaches to linguistic meaning. The mind-external approach, stemming from logic and the philosophy of language, maintains that linguistic expressions are directly connected to the world. This approach remains agnostic about any connection expressions bear to the mental lives of speakers. In contrast, the mentalistic approach places the focus on how expressions connect to speakers' minds (and maintains that any connections to the mind-external world are mediated by mental representations). As such, the mentalistic view is better suited to handle two 'interface' issues: how semantic representations make contact with pragmatics (the way meaning is used and understood in context) and how they make contact with non-linguistic cognition.[2] But few if any studies have demonstrated how the mentalistic view can successfully handle both interface issues *for a single case*.

Here, we aim to present a unified mentalistic approach to semantics, pragmatics, and cognition that retains the formal precision of the mind-external approach to semantics. We do so by way of a case study: the meanings of the quantifiers "each" and "every". It has recently been proposed that formal differences between the meanings of these seemingly similar quantifiers lead them to differentially interface with two systems of non-linguistic representation during sentence verification. In particular, Knowlton (2021) and Knowlton, Pietroski, Halberda, and Lidz (2022) report that when asked to evaluate sentences like "each circle is green", participants treat the relevant circles as a series of independent object-file representations, but when shown the same picture and asked to evaluate whether "every circle is green", participants represent the circles as an ensemble collection. They argue that these results stem from a formal difference between the semantic representations of "each" and "every". Here, we propose that the same difference in meaning and resulting connection to non-linguistic cognition has consequences beyond sentence verification. Namely, we argue that the hypothesized mentalistic meanings and consequent links to non-linguistic cognitive systems for object and group representation have downstream effects on how "each" and "every" are typically used.

This case study demonstrates how thinking about meanings as mental representations that serve as instructions to cognition – as opposed to names for mind-external referents – paves the way for connecting semantics and pragmatics to cognitive representations in a psychologically plausible fashion. Composing linguistic meanings, on this view, can be thought of eventuating in complex concepts whose formal properties play a role in triggering different non-linguistic systems, which themselves have repercussions for conditions of use. This melding of semantics, cognition, and pragmatics not only represents an example of how the study of linguistic meaning can be fruitfully brought under the purview of cognitive science, but also offers a new tool for integrating and explaining an array of previously disconnected phenomena along distinct dimensions of meaning. Moreover, it offers benefits for each of the three levels: studying the semantics-cognition-pragmatics connection in this way has implications for what the semantic representations look like, suggests new directions for studying the related non-linguistic cognitive systems, and adds a new tool to the pragmatists' arsenal.

Below, Section 1.1 provides a brief introduction to the "each" versus "every" case study. Section 1.2 then introduces a mentalistic

[1] An alternative version of mentalistic semantics, situated within the tradition of 'cognitive linguistics' (e.g., Fauconnier, 1984; Lakoff, 1987; Langacker, 1987; Talmy, 2000), more radically distances itself from the mind-external approach discussed above. In addition to a focus on details of human psychology, this 'cognitive' approach largely jettisons the precise formalisms proffered by the mind-external view. As Talmy (2019) puts it: "most cognitive linguists share the sensibility that such formal representations poorly accord with the gradients, partial overlaps, interactions that lead to mutual modification, processes of fleshing out, inbuilt forms of vagueness, and the like that they observe" (p.3).

[2] We use 'semantic representation' here to mean *mental representation that serves as an expression's meaning*. This differs from standard use of the term in semantics, where it usually means *a representation of the meaning, distinct from the meaning itself*. On this more standard usage, the meaning of an expression is taken to be a mind-external object. So a 'semantic representation' is a representation of that meaning, either in the mind of a speaker when they token the expression or in the mind of a theorist when they theorize about it. See Williams (2015) for careful discussion of this issue.

proposal about their meanings. On this proposal, the concepts that serve as the meanings of these expressions formally differ in a way that causes them to (non-deterministically) interface with particular non-linguistic representational systems. "Each" is argued to bias speakers to represent the domain of quantification as a series of object-files, individual indices with associated properties. "Every", in contrast, is argued to bias speakers to represent the domain of quantification as an ensemble, a summary representation of a collection. In Section 2, the properties of these different representational systems are argued to have downstream consequences on the expressions' contexts of use. Being more likely to treat the domain as a series of object-files makes "each" preferred for quantifying over small, local domains, whereas being more likely to treat the domain as an ensemble makes "every" better for larger domains, as well as for generalizing beyond the initially-established one. These predictions are then confirmed with the seven experiments reported in Section 3. Section 4 concludes by comparing the present approach to the standard, mind-external approach and other alternatives and discussing how the three-pronged approach presented here could be generalized beyond the domain of quantification.

1.1. *"Each" and "every"*

The English universal quantifiers "each" and "every" can often be used to describe the very same state of the world. It would be appropriate to say "each frog is green" or "every frog is green" upon encountering four green frogs, for example.[3] But it has long been observed that these quantifiers nonetheless differ in subtle ways. Vendler (1962) asks readers to imagine gesturing toward a basket of apples and either saying "take every one of them" or "take each one of them". Whereas the "every"-version sounds natural, the "each"-version, he notes, sounds incomplete: something like "take each one of them and examine it in turn" would be expected. These sorts of subtle contrasts suggested to Vendler (and many authors since) that "each" is in some sense *more individualistic* than "every". We can see the same point clearly in an example like (1), from Knowlton, Trueswell, and Papafragou (2022).

(1) a. Each martini needs an olive.
b. Every martini needs an olive.

The sentence in (1a) calls to mind a scene in which a few martinis were made but lack garnishes. One could imagine a bartender saying it to remind their assistant that the drinks are not yet ready to serve. On the other hand, (1b) is more naturally understood as a general claim, meant to apply to vast quantities of drinks. One could imagine encountering it as part of a martini recipe. And some cocktail purists might even offer (1b) as a kind-claim: a drink cannot be called a martini if it lacks an olive.

Another contrast can be observed in (2). As Beghelli (1997) pointed out, (2a) can be felicitously answered by mentioning, for each student, which book they were loaned: "I loaned *Frankenstein* to Frank, *Persuasion* to Paula, and *Dune* to Dani". But (2b) seems to demand an answer about the list as a whole: "There's no one book that I loaned to every student, I loaned a different book to each one".

(2) a. Which book did you loan to each student?
b. Which book did you loan to every student?

There are many other differences between these expressions (for reviews, see e.g., Vendler, 1962; Beghelli & Stowell, 1997; Knowlton, Halberda, Pietroski & Lidz under review). Some of the noted differences are syntactic ones, for example, that "each" can be used as an adverb (e.g., "the students each received a book") whereas "every" cannot, and that "every" requires support from "one" when used in the partitive (e.g., "every one of the students left") whereas "each" does not (e.g., "each of the students left"). But one suspects that at least some of the observed differences are illustrative of a difference in meaning between "each" and "every".

1.2. *A mentalistic proposal*

In part motivated by some of the apparent semantic differences between "each" and "every", Knowlton (2021) and Knowlton et al. (2022) offer a mentalistic proposal about these quantifiers' meanings. The proposal has two parts: first, the idea that "each" and "every" have formally distinct concepts of universal quantification – distinct semantic representations – as their meanings, and second, the idea that these distinct meanings naturally invite (but do not entail) the use of distinct cognitive systems.

Starting with the formal distinction, the leading idea is that a sentence like "each frog is green" has a semantic representation that implicates only individuals and avoids any notion of grouping them (*any thing that's a frog is such that it is green*). That is, the concept named by "each" calls for individuating the things quantified over and attributing some property to each one of those individuals (much like a series of conjunctions: *$frog_1$ is green & $frog_2$ is green & …*). On the other hand, a sentence like "every frog is green" has a semantic representation that implicates a single group and attributes a property exhaustively to its members (*the frogs are such that they are each green*). In other words, the concept named by "every" calls for grouping the things quantified over and distributively applying a

[3] For the time being, we set aside the other main English universal quantifier, "all". Our reason for this choice is that "each" and "every" are more similar to each other than either are to "all". For one thing, both require singular agreement ("each/every frog" vs. "all frogs"). For another, both are often described in the literature as being 'distributive' to the exclusion of "all". Moreover, some work in linguistics suggests that "all" is not even a genuine quantifier, but an intensifier of sorts (e.g., Baker, 1995; Partee, 1995a; Brisson, 1998). Lastly, "all" is orders of magnitude more frequent in speech (including child-directed speech; see Knowlton & Lidz, 2021) and its uses more varied. For these reasons, "each" and "every" form a more compelling minimal pair.

predicate to each one of them (as if "every frog" abbreviated "the frogs each"). These semantic representations can be characterized formally as in (3a) and (3b).

(3) a. ∀x:Frog(x)[Green(x)] "each frog is green"
≈*any thing$_x$ that is a frog is green*
b. TheX:Frog(X)[∀x:X(x)[Green(x)]] "every frog is green"
≈*the frogs$_X$ are such that any thing$_x$ that is one of them$_X$ is green*

The difference between (3a) and (3b) is that the former has no constituent corresponding to the plurality of frogs, whereas the latter does (which we indicate here with 'TheX:Frog(X)'). More formally, (3a) only uses the tools of first-order logic whereas (3b) also makes use of second-order logic: (3a) only exhibits quantification into lowercase variable positions, like 'x', which can stand for one thing at a time, whereas (3b) also makes use of quantification into an uppercase variable position, 'X', which can stand for multiple things at once.[4] To reiterate then, the idea is that both meanings are distributive – they apply a predicate to each member of the domain of quantification – but only "every" calls for first grouping its domain. In both (3a) and (3b) the domain of quantification consists of individual frogs, but in (3b) those individuals are initially grouped into the plurality "the frogs". To be clear, the idea is not that "every frog" in any sense quantifies over multiple sets of frogs. Rather, "every frog" and "each frog" merely call for different ways of treating the same domain.

Turning to the distinction between related cognitive systems, the thought is that the formal difference described above leads these two meanings to more naturally connect to two different non-linguistic cognitive systems during sentence verification. That is, the formal properties of these expressions' semantic representations have an effect on how speakers will choose to 'reach out' into the world to determine whether a sentence like "each/every frog is green" is true. The proposal is not that linguistic meanings are verification strategies, nor that meanings are grounded in or built out of non-linguistic cognitive systems. Instead, the idea is that the meaning representations are separate from but readable by non-linguistic systems (for discussion on this point, see Lidz et al., 2011).

In particular, the individualistic "each" meaning is argued to naturally interface with the cognitive system for representing object-files (e.g., Kahneman & Treisman, 1984; Pylyshyn & Storm, 1988; Kahneman, Treisman, & Gibbs, 1992; Xu & Carey, 1996; Pylyshyn, 2001; Xu & Chun, 2009; Green & Quilty-Dunn, 2021). An object-file is essentially a pointer to an object to which properties like size and color are bound. The group-implicating "every" is instead argued to interface with the cognitive system for representing ensembles (e.g., Ariely, 2001; Chong & Treisman, 2003; Alvarez & Oliva, 2008; Alvarez, 2011; Haberman & Whitney, 2012; Sweeny, Wurnitsch, Gopnik, & Whitney, 2015; Ward, Bear, & Scholl, 2016; Whitney & Yamanashi Leib, 2018). Unlike object-files, ensembles abstract away from individual properties and encode the collection in terms of summary statistics like the group's average hue and average size.

These two cognitive systems are operative in humans as early as infancy (for reviews, see Feigenson, Dehaene, & Spelke, 2004; Carey, 2009). And while much work on object-files and ensembles has focused on the visual modality, we assume that these representations are more general throughout cognition. In addition to visually-presented objects, these systems can be used to represent visual events (Wood & Spelke, 2005), auditory beeps (Izard, Sann, Spelke, & Streri, 2009; Kanjlia, Feigenson, & Bedny, 2021), and touches on the skin (Plaisier, Tiest, & Kappers, 2009; Riggs et al., 2006). Going beyond multi-modality, our own results suggest an even more general notion of these two distinct representational systems. As discussed below, two of the experiments presented in Section 3 involve domains that are not perceived at all but still show effects of object-files and ensembles.

The proposed meanings of "each" and "every" along with the proposed connection to distinct non-linguistic systems (object-files and ensembles) are argued to explain some, though not all, of the previously-observed semantic differences between these quantifiers.[5] But the main evidence for the above proposal comes from experiments probing what people remember after evaluating sentences. Knowlton et al. (2022) report that participants recall cardinality – an ensemble summary statistic – better after evaluating sentences like "every circle is green" than sentences like "each circle is green", despite being shown the same images. Knowlton (2021) likewise asked children between 5 and 8 years of age whether "each/every circle is blue" and showed them an image on a tablet. After answering, the image disappeared, and participants were asked to touch where the center of the circles had been. Since center of mass is another ensemble summary statistic, those who had initially been asked the question with "every" gave more accurate responses. Knowlton, Halberda, Pietroski & Lidz (under review) test the opposite prediction: evaluating an "each" sentence should make participants better at recalling individual properties. Indeed, they report that participants who evaluated "each circle is green" were better able to succeed at a change detection task in which one circle's hue changed compared to participants shown the same three circles but given a sentence like "every circle is green".

These sentence-verification experiments provide participants with visual representations of things quantified over and show that

[4] On some views (e.g., Boolos, 1984), first-order and second-order variables (e.g., 'x' and 'X' in (3a-b)) range over the same domain and the difference between them is that the former can 'point to' one value at a time whereas the latter can 'point to' multiple values at once. Another approach treats first-order and second-order variables as ranging over separate domains, like individuals versus sets of individuals (e.g., Frege 1879; 1893). See Schein (1993), Boolos (1998), and Pietroski (2003) for arguments against the 'different domains' view. For our purposes, 'TheX:Frog(X)' in (3b) is intended to be understood as "The Frogs". But depending on the perspective one takes on second-order variables, this could be spelled out in various ways. For example, one might prefer to think of it as an abbreviation for the plural logic expression "the things X such that for each thing x, x is one of them (i.e., one of the Xs) if and only if x is a frog" or as shorthand for "the set of frogs".

[5] To be sure, the proposal outlined above likely does not obviate the need for previously-posited syntactic distinctions (e.g., Beghelli, 1997). However, it has been argued to shift some of the explanatory burden onto properties of non-linguistic cognition (see e.g., Knowlton & Gomes, 2021; Knowlton et al., under review).

the choice of expression leads to different construals of the same scene. Participants group the things quantified over to a greater extent when universal quantification is indicated with "every" than "each". This makes sense if "every" has a meaning that explicitly represents the things quantified over as a group whereas "each" does not. To be clear though, the link between these semantic representations and non-linguistic cognitive systems is not a deterministic one. The finding is not that participants *always* use object-files when encountering "each" and ensembles when encountering "every". After all, the formal properties of the semantic representation constitute just one force pushing participants to deploy a certain cognitive system. And a speaker could very well understand "every frog is green" along the lines of (3b) – they could build the semantic representation in (3b) upon hearing the sentence – but nonetheless opt to individuate the frogs and represent each one as an independent object file. The idea is just that participants will be more likely to group the frogs, and thus deploy the ensemble system, if they are quantified over with "every", since its meaning implicates "the frogs" as a group, whereas the meaning of "each" does not. Put another way, if one understands "each frog is green" along the lines in (3a), this representation won't be as likely to call to mind the idea of grouping the frogs as an ensemble, because it has no constituent corresponding to "the frogs".

If the posited link between linguistic meanings and non-linguistic cognition is the correct explanation of the sentence verification performance described above, this link should be detectable in other ways. Relatedly, if the only observable consequence of the representations in (3) was their effect on sentence verification strategies, one might begin to doubt that these effects are due to different semantic representations in the first place. An added difficulty comes from the fact that the proposed representations in (3) are logically equivalent (i.e., whenever it's true that "every F is G" it will likewise be true that "each F is G"). So evidence for the proposed distinction will not be found via the usual method of eliciting or reflecting on truth-value judgments. Instead, our aim here is to demonstrate that the non-deterministic link between semantic representations and non-linguistic cognition described above can also explain some aspects of pragmatic use. The idea is that the same effects that occur in sentence verification – speakers more often opting to deploy object-files given "each" and ensembles given "every" – will occur in the absence of sentence verification, and, more generally, in the absence of visual/perceptual information about the things quantified over.

2. Current study: Quantifier meaning, cognitive content, and pragmatic use

In this study, we propose to bridge the mentalistic account of quantifiers sketched in the previous section with a pragmatic perspective aiming to explain what information gets conveyed in the use of an expression beyond its 'literal' (semantic) meaning (e.g., Grice, 1967; Gazdar, 1979; Levinson, 1983; Horn, 1984; Sperber & Wilson, 1986; Carston, 1988; 2002; Clark, 1996; Bach, 1997). We adopt an explicitly mentalistic approach to pragmatics on which narrow semantic meaning needs to be augmented with information from context, including listeners' world knowledge and general cognitive abilities (Carston, 2008). Semantics, on this view, offers a mental object that serves as a 'schema or template' for thought-building, whereas pragmatics (in a general sense that includes aspects of human cognition) is then recruited to 'fill in the gaps' inherent in that schema. The end result is a complete thought. In the case of quantifiers, differences in pragmatic contexts of use have been used to argue for particular ways of specifying a quantifier's meaning (e.g., for "most", see Papafragou & Schwarz, 2006; Solt, 2016; Denić & Szymanik, 2022). Our current perspective views quantifier interpretation in context from a novel angle: formal properties of a quantified expression's meaning lead it to routinely interface with particular non-linguistic systems and in turn have downstream consequences on pragmatic contexts of use.

Specifically, we rely on known properties of object-files and ensembles (themselves connected to quantifier semantics in the way described above) to make two novel predictions about the typical contexts of use for "each" and "every". First, ensemble representations are not subject to strict working memory constraints like object-files, so "every" should be preferred when the number of things being quantified over is larger. Second, ensembles support generalizing in a way that object-files do not, so "every" should be preferred when quantification is intended to generalize beyond the locally-established domain. These are predictions about which the mind-external view is mostly silent.

2.1. Domain size

An important and well-studied property of object-files is that they are subject to stringent working memory constraints. This working memory limit has often been observed in adults in multiple object tracking (e.g., Pylyshyn & Storm, 1988) and change detection (e.g., Vogel, Woodman, & Luck, 2001) paradigms. In both cases, performance sharply declines when participants are asked to track 5 or more objects at once. In infants, a working memory limit of 3 object-files has been well-documented. Feigenson and Carey (2005), for example, show that infants fail to distinguish 1 from 4 objects. This catastrophic nature of this failure to represent more than 3 object-files simultaneously cannot be overstated: infants reliably choose 3 crackers over 1, but perform at chance when given the choice between 4 and 1.

Ensemble representations are not subject to this sort of working memory constraint. Though there is a limit on the number of distinct ensembles adults and children can represent simultaneously (Halberda, Sires, & Feigenson, 2006; Zosh, Halberda, & Feigenson, 2011), there seems to be no upper limit on the number of objects that can be grouped as a single ensemble. Indeed, one of the benefits of grouping objects into an ensemble representation is the ability to represent them in terms of summary statistics and thus avoid the working memory limit. Using ensemble representations, infants can successfully represent large numbers of objects simultaneously (Xu & Spelke, 2000). To take one illustrative case, Wood and Spelke (2005) show that 6-month-olds successfully distinguished 8 versus 4 actions (jumps of a puppet) but failed to distinguish 4 versus 2 actions in an otherwise identical experimental setup. It seems, then, that presenting small numbers of actions caused the infants to attempt to individuate them and treat them as object-files (or 'event-files', in this case). In doing so, infants ran up against their working memory limit and failed. But seeing larger

numbers of events naturally triggered grouping those events as two distinct ensemble representations, which could be compared, leading to success at the task.

This working memory contrast gives rise to our first prediction. Ensembles are better able to – and perhaps even tailor-made for – representing large numbers of things. By hypothesis, the meaning of "every" serves as a call to group the things being quantified over. And by hypothesis, this pushes participants to represent the things being quantified over with the system for ensemble representation. So, "every" should be preferred over "each" when the domain of quantification is large. That is to say, when a speaker finds themselves wanting to state a universal generalization and is faced with a large number of things over which to state that generalization, they will likely be representing those things as an ensemble (since large numbers of things cannot simultaneously be represented as independent object-files). The meaning of the expression "every thing" has a constituent that corresponds to "the things" whereas the meaning of "each thing" does not. So the meaning of "every" suggests treating the relevant things as a group. Given that the mind's way of representing groups is with the system for ensemble representation, "every" is a more suitable choice than "each" for large domains, all else equal.

To take a more concrete example, imagine a parent and their child are on a walk through the forest when they encounter some frogs. Suppose the parent wants to offer a universal generalization about those frogs: they're green. Our claim is that the parent would experience pressure to quantify over the frogs with "each" if there were, say, three of them; but would experience pressure to instead use "every" if there were thirty. This is not to say speakers couldn't or wouldn't use "each" given thirty frogs (they may do just that if they want to highlight the individual frogs). And it is not to say that speaker couldn't or wouldn't use "every" given only three frogs. Here, the non-deterministic nature of the proposed link between semantic representations and non-linguistic systems is important: the proposed representations in (3) do not strictly preclude the deployment of any particular cognitive system. The claim is just that "each" and "every" suggest the use of different cognitive systems. Smaller domains are a better fit for "each" given known properties of object-files, which "each", by hypothesis, invites to a greater degree than "every".[6]

Knowlton and Gomes (2022) offer some initial support for this prediction by analyzing a corpus of child-directed speech: parents are overwhelmingly more likely to quantify over small numbers of things (fewer than 3; within the child's working memory limit) with "each" than with "every" (or "all"). Of course, other things differ about these contexts as well (e.g., the physical presence of the domain; intention to generalize beyond the domain). So in Experiments 1 and 2 below, we directly test the domain size prediction by contrasting "each" and "every" in otherwise identical and controlled contexts.

2.2. Generalization

Our second prediction comes from another difference between object-files and ensembles: the former represent individuals and their associated properties independently of one another, whereas the latter describe many individuals in terms of their shared summary statistics. Moreover, representing an ensemble does not require representing the individuals that constitute it. Consider an example from Haberman and Whitney (2011). They show that participants can tell which of two collections of faces are happier on average despite not being able to identify which individual faces changed from one image to the next. That is, participants can represent some faces as an ensemble – and thus encode properties like the group's average happiness – without encoding information about each individual face's expression. Likewise, Sweeny et al. (2015) report that 4- and 5-year-old children are accurate at determining which of two trees had larger oranges on average even when showing relatively poor accuracy at discriminating individual orange sizes. These sorts of results suggest that arriving at a representation of an ensemble summary statistic like average size is more akin to texture processing than to literally averaging the sizes of a sample of individuals (Im & Halberda, 2013).

The important point for our purposes is that ensemble representations are essentially generalizations. They are defined in terms of summary properties, like the average color and range of colors of the group members. An ensemble representation of some things thus naturally licenses a prediction about new members of that group of things. If one is told that a sixth lemon is a member of the same ensemble as five other lemons, it's likely that sixth lemon will have similar properties to the rest of the collection (e.g., it will be a similar hue), because that's what it means to be a part of the ensemble. Moreover, ensembles permit somewhat vague boundaries about which particular things are included in the group (within reason; an ensemble representation of five lemons could be expanded to include a sixth, but could not be extended to include a cutting board). The same cannot be said of object-files, which are first and foremost indices to individuals. Individual properties are bound to these indices, but that two indices share some properties (e.g., being yellow) is an accident as far as the object-file system is concerned. Representing three lemons as independent object-files, on its own, does not license a prediction that the next object encountered will also be yellow. For a prediction about the next object, one would have to employ additional cognitive systems beyond parallel individuation (e.g., pattern recognition). Put another way, it seems reasonable to us to assume that generalization requires categorization (e.g., Waxman & Markow, 1995; Jaswal & Markman, 2007). But while object-files, on their own, do not inherently link to categorization, ensembles do.

[6] The idea might be put as follows. First-order variables present in representations like (3a-b) invite representing the things that serve as their values as independent object-files; second-order variables present in representations like (3b) invite grouping the things that serve as their values as an ensemble collection. Since (3b) has a second-order variable, only "every", by virtue of its semantic representation, calls to mind ensembles. But "every" might also call to mind object-files, seeing as there is a first-order variable present in (3b) as well. And in any case, object-files or ensembles might be triggered in a given situation for reasons independent of the semantic representation in question (e.g., details of the visual scene). So even in a case where "each thing" is used to quantify over some things, there is no hard and fast restriction against those things being represented as an ensemble collection.

This difference leads to our second prediction: "every" should be preferred when universal quantification is meant to extend beyond the locally-established domain. To see how this differs from the domain size prediction above, consider again the scenario of a parent walking their child through the woods and encountering some frogs. The parent might want to convey the idea that these particular frogs are green because in general, all frogs are green. And they might have that desire to project beyond the local domain regardless of whether that domain consists of three or thirty frogs. This is not to say that "every" is always used to project beyond the local domain, nor that "every" has a generic meaning in any sense. The claim is just that "every", because its meaning calls to mind treating the things quantified over as an ensemble, is more likely to enable projection than "each". The reason is that ensembles' boundaries can be extended to include more things. In our view, projection beyond the local domain explains the martini example from Section 1 ("every martini needs an olive" is naturally understood as being a component of a drink recipe as opposed to being a claim about a handful of particular martinis).

Of course, one could also represent things quantified over as an ensemble and opt not to project beyond the local domain. That is, ensembles don't require extending boundaries to include more things. In this case, an "every"-sentence would be interpreted as no more of a broad generalization than the corresponding "each"-sentence. Again, the predicted effect is a non-deterministic one. But the prediction is that "every" will be more likely than "each" to support expanding the domain of quantification arbitrarily, even beyond the things initially considered. We test this prediction in Experiments 3 and 4a–d.

2.3. Experimental strategy

These two predictions are fundamentally about the relationship between quantifier meanings and their contexts of use. As such, our strategy for testing these predictions is to focus on both directions of this relationship. For domain size, Experiment 1 manipulates the linguistic context (the size of the domain) and asks how this change affects participants' preference for using "each" or "every" in a forced-choice task. Inversely, Experiment 2 manipulates the quantifier used, and asks what effect such use has on participants' reasoning about the context (specifically, their estimates of how large the domain of quantification is).

Similarly, for generalization, Experiment 3 manipulates the linguistic context to indicate whether a generalization beyond the stated quantificational domain is intended and asks how this change affects participants preference for using "each" or "every". Inversely, Experiments 4a-d manipulate the quantifier used in conjunction with a visual scene, and ask what effect the use of "each" or "every" has on participants' willingness to draw a generalization beyond the initial scene in a forced-preference task. These manipulations and tasks are briefly summarized in Table 1. It's worth noting that the dependent measures used here differ from more standardly used plausibility judgments or intuitions about semantic anomaly. Instead, these experiments probe general preferences about appropriate context of use.

3. Experiments

3.1. Experiment 1: Domain size (forced-choice)

Experiment 1 tests the novel prediction that the size of the domain of quantification should impact preferences for using "each" or "every" in a forced-choice task. In particular, "every" should be preferred given larger domains (due to the lack of working memory limit on ensemble representations), whereas "each" should be preferred given fewer things to quantify over (due to the strict working memory limits of object-files).

3.1.1. Method

3.1.1.1. Subjects. One hundred adults participated over the internet. All participants (here and in the experiments that follow) were native English speakers living in the United States. They were recruited on Prolific (https://www.prolific.co) and gave informed consent prior to participating.

3.1.1.2. Procedure and materials. The experiment was designed using PCIbex (Zehr & Schwarz, 2018), as were all subsequent experiments reported in this paper. In this experiment, participants completed a forced-choice judgment task. They were asked to choose between two sentences that served as possible continuations to a context sentence. They were instructed as follows: "Type '1' if you think the first sentence is a better fit or '2' if you think the second sentence is a better fit". Sentence order was counterbalanced (i.e.,

Table 1
Summary of experimental manipulations.

Prediction	Exp	Manipulation	Task
Domain size	**1**	**Small vs. large domain** ("three/three thousand martinis")	Forced-choice task ("each/every martini")
	2	**Quantifier used in text** ("each/every martini")	Estimate of domain size ($\leq 3/\geq 4$)
Generalization	**3**	**Local vs. global domain** ("martini that he made/that's worth drinking")	Forced-choice task ("each/every martini")
	4a-d	**Quantifier used in text alongside a visual scene** ("each/every dax is green")	Forced-preference task (unsure/certain about generalizing beyond daxes seen in initial scene)

half of the time "each" was read first; half of the time "every" was read first). On target trials, the context sentence either established a small or large domain of quantification. For example, in one item, the domain consisted of either "three martinis" or "three thousand martinis" (small and large cardinalities were selected based on what felt natural given the particulars of the stimuli; see the Appendix for a complete list of materials). The two continuation options contained quantification over this domain, with the sentences being identical except for the use of "each" vs. "every", as in (4).

(4) a. The bartender at the local tavern has made three martinis.
He said that {each/every} martini he made had an olive.
b. The bartender at the local tavern has made three thousand martinis.
He said that {each/every} martini he made had an olive.

Twelve target items were created. In all cases, the quantified expression contained a relative clause ensuring that the domain of quantification is restricted to the initially-established group (e.g., "each/every martini **he made**"). The purpose of this addition was to ensure that the domain-size effect at issue here is separate from the propensity to use "every" for generalization, at issue in later experiments. No generalization is possible in the second sentences of (4), for example, since the domain is explicitly restricted to martinis that the bartender made (*cf.* "he said that every martini has an olive", which might invite generalization).

An experiment list consisted of these items (6 with a small domain; 6 with a large domain) plus 24 filler items, which were designed to provide participants choices between two possible answers that differed in form but were similar in content (e.g., "In her favorite book, the main character is a talking dog" versus "The main character is a talking dog in her favorite book"; all Filler items are listed in the Appendix). A second experiment list was created from the first by swapping the context condition (small vs. large domain) of each target item but otherwise leaving the lists identical. Participants were assigned randomly to one of the two lists. Filler and target trials were intermixed, with each participant receiving a different random order.

3.1.2. Results

As seen in Fig. 1, our prediction was borne out. Participants were more likely to pick "every" when given a context sentence with a large domain compared to when given a context sentence with a small domain. To test this prediction statistically, responses were fit with a mixed-effects binomial regression model with maximal random effects structure (Barr, Levy, Scheepers, & Tily, 2013): random slopes and intercepts for participant and item. The dependent measure was whether the participant selected the "every" sentence (as opposed to the "each" sentence) and the independent variable of interest was context (small versus large domain, coded as −1 and 1, respectively). No participants were removed from analysis but particular trials within each participant were removed if that trial's reaction time was 2.5 standard deviations higher or lower than the average reaction time. This removed a total of 21 trials.

We tested for significance of context by conducting a likelihood ratio test on the Chi-square values from model comparison (comparing the model with context as a fixed effect against an intercept-only model with the same random effects structure). This confirmed that the large domain made participants more likely to select "every" over "each" ($\chi^2(1) = 10.67$, $p < .01$; main effect of context: $\beta = 0.36$ [95% CI 0.27–0.46], $z = 3.95$, $p < .001$).

Participants were not more likely than chance to use "each" for small domains: they selected "each" about 51% of the time in this condition. This probably reflects a baseline preference for "every" in both conditions (perhaps owing to relative frequency). The

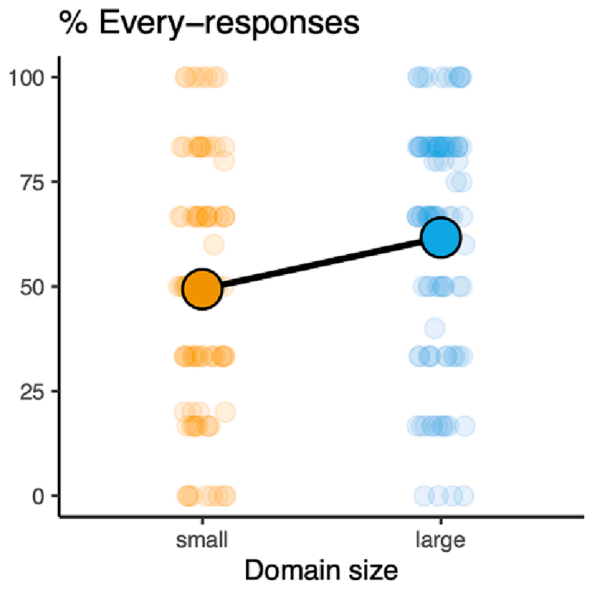

Fig. 1. Rate of choosing the "every" sentence given a small or large domain of quantification (e.g., "three martinis" vs. "three thousand martinis"). Large points represent group means; small points represent individual averages.

important point for our purposes, though, is this: despite any baseline preference for "every", manipulating domain size affects preferences in the predicted way. Namely, increasing the domain size encourages using "every" instead of "each".

3.2. Experiment 2: Domain size (free-response)

Experiment 2 aimed to confirm the results of Experiment 1 in a more direct way. In particular, participants were simply asked how large they thought the domain of quantification was given a sentence with "each" or "every". We predict participants to respond with smaller domains in the "each" condition than the "every" condition, given the working memory limit associated with object-file but not ensemble representations.

3.2.1. Method

3.2.1.1. Subjects. One hundred ninety eight adults participated over the internet. All participants were native English speakers living in the United States. They were recruited on Prolific (https://www.prolific.co) and gave informed consent prior to participating.

3.2.1.2. Procedure and materials. Participants completed a one-trial open-ended response task. They were given the prompt in (5) with one of the two quantifiers and asked to type in an answer. Quantifier was manipulated between-subjects and participants were randomly assigned to one of the two conditions.

(5) Imagine that a bartender said the following: "{each/every} martini I made had an olive". How many martinis would you guess they have in mind?

3.2.2. Results

As seen in Table 2, participants were more likely to provide an answer of three or fewer (within the working memory limit for object-files) in the "each" condition than in the "every" condition ($\chi^2 = 12.19$, $p < .001$). This corroborates the findings from Experiment 1. The two experiments together bear out the predictions discussed above: all else equal, "each" is well-suited to quantifying over small numbers of things whereas "every" is preferred when dealing with larger domains of quantification.

3.3. Experiment 3: Generalization (forced-choice)

In Experiment 3, we aimed to test the generalization prediction – that "every" should be preferred in contexts that call for projecting beyond the locally-established domain – using the same forced-choice paradigm from Experiment 1. Such a prediction is expected if "every" encourages engagement of ensemble representations which encode generalizations over individuals.

3.3.1. Method

3.3.1.1. Subjects. One hundred adults participated over the internet. All participants were native English speakers living in the United States. They were recruited on Prolific (https://www.prolific.co) and gave informed consent prior to participating.

3.3.1.2. Procedure and materials. Participants completed a forced-choice judgment task similar to Experiment 1. They were asked to choose between two sentences that served as possible continuations to a context sentence. They were instructed as follows: "Type '1' if you think the first sentence is a better fit or '2' if you think the second sentence is a better fit". The sentences corresponding to '1' and '2' were counterbalanced (i.e., half of the time "each" was read first; half of the time "every" was read first). The context sentence always established a constant domain ("a few martinis"). What differed between conditions was whether the quantificational phrase referred back to that domain explicitly or went beyond it, as signaled by one of two relative clauses. In the 'local' condition in (6a), for example, "martini that he made" refers back to those few drinks mentioned in the context sentence. But in the 'global' condition in (6b), "martini that's worth drinking" projects beyond that locally-established domain to make a far more general claim.

(6) a. The bartender at the local tavern made a few martinis.
He said that {each/every} martini that he made has an olive.
b. The bartender at the local tavern made a few martinis.
He said that {each/every} martini that's worth drinking has an olive.

Table 2
Participants' responses to being asked the question in (5). "Exhaustive" responses included answers like "all of them" and "every single one".

	≤3	4–5	≥6	Infinitely many	Exhaustive
Each	49	14	31	0	1
Every	20	18	55	1	3

In this experiment, then, the initial sentence was identical across conditions. What differed (within-subjects) was the relative clause modifying the noun phrase quantified over.

As in Experiment 1, here, twelve target items were created (see Appendix for full list of materials). An experiment list consisted of these items (6 with a small domain; 6 with a large domain) plus the same 24 filler items from Experiment 1. A second experiment list was created from the first by swapping the relative clause condition ('local' vs. 'global') of each target item but otherwise leaving the lists identical. Participants were assigned randomly to one of the two lists. Filler and target trials were intermixed, with each participant receiving a different random order.

3.3.2. Results

As seen in Fig. 2, our prediction was borne out. Participants were more likely to pick "every" given the 'global' continuation (i.e., given a relative clause that explicitly goes beyond the established domain) compared to the 'local' continuation (i.e., given a relative clause that refers back to that domain). To test this prediction statistically, responses were fit with a mixed-effects binomial regression model with maximal random effects structure, as in Experiment 1. The dependent measure was whether the participant selected the "every" sentence (as opposed to the "each" sentence) and the independent variable of interest was the relative clause continuation ('local' versus 'global' continuation, coded as −1 and 1, respectively). No participants were removed from analysis but particular trials within each participant were removed if that trial's reaction time was 2.5 standard deviations higher or lower than the average reaction time. This removed a total of 25 trials.

We tested for significance of continuation ('local' vs. 'global') by conducting a likelihood ratio test on the Chi-square values from model comparison (comparing the model with continuation as a fixed effect against an intercept-only model with the same random effects structure). This confirmed that the 'global' continuation made participants more likely to select "every" over "each" ($\chi^2(1) = 9.95$, $p < .01$; main effect of continuation: $\beta = 0.29$ [95% CI 0.21–0.37], $z = 3.72$, $p < .001$).

These results confirm the prediction that "every" should be preferred in cases calling for generalization. However, there is the potential worry that this result reduces to the domain size result discussed in Section 2. After all, the 'global' cases here implicate a larger domain than the 'local' cases: the martinis worth drinking are likely to outnumber the martinis that the bartender just made. So even though the initially-established domain was the same in both cases ("a few martinis"), the domain of quantification is nonetheless larger in the 'global' case. Experiments 4a-d rule out this potential confound by explicitly asking participants how comfortable they are generalizing beyond a domain initially quantified over with "each" or "every".

3.4. Experiment 4a: Generalization (forced-preference)

The four versions of Experiment 4 (4a, 4b, 4c, and 4d) also tested the prediction that "every" should be preferred for generalization beyond the locally-established domain. But instead of asking participants to choose between "each" and "every", participants were given sentences with either "each" or "every" that explicitly quantified over a small number of visually-presented items. They were subsequently asked about their confidence generalizing the property attributed to those items (that they are all green) to a new, unseen item. We expect participants to be more likely to do so when the objects are quantified over with "every" than with "each", given that "every" is more likely to encourage ensemble representation, which readily supports generalization.

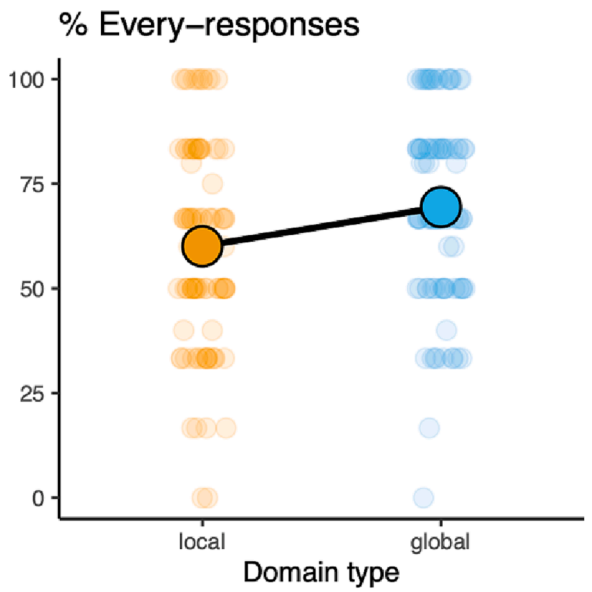

Fig. 2. Rate of choosing the "every" sentence given a local or global domain of quantification (e.g., "martini that he made" vs. "martini that's worth drinking"). Large points represent group means; small points represent individual averages.

3.4.1. Method

3.4.1.1. Subjects. Three hundred adults participated over the internet. All participants were native English speakers living in the United States. They were recruited on Prolific (https://www.prolific.co) and gave informed consent prior to participating.

3.4.1.2. Procedure and materials. Participants completed a one-trial forced-preference experiment in which they were introduced to three novel creatures called "daxes" (Fig. 3). They were subsequently told either that "each dax is green" or that "every dax is green". Participants were randomly assigned to one of the two quantifiers. After being introduced to the novel creatures, participants were shown the silhouette of another dax, whose color was hidden by a shadow. They were then asked how confident they were that this hidden dax was also green and responded using a 7-point scale (with "totally unsure" at one end and "completely certain" at the other). They could not select the midpoint of the scale where the slider began, they had to move the slider at least somewhat to the left (toward "totally unsure") or somewhat to the right (toward "completely certain").

3.4.2. Results

As seen in Fig. 4, participants who were initially introduced to the fact that all three daxes were green with the "every" sentence were more confident generalizing about the unseen dax's color than those who were initially introduced to the daxes' colors with the "each" sentence (Welch's $t(284.42) = 4.51$, $p < .001$).

This result corroborates the findings from Experiment 3: "every" is preferred for generalizing beyond the locally-established domain. And importantly, Experiment 4a controlled for the potential domain size confound. That is, in both the "each" and "every" conditions in Experiment 4a, the domain directly quantified over was the three initial daxes. So the finding that "every" made participants more comfortable generalizing to a novel instance cannot be explained as another instance of the domain size difference established in Section 2. Experiment 4b shows that this effect disappears when the domain is explicitly restricted in the quantified sentence. Experiments 4c and 4d probe the robustness of this effect by adding other cues to promote generalization.

3.5. Experiment 4b: Generalization (restricted domain)

Here we attempted to erase any difference between "each" and "every" with respect to propensity to generalize. If the results of Experiment 4a are due to "every" being preferred for generalizing beyond the locally-established domain, as hypothesized, then the

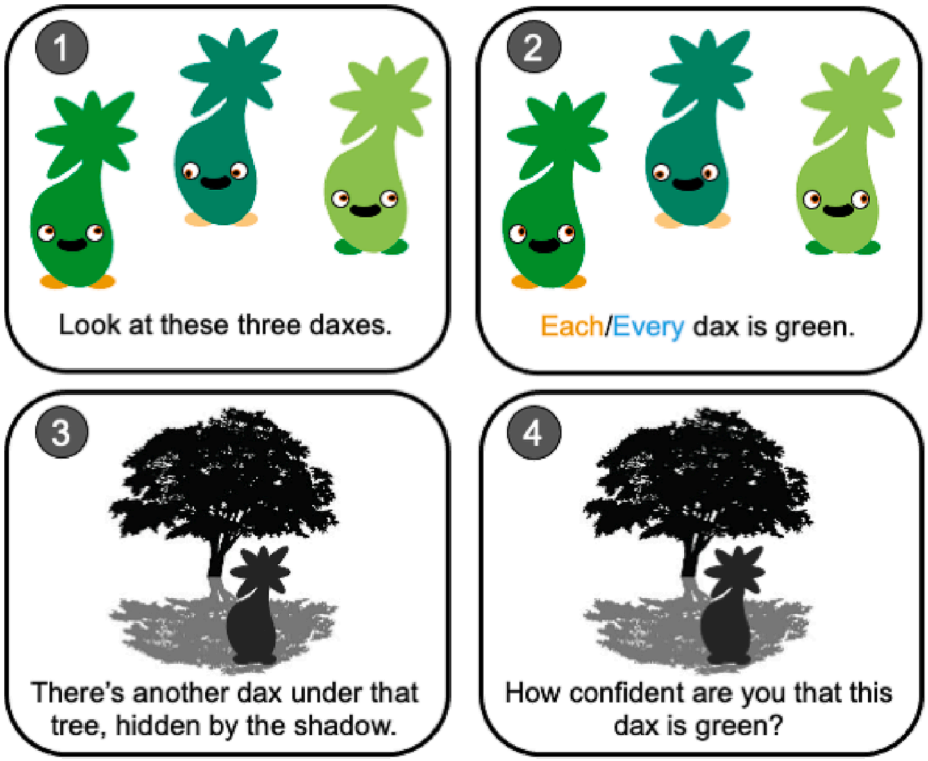

Fig. 3. Trial structure of Experiment 4. Participants were first introduced to three novel creatures, then either told "each dax is green" or "every dax is green" before being asked to provide their confidence generalizing to a novel dax whose color was hidden by a shadow. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

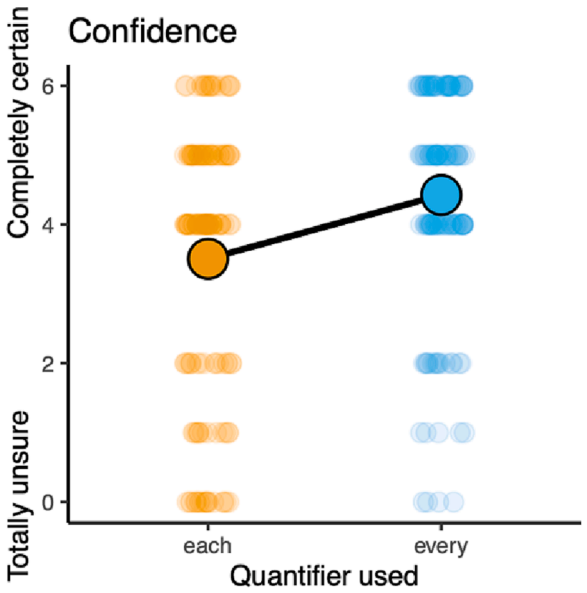

Fig. 4. Confidence (0–6 scale) that the hidden dax is green in Experiment 4a. Large points represent group means; small points represent individual responses. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

effect disappear by when the domain is explicitly restricted. We accomplished this by simply adding the word "here" but otherwise keeping the task exactly the same. We predict this addition should remove the asymmetry observed between "each" and "every" in Experiment 4a.

3.5.1. Method

3.5.1.1. Subjects. Three hundred adults participated over the internet. All participants were native English speakers living in the United States. They were recruited on Prolific (https://www.prolific.co) and gave informed consent prior to participating.

3.5.1.2. Procedure and materials. Participants completed a one-trial experiment with the exact same structure as Experiment 4a. The only difference was the addition of the word "here" in the quantified sentence. That is, participants were introduced to the daxes either by being told that "each dax here is green" or that "every dax here is green". Participants were randomly assigned to one of the two quantifiers. After being introduced to the novel creatures, participants were shown the silhouette of another dax, whose color was hidden by a shadow. They were then asked how confident they were that this hidden dax was also green and responded using a 7-point scale (with "totally unsure" at one end and "completely certain" at the other).

3.5.2. Results

As seen in Fig. 5, the effect observed in Experiment 4a disappeared with the addition of "here". The "every" sentence and the "each" sentence did not significantly differ in terms of generalization confidence (Welch's $t(198.78) = 0.83$, $p = .409$). This result is as expected: "every" invites generalization in the sense that it invites expanding the domain of quantification, in this case from the daxes seen to all daxes. But when the domain is explicitly restricted to just the daxes seen, any difference between "each" and "every" disappears.

3.6. Experiment 4c: Generalization (homogeneous display)

Here we attempted to increase participants' propensity to generalize in the "each" condition by increasing the level of homogeneity in the initial presentation. The intuition was that being presented with novel creatures of the exact same shade of green would make it seem more likely that the color was a property of the kind. If the linguistic framing effect is strong enough, it might withstand this manipulation; but in any case, we predict the affect to be ameliorated compared to Experiment 4a given the independent cue to represent the objects as an ensemble in both conditions.

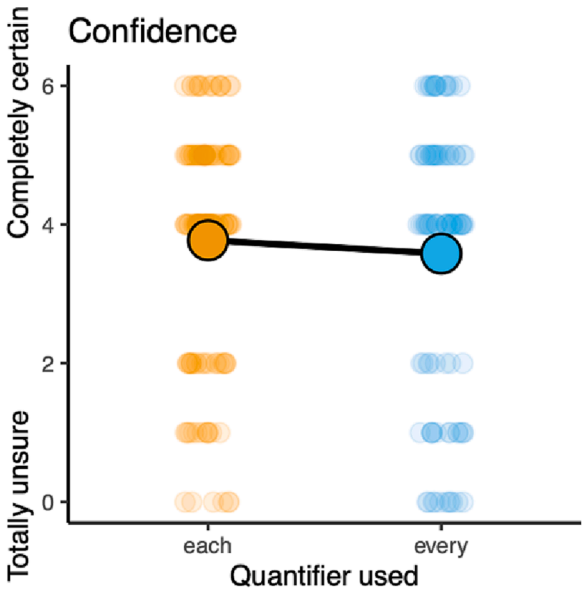

Fig. 5. Confidence (0–6 scale) that the hidden dax is green in Experiment 4b. Large points represent group means; small points represent individual responses. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.6.1. Method

3.6.1.1. Subjects. Three hundred adults participated over the internet. All participants were native English speakers living in the United States. They were recruited on Prolific (https://www.prolific.co) and gave informed consent prior to participating.

3.6.1.2. Procedure and materials. Participants completed a one-trial experiment with the exact same structure as Experiment 4a, save for the fact that the three daxes were the same shade of green (all daxes were the color of the leftmost dax in Fig. 3). As in Experiment 4a, participants were introduced to the daxes either by being told that "each dax is green" or that "every dax is green". Participants were randomly assigned to one of the two quantifiers. After being introduced to the novel creatures, participants were shown the silhouette of another dax, whose color was hidden by a shadow. They were then asked how confident they were that this hidden dax was also green and responded using a 7-point scale (with "totally unsure" at one end and "completely certain" at the other).

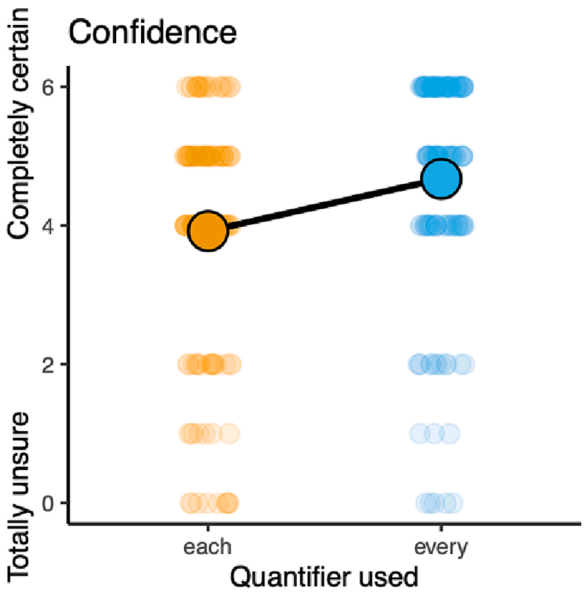

Fig. 6. Confidence (0–6 scale) that the hidden dax is green in Experiment 4c. Large points represent group means; small points represent individual responses. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.6.2. Results

As seen in Fig. 6, the same effect was observed here as in Experiment 4a: the "every" sentence led to greater confidence generalizing about the unseen dax's color than the "each" sentence (Welch's $t(298.63) = 4.12$, $p < .001$). This amounted to a replication of the original effect under even more stringent conditions. Experiment 4c further pushed the boundaries of the effect by providing an even stronger cue to generalizability.

3.7. Experiment 4d: Generalization (larger domain)

In Experiment 4d, we endeavored to increase participants' eagerness to generalize by showing more creatures within the visual scene at the outset (five instead of three). This was expected to create better conditions for generalization for a few reasons. First, as discussed in Section 2, five objects lie slightly beyond the working memory limit for simultaneous representation of object-files. As such, presenting participants with five objects usually triggers ensemble representation of those objects. And, by hypothesis, ensemble representations support generalization in a way that object-files do not. As a result, increasing the pressure to treat the novel creatures as a single ensemble should be expected to increase generalization as a consequence. Second, even setting aside the additional bias to represent ensembles, seeing many objects (crucially all at once; see Wang & Trueswell, 2022) with the same property gives rise to a "suspicious coincidence" (e.g., Xu & Tenenbaum, 2007). All else equal, it would be more surprising to encounter five objects of the same kind with the same property if that property was not fundamental to the kind than it would be to encounter three objects with that property.

As in Experiment 4c then, if the linguistic framing effect is strong enough, it might withstand this manipulation. But we predict the affect to be ameliorated compared to Experiment 4a given the independent cue to represent the objects as an ensemble in both conditions.

3.7.1. Method

3.7.1.1. Subjects. Three hundred adults participated over the internet. All participants were native English speakers living in the United States. They were recruited on Prolific (https://www.prolific.co) and gave informed consent prior to participating.

3.7.1.2. Procedure and materials. Participants completed a one-trial experiment with the same structure as Experiment 4a, save for two changes. First, instead of showing three daxes, as in Fig. 3, participants were initially shown five daxes. Second, instead of introducing the creatures with the phrase "look at these three daxes", the creatures were introduced with the phrase "look at these daxes". Otherwise, the experiment was identical to Experiment 4a.

3.7.2. Results

As seen in Fig. 7, the same effect is again present (albeit ameliorated). Participants exposed to the "every" sentence were on average more confident in generalizing than those exposed to the "each" sentence (Welch's $t(280.89) = 2.45$, $p < .05$). This further

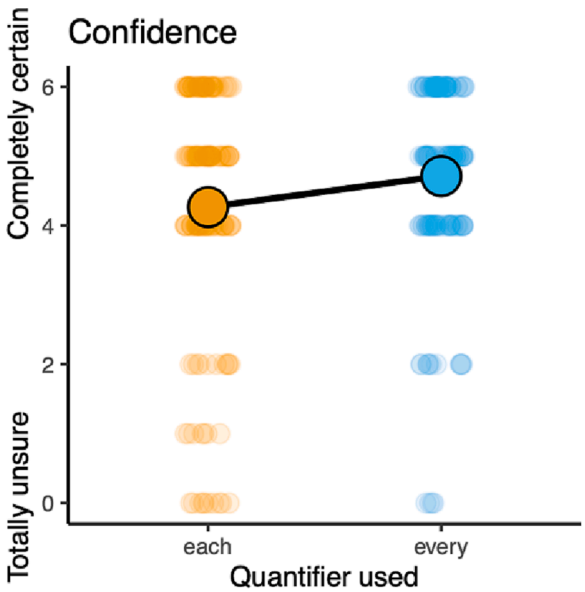

Fig. 7. Confidence (0–6 scale) that the hidden dax is green in Experiment 4d. Large points represent group means; small points represent individual responses. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

confirmation of the effect suggests that the linguistic framing (and in particular using "each" versus "every" to encourage treating the creatures as individuals or as a single group) matters even when other cues to generalizability are manipulated.

4. General discussion

Broadly speaking, there are two contrasting views about what linguistic meanings are. The 'mind-external' view, standard in many corners of formal linguistics, holds that meanings are situated in the world, whereas the 'mentalistic' alternative maintains that meanings are psychological objects. These different approaches to meaning imply different approaches to the relationship between semantics, pragmatics, and cognition. We believe the experiments presented above support a particular version of the mentalistic approach in which an expression's meaning is a mental representation that has a particular formal structure which causes it to interface with certain non-linguistic systems. The details of those interfacing systems, in turn, have consequences for how the expression gets pragmatically used.

More narrowly, the particular experiments presented here show that "every" is preferred over "each" when the domain of quantification is large as opposed to small; furthermore, "every" is preferred when quantification is intended to generalize beyond a locally-established domain. These differences in the pragmatic use of "each" and "every" are predicted by the mentalistic proposal about their meanings given in Section 1.2. Namely, these contrasts can be explained by making appeal to the properties of object-file and ensemble representations, the two cognitive systems that, by hypothesis, naturally interface with the meanings of "each" and "every" due to the formal differences between them.

Though differences in the use of "each" and "every" have been noted in various forms over the years (e.g., Vendler, 1962; Beghelli & Stowell, 1997; Tunstall, 1998), they have not yet been part of a unified picture of meaning; part of an explanatory story about how meanings link to contexts of use. The view offered here attempts to do just that by pushing some of the explanatory burden for interpretive variability to non-linguistic cognition. The usage differences considered are not explained as a matter of grammar *per se*, but as a matter of what pieces of cognition linguistic meanings call upon and the pragmatic purposes these pieces most appropriately align with. This view thus offers a picture of how formal properties of meanings and psychological details of non-linguistic cognitive systems both play a role in explaining the particulars of language use.

4.1. Is a sparser mentalistic account possible?

Given the three-pronged nature of our proposal, a question that naturally arises is whether a sparser, two-pronged mentalistic account that omits one of the interface levels is possible. In our view, the answer to this question is no.

To begin with, one might wonder if the present results could be explained without making any appeal to distinct non-linguistic representational systems. After all, a standard approach to pragmatics is to compute predictions about language use directly from formal semantic representations. For example, in work on alternatives – what is said versus what could have been said – the relevant alternatives to what is said are taken to be competing expressions (Katzir, 2007; Fox & Katzir, 2011; *cf.* Buccola, Križ, & Chemla, 2022). Thus one could obviate the need for invoking object-files and ensembles by trying to derive the predictions about domain size and generalization from the proposed semantic representations in (3), repeated below.

(3) a. $\forall$x:Frog(x)[Green(x)] "each frog is green"
≈*any thing$_x$ that is a frog is green*
b. TheX:Frog(X)[$\forall$x:X(x)[Green(x)]] "every frog is green"
≈*the frogs$_X$ are such that any thing$_x$ that is one of them$_X$ is green*

Assuming that "each" and "every" have different representations along these lines, could the same predictions have been captured by this approach? It's far from obvious that they could have been. The formal distinction between quantifying into first-order variables, as in (3a), and second-order variables, as in (3b), just requires that the variables in question either take on one value at a time (i.e., per assignment of values to variables) or potentially more than one value at a time (alternatively, that they either take on individuals as their values or some sort of plural object as their values). There is nothing about this formal distinction that suggests a differential limit on the number of values either variable can take on. It is only when a connection is made between these types of variables and distinct non-linguistic representational systems that predictions about working memory limitations are motivated.

Assuming that there is a need to appeal to the posited non-linguistic representational systems to account for the present results, one might wonder if there is a need for the posited formal distinction. As noted in footnote 1, avoiding formal commitments is a common tack taken by 'cognitive' approaches (for review, see Talmy, 2019). In the present case, the question is whether anything is added by positing that the formal distinction between first-order and second-order quantification is at play in the meanings of "each" and "every". That is, what is gained by proposing the specific formalisms in (3) as opposed to merely supposing that "each" directly connects to object-files whereas "every" connects to ensembles?

In the first place, some kind of formalism is needed to deal with the fact that the universal quantifiers "each" and "every" are universal. That is, "every frog" differs from "some frog", "no frog", and "most frogs" in that the former has universal force. The meaning of "every frog" cannot just be "treat the frogs as an ensemble", as that would leave the universality aspect unexplained. And formal logic (and by extension formal semantics) provides tools ripe for expressing this sort of content. The distinction relied on here – between first-order and second-order variables and quantification into those variable positions – has been an important one in logic since Frege (1879). But in any case, some formal distinction is needed to capture the precise nature of quantifiers. After all, they express a

kind of abstraction over things in the world: one can't literally see "every" in the same way that one can see "frog". So using perception as grounding for the semantics of these sorts of logical expressions is hopeless from the outset.

More generally, we take it that one of the main goals of a comprehensive theory of meaning is offering an explanation of compositionality: how smaller parts get combined to yield the meaning of the whole. Of course, there are differing views about the nature of compositionality and how it should be explained (e.g., Hodges, 1998, 2001; Szabo, 2000; Pagin, 2003; 2012; Pelletier, 2004; Pietroski, 2004; Barker & Jacobson 2007). But at a minimum, a complete theory of meaning needs to allow for the meaning of a complex expression to be determined by the meanings of its constituent parts and their syntactic arrangement. We did not attempt to tackle this aspect of meaning in the present paper, but our proposed semantic representations are compatible with multiple formal approaches to meaning composition, including standard approaches to compositional semantics or more radical departures from the standard view (e.g., Pietroski, 2018).

Importantly though, in saying that appeal to formal differences is needed, we are not suggesting that standard formal semantic theorizing can be taken off the shelf and 'pushed into the head'. The particular mentalistic approach advocated here departs quite a bit from the textbook semantic view on quantifier meanings. As noted in the introduction, the standard view maintains that quantifiers express second-order relations between two independent sets. But neither of the proposed mentalistic meanings in (3a–b) describe relations, and the meaning for "each" in (3a) is completely first-order. What we want to retain from the formal approach is not the details of (any particular version of) that approach, just the idea that there is a level of representation where precise formalization is possible, useful, and amounts to an accurate description of the representation.

4.2. A non-mentalistic alternative?

Could the more traditional mind-external view account for the present results? It is not obvious to us that it could, at least not without ad-hoc amendments. To see why this is so, consider a well-known spelling-out of the standard, mind-external view from Beghelli and Stowell (1997). Their proposal captures differences between "each" and "every" in terms of differences in associated syntactic features. The gist is that "each" comes with a feature that causes it to associate with a distributivity operator – an unpronounced syntactic element – which is responsible for ensuring that predicates apply to individuals. This operator also happens to be located higher in the syntactic tree than the generic operator – another unpronounced syntactic element – which is responsible for giving sentences generic meanings. Since "every" lacks this feature, it can remain lower in the syntactic tree, beneath these operators. This view was designed to accommodate differences between "each" and "every", like those raised in Section 1.1 (though a frequent criticism of this sort of syntactic approach is that it often recapitulates the data as opposed to offering an explanation for it; Larson, 2021).

Given this view, consider the domain size result of Experiments 1 and 2. Being lower in the tree than "each" does not obviously predict that "every" would be preferred for larger domains. And it is likewise non-obvious that the propensity of "each" to associate with the distributivity operator would help explain the domain size difference (all this operator does is ensure that predicates apply to individuals in the domain; it says nothing about what size that domain is likely to be). One could consider the idea that the distributivity operator itself serves as a call to the object-file system, and thus explain the domain size difference in mentalistic terms similar to our own. But this would be to admit that appeal to non-linguistic cognition is needed. And moreover, this approach would have a hard time explaining why "every" doesn't behave exactly like "each" when the predicate used is distributive. Put another way, if distributive readings arise from association with the distributivity operator, and if associating with the distributivity operator gives rise to the domain-size effect, then why do we observe the effect even for sentences that are fully distributive ("every martini needs an olive" is no less distributive than "each martini needs an olive" since the predicate is meant to apply to each one of the martinis individually)?

Second, consider the generalization contrast from Experiments 3 and 4a-d. Here, the mind-external view could offer an explanation in terms of the generic operator. "Each", on this view, must undergo syntactic movement above the generic operator, and in doing so avoids its influence. "Every", on the other hand, is trapped beneath the generic operator, which puts "every" under its spell. But the relevant phenomenon seems to be more restricted than genericity writ large. Indeed, the syntactic view has been argued to make incorrect predictions (Brendel, 2019; Knowlton, 2021; Surányi, 2003). For example, as Knowlton (2021) points out, "every" does not give rise to 'generic' understandings in all the places where the generic operator has been posited. A sentence like "every tick carries Lyme disease" is seemingly false, whereas a sentence like "ticks carry Lyme disease" is easily understood generically (by hypothesis, because of the generic operator). The same goes for sentences like "humans have hair" (which is generic; *cf.* "every human has hair", which is false). These examples suggest that, when it comes to "every", genericity might not be the right generalization.

On our view, "every" is compatible with generalization because it encourages grouping the things quantified over as an ensemble. Sentences like "every tick carries Lyme disease" have no such interpretation because projection beyond the local domain is not at issue, attributing a property to a kind is. In contrast, a 'generic' understanding is present in a sentence like "every tick can carry Lyme disease", where expanding the domain to include all ticks in existence is at issue. The same can be said for a case like "every dog loves bacon", which is naturally understood as 'generic' in the relevant sense and "every one of the dogs loves bacon", which has no such interpretation (it seems to be a claim about some specific dogs), potentially because "one of the" effectively extinguishes the bias to treat the dogs as an ensemble by promoting their individuation.

In sum, the two predictions discussed here are puzzling on standard views about how to deal with differences between "each" and "every". By contrast, they follow directly from properties of the proposed interfacing non-linguistic representational systems.

4.3. Generalizing the mentalist account beyond quantifiers

As noted at the outset, we focused on quantifiers – and "each" and "every" in particular – as a case study in giving a three-pronged theory of meaning that considers formal semantic representation, related cognitive underpinnings, and corresponding contexts of use. Aside from their historical theoretical importance in formal semantics, there are at least two empirical facts that made this particular case study a fruitful one.

First, mathematically precise hypotheses about quantifier meanings can be easily stated. This is not always the case. Formally spelling out a meaning hypothesis – offering a statement of independently necessary and jointly sufficient conditions – for open-class vocabulary items like "paint" or "chair" has proven notoriously difficult, if not downright impossible (Fodor, 1998). But quantifiers represent a case where candidate hypotheses are easy to come by. In addition to the proposed representations for "each" and "every", we could imagine variants that implicate negation ('¬∃x:Frog(x)[¬Green(x)]'), or that implicate the identity relation ('{x: Frog(x)}= {x: Green(x)} ∩ {x: Frog(x)}'), or the textbook treatment which implicates the subset relation ('{x: Frog(x)}⊆{x: Green(x)}'). On the mind-external view, these representations are all different names for the same state of the world. But if they are taken to be 'psychological forms', they correspond to distinct psychological hypotheses (for an argument against the latter two representations, see Knowlton, Pietroski, Williams, Halberda, & Lidz, 2021).

Of course, this is not to say that all formally distinct specifications necessarily correspond to psychologically distinct hypotheses. A claim throughout this paper has been that the difference between the first-order (3a) – '∀x:Frog(x)[Green(x)]' – and the second-order (3b) – 'TheX:Frog(X)[∀x:X(x)[Green(x)]]' – is a psychologically substantive one despite their logical equivalence. But we doubt that the difference between the brackets in '∀x:Frog(x)[Green(x)]' and '∀x:Frog(x)(Green(x))' would correspond to any such mental distinction. This difference likely has no more cognitive significance than a change in font. Taking formalism seriously (as a mentalistic proposal) is difficult, in part because we cannot say in advance which distinctions have psychological reality and which are mere notational variants of the same 'psycho-logical form'. In the case of "each" and "every", a number of proposals had already been tested and rejected, leading to the current form of (3a) and (3b).

The second important feature of the current case study is that much is known about the interfacing non-linguistic cognitive systems (in this case, object-files and ensembles). Enough is understood about these systems to reliably derive predictions about how their engagement might be more or less pragmatically appropriate in a given context. Again, this is not universally the case. It's likely true that not enough is known about 'dog cognition' to build a research project on the mental representation of the expression "dog", even if suitable hypotheses about its formal specification were in place. To move away from toy examples, consider natural language conjunction. Expressions like "and" have received exquisite attention on the formal side – Schein (2017)'s book on the topic is over 1,000 pages long – and there are multiple formally distinct ways of specifying its meaning. But it is not clear that much is known about 'conjunction cognition', meaning it might not yet be possible to pursue the sort of project presented here for that particular case.

It is only when these two criteria are met – availability of formally explicit hypotheses and understanding of potentially related non-linguistic cognition – that the sort of three-pronged approach pursued here will be feasible. Fortunately, it is unlikely that quantifiers are the sole domain in which both of these conditions are met. Modal expressions such as *can*, *must*, and *may* represent another area rich with formal theories of semantics and the semantics-pragmatics connection (e.g., Kratzer, 1991; Papafragou, 2000; Portner, 2009), as well as a developing understanding of related cognitive systems (Phillips & Knobe, 2018). Verbs are another territory where a three-pronged meaning-cognition-pragmatics approach might be fruitfully pursued. There are various formal approaches to verbal semantics (see Williams, 2021 for review), and a growing literature on the relationship between aspects of verb meanings and event cognition (e.g., Wellwood, Hespos, & Rips, 2018; Papafragou & Grigoroglou, 2019; Ji & Papafragou, 2020; Ünal, Ji, & Papafragou, 2021). The same can be said for the notion of symmetry, which has received attention both linguistically and cognitively (e.g., Gleitman, Gleitman, Miller, & Ostrin, 1996; Partee, 2008; Gleitman, Senghas, Flaherty, Coppola, & Goldin-Meadow, 2019; Hafri, Gleitman, Landau, & Trueswell, 2023).

In these cases, formally explicit proposals about meaning that relate in a principled way to non-linguistic cognitive systems can plausibly lead to testable predictions about pragmatic contexts of use. As with quantifiers, simply viewing meaning from the perspective of how expressions relate to the mind-external world seems limiting, as it threatens to miss generalizations about how language is understood (i.e., mentally represented), how it interfaces with non-linguistic systems, and how it consequently is put to use by speakers.

4.4. Conclusions

The present results provide a case study of a three-pronged approach to meaning: linking formal semantic representations, non-linguistic cognitive systems, and details of pragmatic use. We hope to have shown that thinking of semantic meanings as finely-articulated and formally explicit mental representations and taking seriously the cognitive systems with which they interface can help explain otherwise puzzling patterns of linguistic use. Moreover, the integration of semantics, cognition, and pragmatics in this approach allows for making and testing novel predictions in a way that supports new discoveries at all three levels.

This is perhaps most obvious at the level of semantics, where the approach here supports a particular proposal about the form of the semantic representations underlying "each" and "every". This proposal differs quite drastically from the standard view in formal semantics that expressions like "each frog is green" and "every frog is green" express relations between two independent sets. Neither of the meanings proposed here are relational, and the one proposed for "each" involves no notion of set or group representations. At the level of cognition, the results reported above suggest a new direction for future work on the interfacing systems, object-files and ensembles. These systems are most often studied in the context of perception, but our results show that properties of these two

representational systems matter even in the absence of perceptual inputs. This suggests the distinction exists in mental architecture more broadly, which could present an interesting direction for future research on these non-linguistic representational systems to explore. Finally, at the level of pragmatics, the program presented here fits well with psychological approaches that treat pragmatics as a way of 'filling in the gaps' inherent in semantic representations to result in complete thoughts (see, e.g., Carston, 2008). But unlike other processes of 'filling in' that rely on encyclopedic or other kinds of shared knowledge, the approach presented here suggests that non-linguistic cognitive systems invited by particular semantic representations feed into the subtle ever-present inferences that listeners draw about what is meant by the use of an utterance in context. Leveraging this semantics-cognition-pragmatics connection can thus be an important tool in the pragmatists' arsenal.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

For helpful feedback, we thank two anonymous Cognitive Psychology reviewers, Paul Pietroski, Jeffrey Lidz, and audiences at CogSci 2022, ELM2, HSP 2022, and the New York Philosophy of Language Workshop. Funding for this project was provided by MindCORE at the University of Pennsylvania.

Appendix

Table A1
Experiment 1 stimuli (differences between conditions **bolded**).

Item	Context sentences	Target sentences
1	An astrophysics team at NASA has been studying a cluster of **four** stars. An astrophysics team at NASA has been studying a cluster of **four thousand** stars.	Based on their calculations, {each/every} star in this group has been burning for more than 20 billion years.
2	A group of well-respected biologists has recently concluded a 10-year study of **three** bird species. A group of well-respected biologists has recently concluded a 10-year study of **three hundred** bird species.	They argued that {each/every} bird they studied is born with the instinct to sing.
3	John is a civil engineer who inspected **five** bridges in the US last year. John is a civil engineer who inspected **five hundred** bridges in the US last year.	He claims that {each/every} bridge he tested holds up to 10 tons.
4	Mary is a food critic for the New York Times who has reviewed **four** bakeries' cakes. Mary is a food critic for the New York Times who has reviewed **four hundred** bakeries' cakes.	She recently wrote that {each/every} cake she's tried goes well with ice cream.
5	Chloe loves math and has read **a few** textbooks. Chloe loves math and has read **a few dozen** textbooks.	She said that {each/every} textbook she's read has practice problems for students.
6	As a professional investor, Todd owns shares in **a couple** companies. As a professional investor, Todd owns shares in **a couple hundred** companies.	He expects {each/every} stock in his portfolio to do well this year.
7	The bartender at the local tavern has made **three** martinis. The bartender at the local tavern has made **three thousand** martinis.	He said that {each/every} martini he made had an olive.
8	Suzie is a doctor who has performed **a dozen** knee replacement surgeries. Suzie is a doctor who has performed **a thousand** knee replacement surgeries.	He says that {each/every} patient walks more easily after the surgery.
9	April works at the post office and has **a handful** of old stamps. April works at the post office and has **a massive collection** of old stamps.	She explained that {each/every} stamp in her collection is inscribed with the letters USA.
10	Chris, an anthropology professor, has participated in **four** archaeological digs. Chris, an anthropology professor, has participated in **forty** archaeological digs.	He mentioned that {each/every} dig site that he's visited provides shade for the workers.
11	A team of linguists has been carefully documenting **three** languages spoken in Africa. A team of linguists has been carefully documenting **three hundred** languages spoken in Africa.	They note that {each/every} language that they've documented has at least twenty color words.
12	Beth has played in **five** chess tournaments. Beth has played in **five hundred** chess tournaments.	She tells young players that {each/every} tournament she's been a part of offers food and lodging for contestants.

Table A2
Experiment 3 stimuli (differences between conditions **bolded**).

Item	Context sentences	Target sentences
1	An astrophysics team at NASA has been studying a cluster of stars.	Based on their calculations, {each/every} star **in that cluster** has been burning for more than 20 billion years. Based on their calculations, {each/every} star **in the universe** has been burning for more than 20 billion years.
2	A group of well-respected biologists has recently concluded a 10-year study of bird species.	They argued that {each/every} bird **that they studied** is born with the instinct to sing. They argued that {each/every} bird **on the planet** is born with the instinct to sing.
3	John is a civil engineer who inspects bridges in the US.	He claims that {each/every} bridge **that he tested** holds up to 10 tons. He claims that {each/every} bridge **in North America** holds up to 10 tons.
4	Mary is a food critic for the New York Times who reviews bakeries in Manhattan.	She recently wrote that {each/every} cake **in Manhattan** goes well with ice cream. She recently wrote that {each/every} cake **in the world** goes well with ice cream.
5	Chloe loves math and has read a good number of math textbooks.	She said that {each/every} book **she's read** has practice problems for students. She said that {each/every} book **in print** has practice problems for students.
6	As a professional investor, Todd owns shares in a few different companies and has done extensive research on them.	He expects {each/every} stock **in his portfolio** to do well this year. He expects {each/every} stock **on the market** to do well this year.
7	The bartender at a local tavern made a few martinis.	He said that {each/every} martini **that he made** has an olive. He said that {each/every} martini **that's worth drinking** has an olive.
8	Suzie is a doctor who has performed knee replacement surgeries on a handful of patients.	She says that {each/every} patient **she's operated on** walks more easily after the surgery. She says that {each/every} patient **who gets a knee replacement** walks more easily after the surgery.
9	April works at the post office and has a small collection of old stamps.	She explained that {each/every} stamp **in her collection** is inscribed with the letters USA. She explained that {each/every} stamp **sold in America** is inscribed with the letters USA.
10	Chris, an anthropology professor, has participated in a few archaeological digs.	He mentioned that {each/every} dig site **that he's visited** provides shade for the workers. He mentioned that {each/every} dig site **that's run by professionals** provides shade for the workers.
11	A team of linguists has been carefully documenting the languages of Africa.	They note that {each/every} language **they've documented** has at least twenty color words. They note that {each/every} language **spoken on Earth** has at least twenty color words.
12	Beth has played in a number of American chess tournaments.	She tells young players that {each/every} tournament **she's been a part of** offers food and lodging for contestants. She tells young players that {each/every} tournament **she's heard of** offers food and lodging for contestants.

Table A3
Filler trials for Experiments 1 and 3.

Item	Context sentence	Target sentence
1	Lisa is an avid reader and has collected thousands of books over the years.	In her favorite book, the main character is a talking dog. The main character is a talking dog in her favorite book.
2	As a volunteer firefighter, Sue has been through a lot of training exercises.	They get to use the firehose for target practice on her favorite day of training. On her favorite day of training, they get to use the firehose for target practice.
3	Lauren loves growing herbs and vegetables in her yard.	In the summertime, her garden flourishes. Her garden flourishes in the summertime.
4	Jeff makes sandwiches that he sells in a deli on his town's main street.	His shop is busiest in the morning. In the morning, his shop is busiest.
5	Matt works at a ski resort in Colorado.	He said that when they visit, most guests stay near the base of the mountain. He said that most guests stay near the base of the mountain when they visit.
6	Sandy grew up on a cattle ranch in the Western US.	He tells friends that it's important to practice when first learning to lasso. He tells friends that when first learning to lasso, it's important to practice.

(*continued on next page*)

Table A3 (*continued*)

Item	Context sentence	Target sentence
7	Laurel enjoys getting together with friends and playing board games.	When deciding what move to make next, most of her friends go silent. Most of her friends go silent when deciding what move to make next.
8	As a sculptor, Sarah loves going to art museums for inspiration.	She's sure to stop at the main museum whenever she visits a new city. Whenever she visits a new city, she's sure to stop at the main museum.
9	Paula likes to recreate famous paintings in her spare time.	In her latest project, the copy is identical to the original. In her latest project, the original is identical to the copy.
10	Anna was giving Donald directions to her house.	She said that the oak tree is near her driveway. She said that her driveway is near the oak tree.
11	During the weekends, a local group of bug collectors goes to the woods to find rare insects.	It's hard because the bugs match the surroundings. It's hard because the surroundings match the bugs.
12	Aaron works in a greenhouse growing and selling plants.	He tells customers that roses are similar to tulips. He tells customers that tulips are similar to roses.
13	Calvin and Andrea run a book club and enjoy reading all genres.	In their opinion, fiction resembles nonfiction. In their opinion, nonfiction resembles fiction.
14	Fran visited Los Angeles and hoped to meet a movie star on the street.	She didn't get to, but Tom Hanks met her sister. She didn't get to, but her sister met Tom Hanks.
15	Betty loves ordering take-out on Sunday nights instead of cooking.	Her favorite restaurant is across from city hall. City hall is across from her favorite restaurant.
16	Justin recently moved to a new city far away from his hometown.	As he walks around, he can't help comparing the town with the city. As he walks around, he can't help comparing the city with the town.
17	Kathleen designs car dashboards for a large automotive company.	Some designs feature speedometers that also display speed digitally. Some of her designs feature speedometers that also display speed digitally.
18	Adam is a huge coffee snob who has tried many different roasts.	He tells his friends that some of the brews are best enjoyed black. He tells his friends that some brews are best enjoyed black.
19	Tess has been hiking the Appalachian trail for three months.	Some food she brought got wet, so she stopped for supplies. Some of the food she brought got wet, so she stopped for supplies.
20	Kathy likes to take her little sister to the local zoo on the weekends.	Some of the animals come right up to the side of their cage and greet the visitors. Some animals come right up to the side of their cage and greet the visitors.
21	Paul worked in a coal mine when he was younger.	He tells his grandchildren that some miners rode in the carts that carried coal. He tells his grandchildren that some of the miners rode in the carts that carried coal.
22	Annie loves playing poker and often gives advice to younger players.	She tells them that only some of the hands you're dealt are worth betting on. She tells them that only some hands you're dealt are worth betting on.
23	Tyler watches a lot of baseball games during the summer.	Some games have pretty exciting conclusions. Some of the games have pretty exciting conclusions.
24	Ellen often travels from America to Europe to visit her family.	Some of her trips have been by plane, but her favorite method is to go by boat. Some trips have been by plane, but her favorite method is to go by boat.

References

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences, 15*(3), 122–131. https://doi.org/10.1016/j.tics.2011.01.003

Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological Science, 19*(4), 392–398. https://doi.org/10.1111/j.1467-9280.2008.02098.x

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science, 12*(2), 157–162. https://doi.org/10.1111/1467-9280.00327

Bach, K. (1997). The semantics-pragmatics distinction: What it is and why it matters. In E. Rolf (Ed.) *Pragmatik* (pp. 33–50). VS Verlag für Sozialwissenschaften Wiesbaden. doi: 10.1007/978-3-663-11116-0_3.

Baker, M. C. (1995). On the absence of certain quantifiers in Mohawk. In E. Bach, E. Jelinek, A. Kratzer, & B. H. Partee (Eds.), *Quantification in natural languages* (pp. 21–58). Dordrecht: Kluwer.

Barker, C., & Jacobson, P. (Eds.). (2007). *Direct compositionality*. Oxford University Press.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Barwise, J., & Cooper, R. (1981). Generalized quantifiers and natural language. In *Philosophy, language, and artificial intelligence* (pp. 241–301). doi: 10.1007/BF00350139.

Beghelli, F., & Stowell, T. (1997). Distributivity and negation: The syntax of each and every. In A. Szabolcsi (Ed.), *Ways of scope taking* (pp. 71–107). doi: 10.1007/978-94-011-5814-5_3.

Beghelli, F. (1997). The syntax of distributivity and pair-list readings. In A. Szabolcsi (Ed.) *Ways of scope taking* (pp. 349–408). doi: 10.1007/978-94-011-5814-5_10.

Boeckx, C., & Hornstein, N. (2010). The varying aims of linguistic theory. In J. Franck, & J. Bricmont (Eds.), *Chomsky notebook* (pp. 115–141). Columbia University Press. https://doi.org/10.7312/bric14474-006.

Boolos, G. (1984). To be is to be a value of a variable (or to be some values of some variables). *The Journal of Philosophy, 81*(8), 430–449.

Boolos, G. (1998). *Logic, logic, and logic*. Harvard University Press.

Brendel, C. (2019). An investigation of numeral quantifiers in English. *Glossa: A Journal of General Linguistics, 4*(1). https://doi.org/10.5334/gjgl.391

Brisson, C. M. (1998). *Distributivity, maximality, and floating quantifiers*. Rutgers The State University of New Jersey-New Brunswick dissertation.

Buccola, B., Križ, M., & Chemla, E. (2022). Conceptual alternatives. *Linguistics and Philosophy*, *45*(2), 265–291. Springer. doi: 10.1007/s10988-021-09327-w.
Carey, S. (2009). *The origin of concepts*. Oxford University Press. doi: 10.1093/acprof:oso/9780195367638.001.0001.
Carston, R. (2002). *Thoughts and utterances: The pragmatics of explicit communication*. John Wiley & Sons. doi: 10.1002/9780470754603.
Carston, R. (2008). Linguistic communication and the semantics/pragmatics distinction. *Synthese, 165*(3), 321–345. https://doi.org/10.1007/s11229-007-9191-8
Carston, R. (1988). Language and cognition. In F. J. Newmeyer (Ed.), *Linguistics: the Cambridge survey* (Vol. 3, pp. 38–68). Cambridge University Press. doi: 10.1017/CBO9780511621062.003.
Chomsky, N. (1964). *Current Issues in linguistic theory*. De Gruyter Mouton. doi: 10.1515/9783110867565.
Chomsky, N. (1965). *Aspects of the theory of syntax*. MIT Press.
Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research, 43*(4), 393–404. https://doi.org/10.1016/S0042-6989(02)00596-5
Clark, H. H. (1996). *Using language*. Cambridge University Press. doi: 10.1017/CBO9780511620539.
Davidson, D. (1967). Truth and meaning. *Synthese, 17*(3), 304–323. https://doi.org/10.1007/BF00485035
Denić, M., & Szymanik, J. (2022). Are most and more than half truth-conditionally equivalent? *Journal of Semantics, 39*(2), 261–294. https://doi.org/10.1093/jos/ffab024
Fauconnier, G. (1984). *Mental spaces: Aspects of meaning construction in natural language*. Cambridge University Press.
Feigenson, L., & Carey, S. (2005). On the limits of infants' quantification of small object arrays. *Cognition, 97*(3), 295–313. https://doi.org/10.1016/j.cognition.2004.09.010
Feigenson, L., Dehaene, S., & Spelke, E. S. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*, 307–314. https://doi.org/10.1016/j.tics.2004.05.002
Fodor, J. A. (1975). *The language of thought*. Harvard University Press.
Fodor, J. A. (1998). *Concepts: Where cognitive science went wrong*. Oxford University Press.
Fox, D., & Katzir, R. (2011). On the characterization of alternatives. *Natural Language Semantics, 19*, 87–107. https://doi.org/10.1007/s11050-010-9065-3
Frege, G. (1879). Begriffsschrift. In J. van Heijenoort (Ed.), *From Frege to Gödel: A source book in mathematical logic* (pp. 1879–1931). Harvard University Press.
Frege, G. (1893/1903). *Grundgesetze der Arithmetik*. Verlag Hermann Pohle, Band I/II. Partial translation of Band I, The Basic Laws of Arithmetic, by M. Furth, University of California Press, 1964.
Gazdar, G. (1979). *Pragmatics: Implicature, presupposition, and logical form*. Academic Press.
Gleitman, L. R., Gleitman, H., Miller, C., & Ostrin, R. (1996). Similar, and similar concepts. *Cognition, 58*(3), 321–376. https://doi.org/10.1016/0010-0277(95)00686-9
Gleitman, L., Senghas, A., Flaherty, M., Coppola, M., & Goldin-Meadow, S. (2019). The emergence of the formal category "symmetry" in a new sign language. *Proceedings of the National Academy of Sciences, 116*(24), 11705–11711. https://doi.org/10.1073/pnas.1819872116
Green, E. J., & Quilty-Dunn, J. (2021). What is an object-file? *The British Journal for the Philosophy of Science, 72*(3), 665–699. https://doi.org/10.1093/bjps/axx055
Grice, P. (1967). Logic and conversation. In D. Davidson, & G. Harman (Eds.), *1975 The logic of grammar* (pp. 64–75). Dickenson: Encino.
Haberman, J., & Whitney, D. (2011). Efficient summary statistical representation when change localization fails. *Psychonomic Bulletin & Review, 18*(5), 855–859. https://doi.org/10.3758/s13423-011-0125-6
Haberman, J., & Whitney, D. (2012). Ensemble perception: Summarizing the scene and broadening the limits of visual processing. In J. Wolfe, & L. Robertson (Eds.), *From perception to consciousness: Searching with Anne Treisman* (pp. 339–349). Oxford University Press.
Hafri, A., Gleitman, L. R., Landau, B., & Trueswell, J. (2023). Where word and world meet: Language and vision share an abstract representation of symmetry. *Journal of Experimental Psychology: General, 152*(2), 509–527. https://doi.org/10.1037/xge0001283
Halberda, J., Sires, S. F., & Feigenson, L. (2006). Multiple spatially overlapping sets can be enumerated in parallel. *Psychological Science, 17*(7), 572–576. https://doi.org/10.1111/j.1467-9280.2006.01746.x
Heim, I., & Kratzer, A. (1998). *Semantics in generative grammar*. Oxford: Blackwell.
Hodges, W. (1998). Compositionality is not the problem. *Logic and Logical Philosophy, 6*, 7–33. 10.12775/LLP.1998.001.
Hodges, W. (2001). Formal features of compositionality. *Journal of Logic, Language, and Information, 10*, 7–28.
Horn, L. R. (1984). Toward a new taxonomy of pragmatic inference: Q-based and R-based implicature. In D. Schiffrin (Ed.), *Meaning, form and use in context* (pp. 11–42). Georgetown University Press.
Hornstein, N. (1984). *Logic as grammar*. MIT Press. doi: 10.7551/mitpress/4287.001.0001.
Im, H. Y., & Halberda, J. (2013). The effects of sampling and internal noise on the representation of ensemble average size. *Attention, Perception, & Psychophysics, 75*(2), 278–286. https://doi.org/10.3758/s13414-012-0399-4
Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences, 106*(25), 10382–10385. https://doi.org/10.1073/pnas.0812142106
Jackendoff, R. S. (1983). *Semantics and cognition*. MIT Press.
Jaswal, V. K., & Markman, E. M. (2007). Looks aren't everything: 24-month-olds' willingness to accept unexpected labels. *Journal of Cognition and Development, 8*(1), 93–111. https://doi.org/10.1080/15248370709336995
Ji, Y., & Papafragou, A. (2020). Is there an end in sight? Viewers' sensitivity to abstract event structure. *Cognition, 197*, Article 104197. https://doi.org/10.1016/j.cognition.2020.104197
Kahneman, D., & Treisman, A. (1984). Changing views of attention and automaticity in varieties of attention. In R. Parasuraman, & D. R. Davies (Eds.), *Varieties of attention* (pp. 29–61). Academic Press. https://doi.org/10.1016/0010-0285(92)90007-O.
Kahneman, D., Treisman, A., & Gibbs, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology, 24*(2), 175–219.
Kanjlia, S., Feigenson, L., & Bedny, M. (2021). Neural basis of approximate number in congenital blindness. *Cortex, 142*, 342–356. https://doi.org/10.1016/j.cortex.2021.06.004
Katzir, R. (2007). Structurally-defined alternatives. *Linguistics and Philosophy, 30*, 669–690. https://doi.org/10.1007/s10988-008-9029-y
Knowlton, T., & Gomes, V. (2022). Linguistic and non-linguistic cues to acquiring the strong distributivity of "each". *Proceedings of the Linguistic Society of America, 7*(1). doi: 10.3765/plsa.v7i1.5236.
Knowlton, T., & Lidz, J. (2021). Genericity signals the difference between "each" and "every" in child-directed speech. In *Proceedings of the Boston University Conference on Language Development* (Vol. 45, pp. 399–412). http://www.lingref.com/bucld/45/BUCLD45-31.pdf.
Knowlton, T., Halberda, J., Pietroski, P., & Lidz, J. (under review). *Individuals versus ensembles and "each" versus "every"*.
Knowlton, T., Hunter, T., Odic, D., Wellwood, A., Halberda, J., Pietroski, P., & Lidz, J. (2021). Linguistic meanings as cognitive instructions. *Annals of the New York Academy of Sciences., 1*, 134–144. https://doi.org/10.1111/nyas.14618
Knowlton, T., Pietroski, P., Halberda, J., & Lidz, J. (2022). The mental representation of universal quantifiers. *Linguistics and Philosophy, 45*, 911–941. https://doi.org/10.1007/s10988-021-09337-8
Knowlton, T., Pietroski, P., Williams, A., Halberda, J., & Lidz, J. (2021). Determiners are "conservative" because their meanings are not relations: Evidence from verification. *Semantics and Linguistic Theory, 30*, 206–226.
Knowlton, T., Trueswell, J., & Papafragou, A. (2022). A mentalistic semantics explains "each" and "every" quantifier use. In *Proceedings of the annual meeting of the Cognitive Science Society* (p. 44).
Knowlton, T. (2021). *The psycho-logic of universal quantifiers*. University of Maryland dissertation. doi: 10.13016/fdr8-3qqh.
Kratzer, A. (1991). Modality. In A. von Stechow, & D. Wunderlich (Eds.), *Semantics: An international handbook of contemporary research* (pp. 639–650). De Gruyter.
Lakoff, G. (1987). *Women, fire, and dangerous things: What categories reveal about the mind*. University of Chicago Press.
Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites*. Stanford University Press.
Larson, R. K. (2021). Rethinking cartography. *Language, 97*(2), 245–268. https://doi.org/10.1353/lan.2021.0018
Levinson, S. (1983). *Pragmatics*. Cambridge University Press. doi: 10.1017/CBO9780511813313.
Lewis, D. (1975). Languages and language. In K. Gunderson (Ed.), *Minnesota Studies in the philosophy of science, 7* (pp. 3–35). University of Minnesota Press.

Lidz, J., Pietroski, P., Halberda, J., & Hunter, T. (2011). Interface transparency and the psychosemantics of most. *Natural Language Semantics, 19*(3), 227–256. https://doi.org/10.1007/s11050-010-9062-6
Montague, R. (1973). The proper treatment of quantification in ordinary English. In *Approaches to natural language* (pp. 221–242). Springer. https://doi.org/10.1007/978-94-010-2506-5_10.
Pagin, P. (2003). Communication and strong compositionality. *Journal of Philosophical Logic, 32*, 287–322. https://doi.org/10.1023/A:1024258529030
Pagin, P. (2012). Communication and the complexity of semantics. In W. Werning, W. Hinzen, & E. Machery (Eds.), *The Oxford handbook of compositionality* (pp. 510–529). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199541072.013.0025.
Papafragou, A. (2000). *Modality: Issues in the semantics-pragmatics interface*. Brill. doi: 10.1163/9780585474199.
Papafragou, A., & Grigoroglou, M. (2019). The role of conceptualization during language production: Evidence from event encoding. *Language, Cognition and Neuroscience, 34*(9), 1117–1128. https://doi.org/10.1080/23273798.2019.1589540
Papafragou, A., & Schwarz, N. (2006). Most wanted. *Language Acquisition, 13*(3), 207–251. https://doi.org/10.1207/s15327817la1303_3
Partee, B. H. (1979). Semantics—mathematics or psychology?. In *Semantics from different points of view* (pp. 1–14). Springer.
Partee, B. H. (2008). Symmetry and symmetrical predicates. In *In Computational linguistics and intellectual technologies: Papers from the international conference "Dialogue"* (pp. 606–611).
Partee, B. H. (2011). Formal semantics: Origins, issues, early impact. *Baltic International Yearbook of Cognition, Logic and Communication, 6*. https://doi.org/10.4148/biyclc.v6i0.1580
Partee, B. H. (2018). Changing notions of linguistic competence in the history of formal semantics. In B. Rabern, & D. Ball (Eds.), *The science of meaning: Essays on the metatheory of natural language semantics* (pp. 172–196). Oxford University Press.
Partee, B. (1995b). Lexical semantics and compositionality. In *An invitation to cognitive science* (pp. 311–360).
Partee, B. H. (1995a). Quantificational structures and compositionality. In E. Bach, E. Jelinek, A. Kratzer, and B.H. Partee (Eds.), *Quantification in natural languages* (pp. 541–601). Springer. doi: 10.1007/978-94-017-2817-1_17.
Pelletier, F. (2004). The principle of semantic compositionality. In S. Davis, & B. Gillon (Eds.), *Semantics: a reader*. Oxford University Press.
Phillips, J., & Knobe, J. (2018). The psychological representation of modality. *Mind & Language, 33*(1), 65–94. https://doi.org/10.1111/mila.12165
Pietroski, P. M. (2003). Quantification and second-order monadicity. *Philosophical Perspectives, 17*, 259–298. https://doi.org/10.1111/j.1520-8583.2003.00011.x
Pietroski, P. M. (2004). *Events and Semantic Architecture*. Oxford University Press. doi: 10.1093/acprof:oso/9780199244300.001.0001.
Pietroski, P. (2005). Meaning before truth. In G. Preyer, & G. Peter (Eds.), *Contextualism in philosophy: Knowledge, meaning, and truth* (pp. 255–302). Clarendon Press.
Pietroski, P. (2017). Semantic internalism. In McGilvray (Ed.), *The Cambridge companion to Chomsky, 2* pp. 196–216). Cambridge University Press. https://doi.org/10.1017/9781316716694.010.
Pietroski, P. M. (2018). *Conjoining meanings: Semantics without truth values*. Oxford University Press.
Pietroski, P., Lidz, J., Hunter, T., & Halberda, J. (2009). The meaning of 'most': Semantics, numerosity and psychology. *Mind & Language, 24*(5), 554–585. https://doi.org/10.1111/j.1468-0017.2009.01374.x
Plaisier, M. A., Tiest, W. M. B., & Kappers, A. M. (2009). One, two, three, many–Subitizing in active touch. *Acta psychologica, 131*(2), 163–170. https://doi.org/10.1016/j.actpsy.2009.04.003
Portner, P. (2009). *Modality*. Oxford University Press.
Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition, 80*(1–2), 127–158. https://doi.org/10.1016/S0010-0277(00)00156-6
Pylyshyn, Z. W., & Storm, R. W. (1988). Tracking multiple independent targets: Evidence for a parallel tracking mechanism. *Spatial Vision, 3*(3), 179–197.
Riggs, K. J., Ferrand, L., Lancelin, D., Fryziel, L., Dumur, G., & Simpson, A. (2006). Subitizing in tactile perception. *Psychological Science, 17*(4), 271–272. https://doi.org/10.1111/j.1467-9280.2006.01696.x
Schein, B. (1993). *Plurals and events*. MIT Press.
Schein, B. (2017). *'And': Conjunction reduction redux*. MIT Press.
Soames, S. (1987). Logic as Grammar by Norbert Hornstein. *The Journal of Philosophy, 84*(8), 447–455. https://doi.org/10.5840/jphil198784891
Solt, S. (2016). On measurement and quantification: The case of "most" and "more than half". *Language, 92*(1), 65–100.
Sperber, D., & Wilson, D. (1986). *Relevance: Communication and cognition*. Harvard University Press.
Surányi, L. B. (2003). *Multiple operator movements in hungarian*. Utrecht University dissertation.
Sweeny, T. D., Wurnitsch, N., Gopnik, A., & Whitney, D. (2015). Ensemble perception of size in 4–5-year-old children. *Developmental Science, 18*(4), 556–568. https://doi.org/10.1111/desc.12239
Szabo, Z. (2000). Compositionality as supervenience. *Linguistics and Philosophy, 23*(5), 475–505. https://www.jstor.org/stable/25001788.
Talmy, L. (2000). *Toward a cognitive semantics*. MIT Press. doi: 10.7551/mitpress/6847.001.0001.
Talmy, L. (2019). Cognitive semantics: An overview. In C. Maienborn, K. von Heusinger, and P. Portner (Eds.), *Semantics – theories*. De Gruyter Mouton. doi: 10.1515/9783110589245-001.
Tunstall, S. (1998). *The interpretation of quantifiers: Semantics and processing*. University of Massachusetts Amherst dissertation. https://scholarworks.umass.edu/dissertations/AAI9909228.
Ünal, E., Ji, Y., & Papafragou, A. (2021). From event representation to linguistic meaning. *Topics in Cognitive Science, 13*(1), 224–242. https://doi.org/10.1111/tops.12475
Vendler, Z. (1962). Each and every, any and all. *Mind, 71*(282), 145–160.
Vogel, E. K., Woodman, G. F., & Luck, S. J. (2001). Storage of features, conjunctions, and objects in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 27*(1), 92–114. https://doi.org/10.1037/0096-1523.27.1.92
Wang, F. H., & Trueswell, J. (2022). Being suspicious of suspicious coincidences: The case of learning subordinate word meanings. *Cognition, 224*, Article 105028. https://doi.org/10.1016/j.cognition.2022.105028
Ward, E. J., Bear, A., & Scholl, B. J. (2016). Can you perceive ensembles without perceiving individuals?: The role of statistical perception in determining whether awareness overflows access. *Cognition, 152*, 78–86. https://doi.org/10.1016/j.cognition.2016.01.010
Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12-to 13-month-old infants. *Cognitive Psychology, 29*(3), 257–302. https://doi.org/10.1006/cogp.1995.1016
Wellwood, A. (2020). Interpreting degree semantics. *Frontiers in Psychology, 10*, 2972. https://doi.org/10.3389/fpsyg.2019.02972
Wellwood, A., Hespos, S. J., & Rips, L. (2018). The object: Substance: Event: Process analogy. *Oxford Studies in Experimental Philosophy, 2*, 183–212. https://doi.org/10.1093/oso/9780198815259.003.0009
Whitney, D., & Yamanashi Leib, A. (2018). Ensemble perception. *Annual Review of Psychology, 69*, 105–129. https://doi.org/10.1146/annurev-psych-010416-044232
Williams, A. (2015). *Arguments in syntax and semantics*. Cambridge University Press. doi: 10.1017/CBO9781139042864.
Williams, A. (2021). Events in semantics. In *Cambridge handbook of the philosophy of language*. Cambridge University Press. doi: 10.1017/9781108698283.021.
Wood, J. N., & Spelke, E. S. (2005). Infants' enumeration of actions: Numerical discrimination and its signature limits. *Developmental science, 8*(2), 173–181. https://doi.org/10.1111/j.1467-7687.2005.00404.x
Xu, F., & Carey, S. (1996). Infants' metaphysics: The case of numerical identity. *Cognitive Psychology, 30*(2), 111–153. https://doi.org/10.1006/cogp.1996.0005
Xu, Y., & Chun, M. M. (2009). Selecting and perceiving multiple visual objects. *Trends in Cognitive Sciences, 13*(4), 167–174. https://doi.org/10.1016/j.tics.2009.01.008
Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*(1), B1–B11. https://doi.org/10.1016/S0010-0277(99)00066-9

Xu, F., & Tenenbaum, J. B. (2007). Sensitivity to sampling in Bayesian word learning. *Developmental Science, 10*(3), 288–297. https://doi.org/10.1111/j.1467-7687.2007.00590.x

Zehr, J., & Schwarz, F. (2018) PennController for Internet Based Experiments (IBEX).

Zosh, J. M., Halberda, J., & Feigenson, L. (2011). Memory for multiple visual ensembles in infancy. *Journal of Experimental Psychology: General, 140*(2), 141–158. https://doi.org/10.1037/a0022925