

# Heuristics in Language Comprehension

Veena D. Dwivedi<sup>1,2\*</sup>, Kaitlin E. Goertz<sup>3</sup>, Janahan Selvanayagam<sup>2</sup>

<sup>1</sup>Department of Psychology, Brock University, St. Catharines, Canada

<sup>2</sup>Centre for Neuroscience, Brock University, St. Catharines, Canada

<sup>3</sup>Department of Applied Linguistics, Brock University, St. Catharines, Canada

Email: \*vdwivedi@brocku.ca

**How to cite this paper:** Dwivedi, V.D., Goertz, K.E. and Selvanayagam, J. (2018) Heuristics in Language Comprehension. *Journal of Behavioral and Brain Science*, 8, 430-446.  
<https://doi.org/10.4236/jbbs.2018.87027>

**Received:** May 25, 2018

**Accepted:** July 9, 2018

**Published:** July 12, 2018

Copyright © 2018 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

We used a sentence-picture matching task to demonstrate that heuristics can influence language comprehension. Interpretation of quantifier scope ambiguous sentences such as *Every kid climbed a tree* was investigated. Such sentences are ambiguous with respect to the number of trees inferred; either several trees were climbed or just one. The availability of the NOUN VERB NOUN (N-V-N) heuristic, e.g., KID CLIMB TREE, should contribute to the interpretation of how many trees were climbed. Specifically, we hypothesized that number choices for these stimuli would be predicted by choices previously made to corresponding (full) sentences. 45 participants were instructed to treat N-V-N triplets such as KID CLIMB TREE as telegrams and select a picture, regarding the quantity (“several” vs. “one”) associated with *tree*. Results confirmed that plural responses to quantifier scope ambiguous sentences significantly predict increased plural judgments in the picture-matching task. This result provides empirical evidence that the N-V-N heuristic, via conceptual event knowledge, can influence sentence interpretation. Furthermore, event knowledge must include the *quantity* of participants in the event (especially in terms of “several” vs. “one”). These findings are consistent with our model of language comprehension functioning as “Heuristic first, algorithmic second.” Furthermore, results are consistent with judgment and decision making in other cognitive domains.

## Keywords

Conceptual Event Knowledge, Language, Quantifier Scope, Scripts, Heuristics

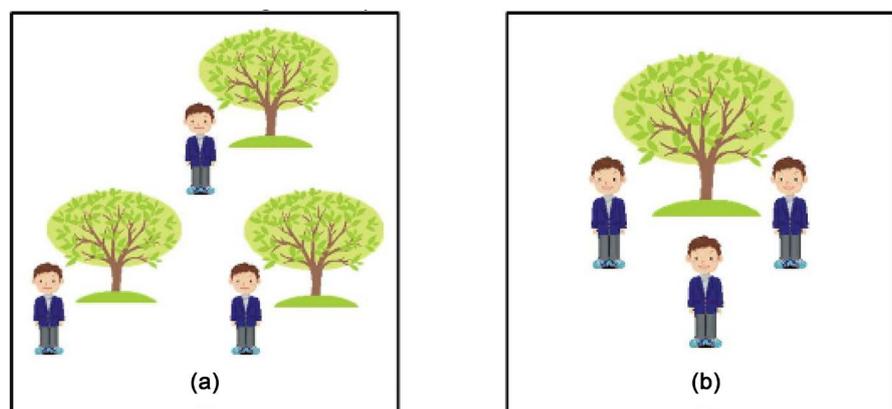
## 1. Introduction

We can interpret “DOG BITE MAN” into a particular scene or context, and furthermore, this context would be easier to understand than “MAN BITE DOG”.

Thus, in English, we can use a simple NOUN VERB NOUN (N-V-N) heuristic [1], or rule of thumb, to create events in our minds, where some events are more likely to occur than others. On the other hand, if the same words, in the same order as DOG BITE MAN, are inserted into grammatically complex sentences, such as *It was the dog that the man bit*, or *The dog was bitten by the man*, the grammar dictates that the interpretation is consistent with the unlikely scene indicated in MAN BITE DOG. It is clear that two systems are important for language comprehension, one relying on context and experience (heuristics), and the other relying on grammatical form and rules (algorithms). In previous work [2], we argued that language comprehension operates via “Heuristic first, algorithmic second” mechanisms. The mental events associated with language comprehension are analogous to those posited in [3], where heuristic computations are called System 1 and algorithmic computations are called System 2. We claim that, at first pass, the processes people use in other cognitive domains are consistent with processes used in language comprehension.

In [2], (see also [4]) sentences such as *Every kid climbed a tree* were investigated. These sentences, which exhibit quantifier scope ambiguity (QSA), are of interest because they have more than one interpretation, despite not being syntactically or lexically ambiguous. The ambiguity has to do with the quantity of entities plausibly inferred. That is, on one reading of *Every kid climbed a tree*, several trees are climbed, on another just one (see Figure 1).

Whereas previous works examining on-line sentence interpretation [5]-[11] have argued that QSA sentences are understood via mechanisms that are sensitive to abstract rules as posited in linguistic and philosophical traditions [12] [13], we have not. This is because such rules would result in an actual (plural) preference (see [14], as well as works mentioned above) which our lab has not empirically shown. That is, we have not observed a (plural) preference using on-line Event Related Potential (ERP) methods or self-paced reading methods (see [2] [4] respectively, for these findings. Also, see [2] for a review of theoretical



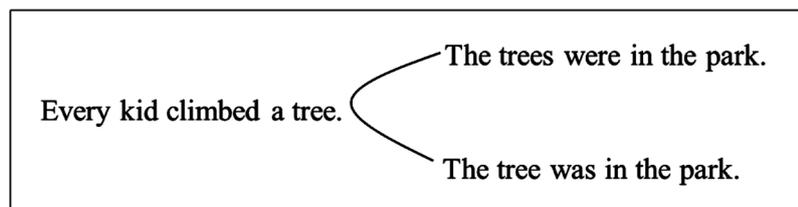
**Figure 1.** Two possible interpretations of a Quantifier Scope Ambiguous sentence. Example of two possible interpretations of Quantifier Scope Ambiguous sentence, *Every kid climbed a tree*, where (a) and (b) correspond with plural and singular interpretations of the word *tree*, respectively.

issues regarding quantifier scope).

In this work, we examine the role that numerical cognition plays in people's conceptual knowledge of events. We hypothesize that real-time sentence interpretation is not derived solely via algorithmic computation but instead from heuristic knowledge regarding events [2]. Specifically, we will claim that conceptual world knowledge contains information regarding numerosity; that is, people can estimate the quantity of participants in an event. If this is correct, then an empirical demonstration whereby variability in number interpretation for sentences is derived from the mental representation of the number of participants in a corresponding event, should follow. We pursue that demonstration here.

We build on our previous work here and take as our starting point an off-line norming task, reported in [4], and discussed in [2]. 32 participants read 160 QSA sentences (*Every kid climbed a tree*) and were asked to circle whether they preferred the (disambiguating) continuation sentence such as *The trees were in the park* versus *The tree was in the park* (see **Figure 2**). Results showed a preference for plural continuation sentences at a rate of 74%, replicating results of [14], who showed a preference at 76% (c.f. [8]).

A follow-up items analysis in [2] revealed that despite the fact that all 160 sentences were of the same syntactic and semantic form (EVERY  $N_1$  VERB(ED) A( $N_2$ )), quantity preferences varied greatly among items. For example, *Every kid climbed a tree* was interpreted as plural at 100% whereas *Every jeweller appraised a diamond* was interpreted as plural at 60% (see **Appendix A** for further stimuli and corresponding plural preferences). It was assumed that the sentences differed in their interpretation because the  $N_1$ -V- $N_2$  content, i.e., the lexical backbone of the sentences (e.g., KID CLIMB TREE vs. JEWELLER APPRAISE DIAMOND) evoked different mental representation of events, where these representations differed with respect to the number associated with  $N_2$ . Given that sentences of the form *Every kid climbed a tree* exhibit  $N_1$ -V- $N_2$  linear order, which is canonical of English sentences, we expect that people could use heuristic strategies to understand them. Based on schema theory [15], we can imagine that our experience in the world tells us that kids tend to climb several trees, but in the case of jewellers appraising diamonds, we might not have expectations regarding several diamonds<sup>1</sup>.



**Figure 2.** Example stimuli from [4] off-line questionnaire study.

<sup>1</sup>Note that on an account where a logical syntactic rule always and only applies for interpretive purposes, the algorithm's application should consistently result in the same interpretation, i.e., plural number. However, recent work by our group shows that this is not the case (for a review of the logical rule application and its empirical predictions, see [2]).

As such, it could be the case that people understand sentences by simply attending to the N-V-N sequence in a sentence, which would then activate conceptual world knowledge, assumed to be built on experience in the world, and is thus independent of grammatical considerations [1] [16] [17] [18]. As such, our work builds on previous literature that argues that event knowledge includes information about stereotypical subjects and objects, instruments, and location (see e.g., [19] [20] [21] [22]).

In the present work, we claim that conceptual world knowledge includes information regarding the quantity of participants in events. Moreover, this information is available immediately in sentence comprehension, and thus would be available in real-time language processing.

We build on recent claims that sentence comprehension can occur without any grammatical analysis; heuristic (word-based) mechanisms alone can be used [3] [18] [23] [24]. In our model of language processing [2], we claim that sentences are understood via heuristic mechanisms first, and algorithmic (*i.e.*, rule-based) processes occur second (and only if required).

In the current work, we build on the findings above by explicitly testing the assumption regarding the mental representation of conceptual events and number, using a novel sentence-picture verification task. Presently, instead of sentences, N-V-Ns evoking a conceptual script (in the form of  $N_1$ -V- $N_2$ ) were presented. Sentences from [4] were stripped of quantifiers, verbal, and nominal inflection, yielding a simple, three word  $N_1$ -V- $N_2$  skeleton. The N-V-N design was borrowed from [25], who argued that (Dutch) scripts containing three words, with no grammatical inflection, could evoke an event interpretation. In their work, the N400 component, an Event Related Potential (ERP) marker of lexical-pragmatic anomaly [26] was elicited at the final word of *VACATION TRIAL DISMISSAL* as compared to the plausible scene computed in *DIRECTOR BRIBE DISMISSAL*. As such, those authors concluded that N-V-Ns could indeed elicit script or schema interpretation [15] [17] [27].

We take this as a starting point, and in the present experiment, have participants choose a picture that best matched their interpretation of N-V-N stimuli. Participants had to respond to the final word ( $N_2$ ) in N-V-N triplet stimuli with respect to singular/plural number. That is, for the  $N_1$ -V- $N_2$  script, KID CLIMB TREE, derived from *Every kid climbed a tree*, participants had to choose a picture which had several trees or just one, in a scene with multiple kids (for details of stimuli, see Methods below).

Given this design, our predictions were straightforward. Judgments for QSA sentences in the previous experiment, regarding plural vs. singular interpretation (e.g., *Every kid climbed a tree*, *Every jeweller appraised a diamond*) should serve as significant predictors of plural vs. singular interpretations of corresponding N-V-N stimuli (e.g., KID CLIMB TREE, JEWELLER APPRAISE DIAMOND) in the current experiment.

If so, this work would show that conceptual knowledge of events not only

includes information about the nature of protagonists, location, and instruments [28] but also numerosity. In other words, the mental representation of events necessarily includes information about the quantity of participants/entities in events. In addition, while it might seem obvious that lexical factors influence variation in sentence acceptability judgments [29] [30], and is indeed assumed to be the case, to our knowledge this has yet to be empirically demonstrated. Finally, these findings would be consistent with the model proposed in [2], which posits heuristic mechanisms as primary in sentence understanding, not the use of grammatical algorithms.

## 2. Material and Methods

### 2.1. Participants

Forty-five Brock University undergraduate students (40 female, mean age 20 years, range 18 to 30 years) were recruited from February to June 2012. Participants were either paid for their participation or received partial course credit. All subjects were native speakers of English, had normal or corrected-to-normal vision and were right handed. None of the participants reported any neurological impairment, history of neurological trauma or use of neuroleptics. Also, none of them had participated in the norming task reported in [4]. Based on past experience in our lab and examples in the literature [31] [32] [33], sample size was deemed as 45 (15 participants for three lists, see below). Data collection was stopped once this was achieved.

This study received ethics approval from the Brock University Social Science Research Ethics Board (SREB) prior to the commencement of the experiment (REB 12-080). Written, informed consent was received from all participants prior to their participation in the experiment.

### 2.2. Materials

#### 2.2.1. Ambiguous Condition

Simple  $N_1$ -V- $N_2$  word triplets (e.g., KID CLIMB TREE, JEWELLER APPRAISE DIAMOND) were constructed by stripping the quantifiers and inflection from the QSA sentences used in [4] (e.g., *Every kid climbed a tree*, *Every jeweller appraised a diamond*). All stimuli were presented in black, upper-case letters in 19 pt Courier New font, vertically and horizontally centered on a white background. The presentation of these linguistic stimuli was followed by two pictures simultaneously presented on either side of the computer screen. The left side of each picture always consisted of three repeated images corresponding to  $N_1$ , consistent with more than one individual, from the original sentence which used *Every*. The right side of each picture consisted of either a single object corresponding to  $N_2$  or five repeated images of  $N_2$ . Participants were required to make a judgment regarding the number associated with  $N_2$ . We note here that we did not choose three repeated images for  $N_2$ , as we were not interested in invoking a distributive reading of the event (see [9] among others for an investigation of distributivity

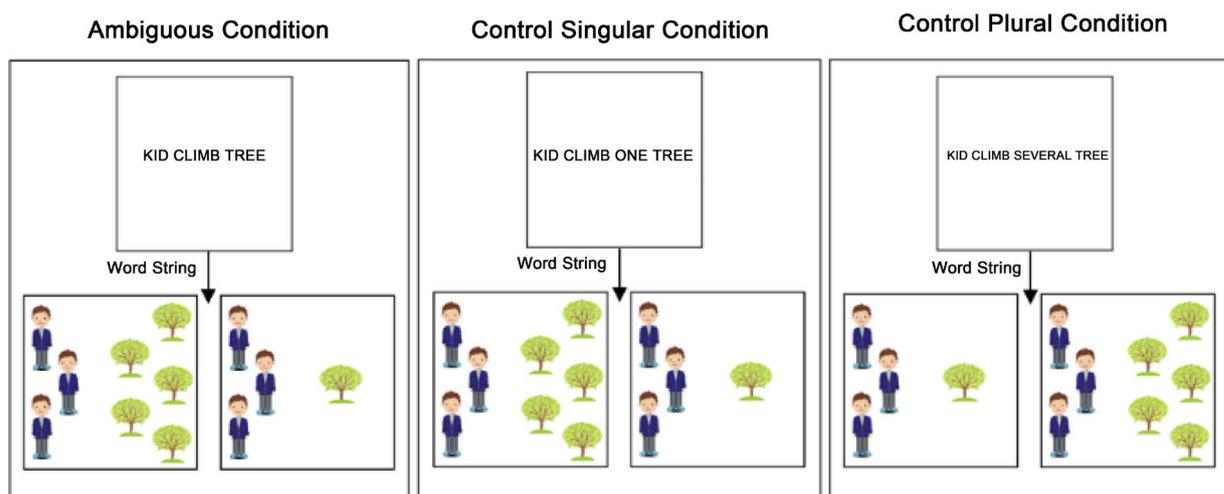
effects, an issue orthogonal to our present study). Next, four items would be numerically too close to three and might therefore involve difficulty [34] [35] [36]. Thus, five objects were chosen to correspond to a plural interpretation of  $N_2$ .

Also note that, in order to divide stimuli evenly into three lists, one of the 160 scenarios from [4] was randomly selected to be removed for the present experiment.

Images used in the pictures were found using various image databases online. **Figure 3** highlights the sequence of N-V-N and picture presentation. The position of the plural versus the singular version of each scenario was counterbalanced so that it was shown an equal number of times on the left hand side versus the right hand side of the screen.

### 2.2.2. Control Conditions

Control conditions were such that  $N_2$  was preceded by a quantifier that unambiguously signaled either singular or plural number. The form of the Control Singular condition was  $N_1$ -V-ONE- $N_2$  (e.g., KID CLIMB ONE TREE, JEWELLER APPRAISE ONE DIAMOND) and the Control Plural condition was  $N_1$ -V-SEVERAL- $N_2$  (e.g., KID CLIMB SEVERAL TREE, JEWELLER APPRAISE SEVERAL DIAMOND). These control linguistic stimuli were followed by exactly the same pictures as those in the Ambiguous condition (see **Figure 3**). See **Table 1** for a summary of experimental stimuli.



**Figure 3.** Examples of critical stimuli. An example of the ambiguous critical stimuli item, KID CLIMB TREE, for picture-matching task, and its singular (KID CLIMB ONE TREE) and plural (KID CLIMB SEVERAL TREE) control conditions.

**Table 1.** Summary of experimental stimuli used in the present study.

| Experimental Condition | Format <sup>a</sup>       | Example Stimuli        |
|------------------------|---------------------------|------------------------|
| Ambiguous              | $N_1$ -V- $N_2$           | KID CLIMB TREE         |
| Control Singular       | $N_1$ -V-"ONE"- $N_2$     | KID CLIMB ONE TREE     |
| Control Plural         | $N_1$ -V-"SEVERAL"- $N_2$ | KID CLIMB SEVERAL TREE |

<sup>a</sup>The column, Format, describes the structure of the "triplet" stimuli.  $N_1$ —first noun; V—verb;  $N_2$ —second noun.

There were 159 N-V-N scenarios for each of the three experimental conditions (Ambiguous, Control Singular, Control Plural) resulting in a total of 477 experimental stimuli. In order to reduce repetition effects, the stimuli were divided into three counterbalanced lists, such that each participant saw an equal number of conditions from each scenario. This resulted in 53 trials per experimental condition (Ambiguous, Control Singular, and Control Plural) per list, so that each participant saw 159 experimental items in total.

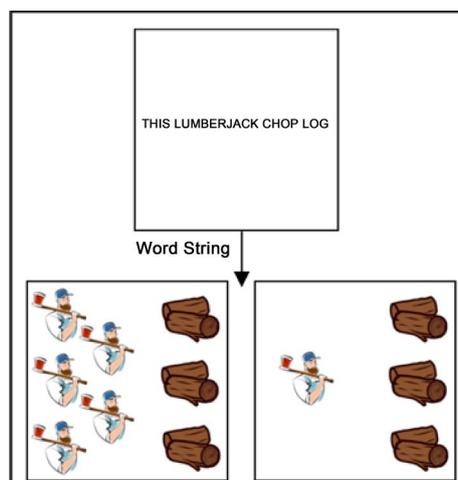
### 2.2.3. Filler Conditions

In addition to the experimental trials, there were 231 filler trials to reduce the predictability of the experimental stimuli and to reduce the chance of participants adopting meta-linguistic processing strategies (see **Figure 4** and **Figure 5**, and **Table 2**). The filler conditions served as additional controls for the experimental conditions. These controlled for type of quantifier (unambiguous plural quantifiers such as *many*, *some*, and *all* were used; unambiguous singular determiners such as *this*, *that*, and *the* were also used), and visual field (judgments on fillers would be in the left or central visual field to counterbalance critical stimuli requiring judgments in the right visual field).

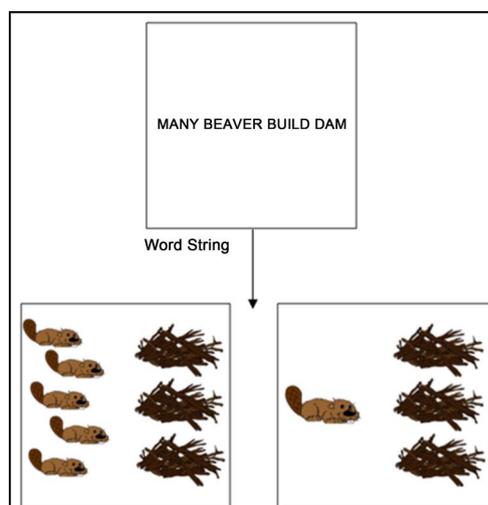
In total, each list viewed by a participant contained 390 stimuli: 159 target experimental stimuli and 231 filler trials as described above. As noted earlier, each participant saw one list only, with sentences presented in a pseudo-random fixed sequence using the program, Mix [37] with the stipulation that no two trials from the same experimental condition or filler condition followed each other. See **Appendix B** for a full list of stimuli.

### 2.3. Procedure

Informed consent was obtained from each participant before the experiment began. All participants completed a short demographics survey on handedness and reading preferences and a short computerized test of working memory



**Figure 4.** Example of filler singular stimuli. An example of the filler singular stimuli item, THIS LUMBERJACK CHOP LOG.



**Figure 5.** Example of filler plural stimuli. Example of filler plural stimuli item, MANY BEAVER BUILD DAM.

**Table 2.** Summary of filler stimuli used in the present experiment.

| Filler Condition   | Format <sup>a</sup>   | Example  |
|--|---|--|
| Filler, Singular Determiner ( <i>the, this, that</i> )                       | DET-N <sub>1</sub> -V-N <sub>2</sub><br>or N <sub>1</sub> -V-DET-N <sub>2</sub> | THIS LUMBERJACK CHOP LOG;<br>NANNY MAKE THAT BREAKFAST |
| Filler, Plural Quantifier Determiner<br>( <i>all, many, two, four, six</i> ) | Q-N <sub>1</sub> -V-N <sub>2</sub><br>or N <sub>1</sub> -V-Q-N <sub>2</sub>     | MANY BEAVER BUILD DAM;<br>BANDIT ROB ALL TRAIN         |
| Content Filler, Left Visual Field  | N <sub>1</sub> -V-N <sub>2</sub>  | MAN ANGER WIFE   |
| Content Filler, Central Visual Field   | Q(NUMERAL)-N-V  | TEN FAX ARRIVE   |

<sup>a</sup>The column, Format, describes the structure of the “triplet” stimuli. DET—determiner; Q—quantifier; Q(NUMERAL)—numeric quantifier; N<sub>1</sub>—first noun; V—verb; N<sub>2</sub>—second noun.

(operation span task [38] before beginning the present experiment. Each participant was then seated in front of a computer monitor that displayed the experiment. Participants were presented with instructions that outlined the task (forced choice sentence-picture matching). The experiment was presented to participants electronically using E-Prime 1.2 software [39]. Instructions informed participants that they would be viewing examples of telegrams, which would immediately be followed by two pictures, and that they would have to choose which of the two pictures best described the telegram. The instructions were presented in black, 14 pt Courier New font, horizontally and vertically centered on a white background. See **Appendix C** for details of experimental instruction.

Participants were then given five examples of the task as practice trials before beginning the experiment, which always began with six non-critical stimuli (*i.e.*, fillers). For each trial, participants saw a fixation cross in the middle of the screen for 500 ms, and then N-V-Ns were presented at a fixed duration of 1000 ms. The pictures appeared immediately after the N-V-N disappeared, and remained on the screen until the participant responded. Participants were instructed to

respond to the pictures using an E-Prime stimulus response box [39], and to press the button labelled “L” if they thought that the picture on the left correctly conveyed the telegram message and to press the button labelled “R” if they thought that the picture on the right correctly conveyed the telegram message. Stimuli were presented in two blocks, each containing 195 stimuli. After the first block, participants were given the opportunity to take a short break and then they were presented with the second block. The entire duration of the experiment, including the preliminary consent, the demographics survey, the working memory test, and the main experiment, was approximately 45 minutes.

## 2.4. Data Analysis

Repeated measures ANOVA were conducted for mean accuracy rates and response times, using IBM SPSS, version 20.0 [40]. We report all significant effects at the 0.05 level, using the mean square error terms from the analysis by participants. Effect size is reported using partial eta squared,  $\eta_p^2$ .

A paired samples t-test was performed to examine apparent differences between word frequencies of singular and plural variations of N<sub>2</sub> words.

Following study completion, it was recognized that items in the Filler Singular condition including the determiner THE (e.g., THE SENIOR WATCH TELEVISION) should not be included in analyses for singular interpretation, since this determiner does not unambiguously indicate singular number.

### Binary Response Data

Binary response data analyses were carried out using the statistical software R (version 3.1.0, [41]). First, we analyzed number inference in the Ambiguous condition using a logistic regression analysis. The log odds (logit) of decisions made in the current picture matching study (plural vs. singular) for the Ambiguous condition was modelled using the norming data from the previous questionnaire study on QSA sentence interpretation as our independent predictor. This analysis was performed using lmer (package “lme4” [42]; *p*-values were estimated using the lmerTest package [43]).<sup>2</sup>

We analyzed our data by modeling responses using a logit mixed-effect model [44]. Starting from the null model, including only our binomial dependent variable (plural picture responses to N-V-N stimuli) and participants as a random factor, we used the glmer function (package “lme4” [42] to analyze the improvement of the model after the predictor variable (plural sentence responses to sentence stimuli) was added. The R formula used was as follows: glmer((plural

---

<sup>2</sup>Inspection of the norming data of the quantifier scope sentences (eg, *Every kid climbed a tree*) revealed that very few full sentence items (namely, 16 out of 160 or 10% of the items) from [2] [4] exhibited a proportion plural bias at less than 40%. Given this low level of bias, in addition to the low number of items, we did not expect that including these items in the model would add predictive value for judgments in the present N-V-N picture matching study. Our expectation that including items exhibiting a plural bias of 40% and below would add no predictive value for judgments in the present study was confirmed via a piecewise binary logistic regression (R package “segmented” [50] [51]). As such, the independent variable was defined as norming data from the previous quantifier scope study with items that were biased for the plural judgment at 40% and above (range 40% - 100%).

picture responses to N-V-N stimuli) ~ (plural sentence responses to sentence stimuli) + (1|Participants in N-V-N study), data = data, family = “binomial”).<sup>3</sup>

We also analyzed the odds of plural number inference in Ambiguous vs Control Singular conditions. The analysis was a logistic regression with the following formula:  $\text{glm}(\text{Number Judgment} \sim \text{Condition}, \text{data} = \text{data}, \text{family} = \text{“binomial”})$  and  $p$ -values were estimated using lmerTest package [43].

Finally, we analyzed the accuracy in Control Singular vs Control Plural conditions in a logistic regression with the following formula:  $\text{glm}(\text{Number Judgment} \sim \text{Condition}, \text{data} = \text{data}, \text{family} = \text{“binomial”})$  and  $p$ -values were estimated using lmerTest package [43].

### 3. Results

#### 3.1. Accuracy and Response Times

Given the novelty of the current paradigm, mean accuracy rates by participant for Control conditions and response times for all critical conditions (in ms) are first examined in order to establish that participants were able to perform the task correctly (see Table 3).

The high accuracy rates for both Control Singular and Control Plural conditions indicate the success of this novel paradigm—participants were able to perform the task appropriately regarding number inference and picture matching. That is, while it could be argued that the plural picture scenario does not rule out the single-tree interpretation, the fact that participants were able to distinguish between these unambiguously marked number conditions shows that they were indeed responsive to the numerical contrast in the experiment (for further evidence of this, see complete filler results in Appendix D which also indicate high accuracy). In addition, participants were clearly sensitive to the ambiguity present in the Ambiguous conditions; RTs for this condition were 425 ms and 335ms longer than Control Singular and Control Plural conditions, respectively ( $F(2, 88) = 143.4, MSE = 18,058, p < 0.001, \eta_p^2 = 0.765$ ).

#### 3.2. Logistic Regression Analysis of Ambiguous N-V-N Responses

Next, we report results directly relevant to our hypothesis regarding plural picture

**Table 3.** Mean accuracy (%), response times (RT) and proportion plural responses (%) for critical conditions.

| Condition        | Accuracy (%) | RT (ms)    | Proportion Plural Judgments (%) |
|------------------|--------------|------------|---------------------------------|
|                  | Mean (MSE)   | Mean (MSE) | Mean (MSE)                      |
| Ambiguous        | ---          | 1541 (63)  | 7.5 (1.0)                       |
| Control Singular | 97 (0.4)     | 1112 (46)  | 2.7 (0.4)                       |
| Control Plural   | 89 (1.4)     | 1213 (47)  | 89.0 (1.4)                      |

<sup>3</sup>Note that the formula does not include word frequency as a random effect. Effectively, word frequency is a quantitative measure of the real-world experience with particular lexical items. Since the question we are asking is whether responses to N-V-Ns can be predicted by sentences that contain those very same lexical items, if we control for word frequency, we would be taking out a fundamental component of the factor that we are interested in modeling.

choices in the current experiment. Results revealed that responses to items from the previous quantifier norming study did serve as a significant predictor of plural judgments in the present experiment ( $b = 1.46$ ,  $SE = 0.60$ ,  $z = 2.41$ ,  $p = 0.02$ ). Thus, according to the present model, a greater proportion of plural responses made to sentences in the previous experiment predicts a greater likelihood of plural picture choice to a corresponding N-V-N in the current experiment. The odds of such a choice are 4.31 times ( $=e^{1.46}$ ) greater for a one-unit increase in plural response to a sentence in the previous experiment (odds ratio,  $OR = 4.31$ ,  $95\% CI = [1.30, 14.64]$ ). Thus, number interpretation to sentential stimuli does serve as a predictor of number interpretation to (conceptual event) N-V-N stimuli.

### 3.3. Other Analyses

Next, we note that we had no other *a priori* hypotheses in the current experiment. We recognize that the plural picture choices for the Ambiguous N-V-N condition in the present experiment are in the opposite direction as compared to responses to quantifier ambiguous sentences. Given that the N-V-Ns had no inflection, this is not surprising. Participants favoured singular interpretations in the current experiment, since in English, plural is overwhelmingly marked via *-s* inflection. Without it, nouns are likely interpreted as singular. Furthermore, plural pictures are necessarily visually more complex than the singular pictures. Thus, at the face of it, a complete lack of inflection (which would heavily bias towards a singular interpretation), along with the less visually complex choice of a singular picture, would explain the bias for singular choices found in the current experiment. That being said, it is worth pointing out that the plural picture choices for the Ambiguous condition were still significantly higher than those for the Control Singular condition ( $b = -1.07$ ,  $SE = 0.15$ ,  $z = -7.17$ ,  $p < 0.001$ ). This suggests, importantly, that participants performed a different number inference for Ambiguous vs. Control Singular conditions. Next, we examine differences in word frequencies between singular vs plural words as a way to understand the bias for singular found in the current experiment.<sup>4</sup>

### 3.4. Word Frequency

Relative log word frequencies of singular  $N_2$  variations ( $M = 0.58$ ,  $SD = 0.05$ ) were found to be significantly greater than those of plural  $N_2$  variations ( $M = 0.42$ ,  $SD = 0.05$ ), resulting in a significant mean difference of 0.16 ( $t = 18.94$ ,  $df = 157$ ,  $p < 0.001$ ,  $95\% CI = [0.15, 0.18]$ ).<sup>5</sup>

<sup>4</sup>Given the novelty of the paradigm, we have included all analyses for filler items in **Appendix D**. These are not discussed here, so as to not detract from the question at hand.

<sup>5</sup>The word frequencies of singular and plural variations of all  $N_2$  words (e.g. “tree” and “trees” in KID CLIMB TREE) in the experimental stimuli were collected using the SUBTLEX American Word Frequency Database [52]. The relative singular and plural word frequencies were calculated for each  $N_2$  word by dividing the singular and plural log word frequency by the sum of both the plural and singular log word frequencies, respectively [53].

Thus, given the fact that the singular form of the words was significantly higher in terms of word frequency, it is not surprising that, overall, a very strong singular bias was found in the current experiment, which used uninflected N-V-Ns. Further evidence of the stark singular bias is revealed via analyses regarding the Control Singular vs. Control Plural conditions. Although accuracy rates for both Control conditions were close to ceiling (see **Table 3**), Control Singular still exhibited a significantly higher rate of accuracy ( $b = 1.5$ ,  $SE = 0.14$ ,  $z = 10.54$ ,  $p < 0.001$ ), confirming the very strong bias for singular picture choices.

#### 4. Discussion

In the current study, a novel conception of the sentence-picture matching task was employed—instead of full sentences, N-V-N triplets were presented. People were asked to choose the picture that matched their interpretation regarding the number of entities inferred for the final item of the triplet. Participants' near ceiling accuracy rates for unambiguous control conditions, as well as their increased response times for ambiguous conditions indicated that they were indeed sensitive to the numerical contrast in the experiment, and that they performed this novel task correctly [35].

Despite a strong bias for choosing a singular interpretation, binary logistic regression analyses revealed that plural interpretation for corresponding sentences did serve as significant predictors of plural picture choices to N-V-N stimuli. The strong singular bias is explained due to the nature of the linguistic stimuli (in English, nouns lacking inflection are interpreted as singular; in addition, singular words were found to be lexically more frequent than plural stimuli), and due to the nature of the current experimental task (choosing between pictures that contained more vs. fewer items).

At the face of it, the findings support the long-held assumption (however, yet to be empirically demonstrated) that variation in sentence interpretation can be explained via lexically based factors.

We interpret the cognitive significance of the predicted statistical effect as follows: these findings indicate that our mental representation of conceptual event knowledge must include numerical information regarding quantity of participants in the event. That is, initial comprehension of QSA sentences of the form *Every kid climbed a tree* consists of a fast-and-frugal interpretation regarding the likely number of participants in the ( $N_1$ -V- $N_2$ ) event, and does not rely exclusively on algorithmic rules for interpretation. These effects are consistent with our recent model that language processing proceeds along a “Heuristic first, algorithmic second mechanism”. We further discuss these issues below.

##### 4.1. The Nature of Conceptual Event Knowledge

This study sheds light on the question of whether people have expectations regarding the likely number of entities in a conceptual event, to date an unexplored aspect of schematic knowledge. That is, it is well-known that understanding

language relies on the mental representation of world knowledge [15] [16] [17] [27]. Over the years, different properties of events have been posited: usually, these are due to the conceptual knowledge associated with specific verbs. It has been shown that people have expectations regarding the nature of typical agents, themes, instruments, and locations, associated with schematic representations [19] [22] [45]. In the present work, another stereotypical property of events is posited: the quantity of objects associated with the event. In other words, people have intuitions about whether certain events involve several objects, or not. This conceptual event information, containing numerical information, informs sentence interpretation.

#### 4.2. Language Processing as “Heuristic First, Algorithmic Second”

The present results are of theoretical importance as they call into question the dominant psycholinguistic perspective that algorithmic syntactic processing drives semantic interpretation of these and other sentences. Instead, the present results show that—at least under certain circumstances—interpretive processes need not include syntactic algorithms at all. That is, for the constructions examined here, experience trumps grammar (c.f. [1] [18] [23] [46] [47] [48]). Regarding QSA sentences, it is assumed that the grammatical/algorithmic rule is the procedure of ordering quantifiers at an abstract level (e.g., Logical Form, see [12]) for sentence interpretation. Results here indicate that these algorithmic mechanisms do not solely (if at all) determine number interpretation. Furthermore, these findings call into question the notion that initial processes in sentence interpretation are informationally encapsulated from contextual influences [8] [49]. The present findings, along with those presented in [2], show that initial processing of sentences can proceed by interpreting relevant lexical items that contribute to event interpretation.

It is important to note that we are *not* claiming that the determiners “every” and “a” play no role in the interpretation of QSA sentences. The stark differences found in interpretation of QSA sentences vs. their corresponding  $N_1$ -V- $N_2$  triplets clearly attests to the important contribution of these determiners. In addition, we note that the heuristic first mechanisms in use for sentences such as *Every kid climbed a tree* would not be in use for more complicated sentences of the form *Every kid climbed at least five trees*, which is logically equivalent to *No kid climbed less than five trees*<sup>6</sup>. These latter sentences would immediately invoke algorithmic mechanisms for comprehension (or System 2, in Kahneman’s terms), due to their complexity. It is our contention that many psycholinguistic experiments have in fact examined highly difficult sentence forms (such as centre-embedded relative clauses, for example). It would make sense if algorithmic processes immediately applied to those sentences, however, those are not the sorts of sentences encountered in day-to-day conversation. Thus, while it is true that language processing is automatic and occurs without effort—in *day to day*

---

<sup>6</sup>Thanks to a previous anonymous reviewer for this insight.

*conversation*, many sentences can indeed be understood in a shallow manner. The claim we make here (and previously) is that even certain QSA sentences may be understood in this way.

## 5. Conclusions

In sum, using logistic regression, this work empirically demonstrated that variation in sentence interpretation with respect to number, for identical syntactic constructions, is a function of lexical factors that contribute to event interpretation. Thus, our findings reveal that conceptual event interpretation must include information regarding the number of participants in a script.

Overall, the results support our recent model of language processing in that it proceeds along a “Heuristic first, algorithmic second mechanism”. That is, our findings confirm that the numerical interpretation of QSA sentences of the form *Every kid climbed a tree* can rely on heuristic processes dependent on event representations, rather than exclusively relying on algorithmic rules. Therefore, it challenges the dominant view that these sentences are primarily interpreted via logical algorithmic rules.

## Acknowledgements

Thanks to Raechelle Gibson, Dilani Balasubramaniam, and Erin Murphy for stimuli preparation assistance, also to Raechelle Gibson for help in experimental set-up. We are grateful to Gary Libben, Harald Baayen, Joffre Mercier, Matthew Walenski and Louis Schmidt for comments. Thanks also to Victoria Witte for manuscript assistance.

## References

- [1] Bever, T.G. (1970) The Cognitive Basis for Linguistic Structure. In: Hayes, J.R., Ed., *Cognition and the Development of Language*, Wiley & Sons, New York, 279-362.
- [2] Dwivedi, V.D. (2013) Interpreting Quantifier Scope Ambiguity: Evidence of Heuristic First, Algorithmic Second Processing. *PLoS ONE*, **8**, 1-20. <https://doi.org/10.1371/journal.pone.0081461>
- [3] Kahneman, D. (2011) *Thinking, Fast and Slow*. Anchor Canada, Toronto.
- [4] Dwivedi, V.D., Phillips, N.A., Einagel, S. and Baum, S.R. (2010) The Neural Underpinnings of Semantic Ambiguity and Anaphora. *Brain Research*, **1311**, 93-109. <https://doi.org/10.1016/j.brainres.2009.09.102>
- [5] Bott, O. and Schlotterbeck, F. (2015) The Processing Domain of Scope Interaction. *Journal of Semantics*, **32**, 39-92. <https://doi.org/10.1093/jos/fft015>
- [6] Filik, R., Paterson, K.B. and Liversedge, S.P. (2004) Processing Doubly Quantified Sentences: Evidence from Eye Movements. *Psychonomic Bulletin and Review*, **11**, 953-959. <https://doi.org/10.3758/BF03196727>
- [7] Paterson, K.B., Filik, R. and Liversedge, S.P. (2008) Competition during the Processing of Quantifier Scope Ambiguities: Evidence from Eye Movements during Reading. *Quarterly Journal of Experimental Psychology*, **61**, 459-473. <https://doi.org/10.1080/17470210701255317>
- [8] Raffray, C.N. and Pickering, M.J. (2010) How Do People Construct Logical Form

- during Language Comprehension? *Psychological Science*, **21**, 1090-1097.  
<https://doi.org/10.1177/0956797610375446>
- [9] Patson, N.D. and Warren, T. (2010) Evidence for Distributivity Effects in Comprehension. *Journal of Experimental Psychology: Learning Memory and Cognition*, **36**, 782-789. <https://doi.org/10.1037/a0018783>
- [10] Anderson, C. (2004) The Structure and Real-Time Comprehension of Quantifier Scope Ambiguity. Ph.D. Dissertation, Northwestern University, Evanston, IL.
- [11] Tunstall, L. (1998) The Interpretation of Quantifiers: Semantics & Processing. Ph.D. Dissertation, University of Massachusetts Amherst, Amherst, MA.
- [12] May, R. (1985) Logical Form: Its Structure and Derivation. MIT Press, Cambridge.
- [13] Russell, B. (1905) On Denoting. *Mind*, **14**, 479-493.  
<https://doi.org/10.1093/mind/XIV.4.479>
- [14] Kurtzman, H.S. and MacDonald, M.C. (1993) Resolution of Quantifier Scope Ambiguities. *Cognition*, **48**, 243-279. [https://doi.org/10.1016/0010-0277\(93\)90042-T](https://doi.org/10.1016/0010-0277(93)90042-T)
- [15] Rumelhart, D. (1980) Schemata: The Building Blocks of Cognition. In: Spiro, R., Bruce, B. and Brewer, W., Eds., *Theoretical Issues in Reading Comprehension*, Erlbaum Associates, Mahway, 33-58.
- [16] Van Dijk, T.A. and Kintsch, W. (1983) Strategies of Discourse Comprehension. Academic Press, New York, 413.
- [17] Schank, R.C. and Abelson, R.P. (1977) Scripts, Plans, Goals and Understanding. Erlbaum Associates, Mahway.
- [18] Townsend, D.J. and Bever, T.G. (2001) Sentence Comprehension: The Integration of Rules and Habits. MIT Press, Cambridge.
- [19] Altmann, G.T.M. and Kamide, Y. (1999) Incremental Interpretation at Verbs: Restricting the Domain of Subsequent Reference. *Cognition*, **73**, 247-264.  
[https://doi.org/10.1016/S0010-0277\(99\)00059-1](https://doi.org/10.1016/S0010-0277(99)00059-1)
- [20] Ferretti, T.R., Kutas, M. and McRae, K. (2007) Verb Aspect and the Activation of Event Knowledge. *Journal of Experimental Psychology: Learning Memory and Cognition*, **33**, 182-196. <https://doi.org/10.1037/0278-7393.33.1.182>
- [21] Graesser, A.C., Singer, M. and Trabasso, T. (1994) Constructing Inferences during Narrative Text Comprehension. *Psychological Review*, **101**, 371-395.  
<https://doi.org/10.1037/0033-295X.101.3.371>
- [22] McRae, K., Spivey-Knowlton, M.J. and Tanenhaus, M.K. (1998) Modeling the Influence of Thematic Fit (and Other Constraints) in On-Line Sentence Comprehension. *Journal of Memory and Language*, **38**, 283-312.  
<https://doi.org/10.1006/jmla.1997.2543>
- [23] Ferreira, F. (2003) The Misinterpretation of Noncanonical Sentences. *Cognitive Psychology*, **47**, 164-203. [https://doi.org/10.1016/S0010-0285\(03\)00005-7](https://doi.org/10.1016/S0010-0285(03)00005-7)
- [24] Gigerenzer, G. (2000) Adaptive Thinking: Rationality in the Real World. Oxford University Press, New York.
- [25] Chwilla, D.J. and Kolk, H.H.J. (2005) Accessing World Knowledge: Evidence from N400 and Reaction Time Priming. *Cognitive Brain Research*, **25**, 589-606.  
<https://doi.org/10.1016/j.cogbrainres.2005.08.011>
- [26] Kutas, M. and Hillyard, S. (1980) Reading Senseless Sentences: Brain Potentials Reflect Semantic Incongruity. *Science*, **207**, 203-205.  
<https://doi.org/10.1126/science.7350657>
- [27] Zwaan, R.A. (1999) Situation Models: The Mental Leap into Imagined Worlds.

*Current Directions in Psychological Science*, **8**, 15-18.

<https://doi.org/10.1111/1467-8721.00004>

- [28] Zwaan, R.A. and Radvansky, G. (1998) Situation Models in Language Comprehension and Memory. *Psychological Bulletin*, **123**, 162-185.  
<https://doi.org/10.1037/0033-2909.123.2.162>
- [29] Dabrowska, E. (2010) Naive v. Expert Intuitions: An Empirical Study of Acceptability Judgments. *Linguistic Review*, **27**, 1-23. <https://doi.org/10.1515/tlir.2010.001>
- [30] Caramazza, A. and Zurif, E.B. (1976) Dissociation of Algorithmic and Heuristic Processes in Language Comprehension: Evidence from Aphasia. *Brain and Language*, **3**, 572-582. [https://doi.org/10.1016/0093-934X\(76\)90048-1](https://doi.org/10.1016/0093-934X(76)90048-1)
- [31] Aaronson, D. and Ferres, S. (1984) Reading Strategies for Children and Adults: Some Empirical Evidence. *Journal of Verbal Learning and Verbal Behavior*, **23**, 189-220. [https://doi.org/10.1016/S0022-5371\(84\)90137-3](https://doi.org/10.1016/S0022-5371(84)90137-3)
- [32] Just, M.A., Carpenter, P.A. and Woolley, J.D. (1982) Paradigms and Processes in Reading Comprehension. *Journal of Experimental Psychology General*, **111**, 228-238.  
<https://doi.org/10.1037/0096-3445.111.2.228>
- [33] Ferreira, F. and Clifton, C. (1986) The Independence of Syntactic Processing. *Journal of Memory and Language*, **25**, 348-368.  
[https://doi.org/10.1016/0749-596X\(86\)90006-9](https://doi.org/10.1016/0749-596X(86)90006-9)
- [34] Henik, A., Leibovich, T., Naparstek, S., Diesendruck, L. and Rubinsten, O. (2012) Quantities, Amounts, and the Numerical Core System. *Frontiers in Human Neuroscience*, **5**, 186. <https://doi.org/10.3389/fnhum.2011.00186>
- [35] Van Opstal, F., Santens, S. and Ansari, D. (2012) The Numerate Brain: Recent Findings and Theoretical Reviews on the Neurocognitive Foundations of Number Processing. *Frontiers in Human Neuroscience*, **6**, 2011-2012.  
<https://doi.org/10.3389/fnhum.2012.00201>
- [36] Feigenson, L., Dehaene, S. and Spelke, E. (2004) Origins and Endpoints of the Core Systems of Number. Reply to Fias and Verguts. *Trends in Cognitive Sciences*, **8**, 448-449. <https://doi.org/10.1016/j.tics.2004.08.010>
- [37] Van Casteren, M. and Davis, M.H. (2006) Mix, a Program for Pseudorandomization. *Behavior Research Methods*, **38**, 584-589. <https://doi.org/10.3758/BF03193889>
- [38] Unsworth, N., Heitz, R.P., Schrock, J.C. and Engle, R.W. (2005) An Automated Version of the Operation Span Task. *Behavior Research Methods*, **37**, 498-505.  
<https://doi.org/10.3758/BF03192720>
- [39] Psychology Software Tools (2006) E-Prime (Version 1.2). Psychology Software Tools, Inc., Pittsburgh.
- [40] IBM Corp. (2011) SPSS (Version 20.0). IBM Corp., Armonk.
- [41] R Core Team (2013) R: A Language and Environment for Statistical Computing (Version 3.1.0). R Foundation for Statistical Computing, Vienna.
- [42] Bates, D., Mächler, M., Bolker, B. and Walker, S. (2015) Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, **67**, 1-48.  
<https://doi.org/10.18637/jss.v067.i01>
- [43] Kuznetsova, A., Brockhoff, P.B. and Christensen, R.H.B. (2016) Tests in Linear Mixed Effects Models. R Package Version, 2.
- [44] Jaeger, T.F. (2008) Categorical Data Analysis: Away from ANOVAs (Transformation or Not) and towards Logit Mixed Models. *Journal of Memory and Language*, **59**, 434-446. <https://doi.org/10.1016/j.jml.2007.11.007>

- [45] Ferretti, T.R., McRae, K. and Hatherell, A. (2001) Integrating Verbs, Situation Schemas, and Thematic Role Concepts. *Journal of Memory and Language*, **44**, 516-547. <https://doi.org/10.1006/jmla.2000.2728>
- [46] Caplan, D., Hildebrandt, N. and Waters, G.S. (1994) Interaction of Verb Selectional Restrictions, Noun Animacy and Syntactic Form in Sentence Processing. *Language and Cognitive Processes*, **9**, 549-585. <https://doi.org/10.1080/01690969408402131>
- [47] Christianson, K., Hollingworth, A., Halliwell, J.F. and Ferreira, F. (2001) Thematic Roles Assigned along the Garden Path Linger. *Cognitive Psychology*, **42**, 368-407. <https://doi.org/10.1006/cogp.2001.0752>
- [48] Ferreira, F. and Patson, N.D. (2007) The “Good Enough” Approach to Language Comprehension. *Language and Linguistics Compass*, **1**, 71-83. <https://doi.org/10.1111/j.1749-818X.2007.00007.x>
- [49] Fodor, J.A. (1983) *The Modularity of Mind*. MIT Press, Cambridge.
- [50] Muggeo, V.M.R. (2003) Estimating Regression Models with Unknown Break-Points. *Statistics in Medicine*, **22**, 3055-3071. <https://doi.org/10.1002/sim.1545>
- [51] Muggeo, V.M.R. (2008) Segmented: An R Package to Fit Regression Models with Broken-Line Relationships. *R News*, **8**, 20-25.
- [52] Brysbaert, M. and New, B. (2009) Moving beyond Kučera and Francis: A Critical Evaluation of Current Word Frequency Norms and the Introduction of a New and Improved Word Frequency Measure for American English. *Behavior Research Methods*, **41**, 977-990. <https://doi.org/10.3758/BRM.41.4.977>
- [53] Magnuson, J.S., Dixon, J.A., Tanenhaus, M.K. and Aslin, R.N. (2007) The Dynamics of Lexical Competition during Spoken Word Recognition. *Cognitive Science*, **31**, 133-156. <https://doi.org/10.1080/03640210709336987>

## Appendix

**Appendix A. Norming study stimuli and values.**

**Appendix B. Experimental stimuli lists.** Lists of experimental stimuli and corresponding picture pairs shown to participants in the present study.

**Appendix C. Experiment instructions.** A detailed description of experiment instructions, including verbatim text, as they were shown to participants in the present study.

**Appendix D. Filler data analysis.** Mean accuracy (%), response times (RT, ms), and proportion plural responses (%) by filler condition along with detailed discussion.

<https://www.researchgate.net/project/the-neural-underpinnings-of-semantic-ambiguity>