Running head: CONTEXTS OF NEGATIVE SENTENCES

A Pragmatic Account of the Processing of Negative Sentences

Ann E. Nordmeyer and Michael C. Frank

Department of Psychology, Stanford University

Abstract

Introduction

The word "no" is a basic linguistic phenomenon, universal in natural languages (Dahl, 2010) and seen early in children's linguistic development (Bloom, 1970, 1993; Pea, 1980). Imagine a world in which our cognitive and linguistic abilities did not include negation. For one thing, this kind of thought experiment would be impossible in such a world, since without negation we would not be able to reason about things that do not exist. Negation is enormously important to our ability to think and form complex thoughts about the world. Even the most simple forms of logic include operations to negate propositions. In addition, negation allows us to communicate about a wide range of concepts. It is negation that allows us to communicate to others about nonexistence, to reject things that we do not want, and to deny statements that are false. While previous work has explored how quickly people are able to evaluate the truth-value of negative sentences, there has been less research on how negation is processed in a communicative setting, and how contextual factors may influence the processing of negation.

There is a consistent finding in the literature on negation that participants are slower to evaluate the truth of a sentence when the sentence contains a negative element. In one such study, conducted by Clark & Chase (1972), adults participated in a sentence verification task in which participants were asked to compare either affirmative or negative sentences to simple pictures which either matched or did not match the sentences. Participants were overall slower to respond to negative sentences, such as "star isn't above plus". In addition, an interaction was found between the type of sentence and the truth of the sentence, such that participants were slower to respond to false affirmative sentences than true affirmative sentences, but were faster to respond to false negative sentences than true negative sentences. (Clark & Chase, 1972). The finding that negative sentences incur a greater response time than positive sentences, and the presence of an interaction between the type of sentence and the truth value is a robust finding that has been

replicated many times (Carpenter & Just, 1975; Just & Carpenter, 1971, 1976). The results of studies such as these have generally been taken as evidence for a propositional view of negation, in which a negative sentence is represented as a negative operator operating over a proposition, leading to greater processing time for negative as opposed to positive sentences.

There is a critical difference between evaluating sentences in a laboratory setting, where sentences are presented without any communicative context, and comprehending speech in the real world. According to Grice's cooperative principle (Grice, 1975), speakers should produce utterances that are appropriately informative, relevant, and unambiguous. However, in many cases producing a negative sentence can be uninformative, ambiguous, and irrelevant. If my friend asks me what I did this week, and I respond "I bought a car!", this piece of new information is informative and relevant to the conversation. However, if I respond "I didn't buy a car this week", my contribution to the conversation is neither relevant nor informative. The fact that unsupported negative utterances are generally rated as more ambiguous than unsupported affirmative utterances makes intuitive sense, and has been documented experimentally (Glenberg, Robertson, Jansen, & Johnson-Glenberg, 1999). In general, negative utterances are produced in contexts where there is some *expectation* that the speaker wishes to negate. For example, if my friend knows that I have been in the market for a new car, then she may be asking about my week because she expects I bought a new car, and my statement "I didn't buy a car" becomes relevant and informative.

Noting that denials are generally produced in response to a violation of expectations, Wason (1965) designed a study to examine whether pragmatic constraints such as context would affect how participants responded to negative sentences. Participants viewed stimuli consisting of 8 colored dots, in which 7 dots were one color and 1 dot was a different color. Participants were asked to describe the stimuli, and then evaluate positive and negative sentences about the stimuli. Wason found that when participants' descriptions of the stimuli highlighted the fact that one of the dots was an exception to the rule, they were faster to evaluate negative sentences about the "odd" dot. Additional work has shown that when participants read sentences that set up a supportive

context, reading times for negative sentences are reduced (Glenberg et al., 1999; Lüdtke & Kaup, 2006), and mouse trajectories to evaluate the truth of negative sentences are smoother and less discrete (Dale & Duran, 2011). Children are also sensitive to context effects when processing negative sentences, responding faster to respond to negative sentences about a toy that is different from other toys in the group (De Villiers & Tager Flusberg, 1975), and producing more spontaneous negatives to describe a stimulus that lacks an obvious characteristic (Watson, 1979). This work suggests that pragmatics and contextual support are particularly important to the ease at which negative sentences are processed.

(there is some more literature that I'm not sure if/how I should discuss here - kaup, ludtke? Tian, Breheny, Ferguson?)

(this paragraph is weak, I'm not really sure what I should be saying here and what work I should be citing): Studies of the effects of pragmatics on linguistic processing exist in other areas of language as well. For example, extensive research has been conducted on the processing of scalar implicatures, such as the inference made by listeners that the word "some" excludes the possibility "all", even thought semantically the word "some" can encompass "all". However, the relationship between pragmatic processing and reaction time is not entirely clear. Work by Huang and Snedeker (CITE) suggests that there is a gap between semantic and pragmatic processing, with participants taking significantly longer to resolve referential ambiguity when the resolution involves computing a pragmatic implicature (i.e. some = not all). In addition, work by Breheny (2006 CITE) demonstrates that the context in which a sentence is presented influences whether participants interpret "some" as lower-bounded (i.e. "some" can mean "all") or upper-bounded (i.e. "some" cannot mean "all"). Participants took longer to read sentences in contexts that promoted an upper-bounded interpretation of "some" and therefore required processing a scalar implicature. However, other work has suggested that listeners are able to rapidly incorporate contextual information into their interpretation of a sentence, e.g. making the inference that adjectives are used contrastively to more rapidly select a target item (CITE Sedivy 1999). (I'm not sure how to

end this paragraph...also, I should work in somewhere the work you and noah have done on pragmatics and how listeners assume speakers are informative and this can help them make predictions about reference etc...I'm not sure how or where to incorporate this though).

This paper examines the possibility of a pragmatic account of negation comprehension. Under this account, the processing delays incurred by evaluating negative sentences (Carpenter & Just, 1975; Just & Carpenter, 1971, 1976; Clark & Chase, 1972) are due to the lack of contextual support, which makes these sentences pragmatically awkward. Since the pragmatics of negation dictates that a negative sentence should occur in response to a violation of expectation, the presence of a visual context setting up expectations before each trial should facilitate the processing of negative sentences. Study 1 examines this possibility. Furthermore, if we vary the degree of expectation that each context sets up - for example, by varying the base rate of a certain "event" occurring - then we should see this variation reflected in reaction times to evaluate negative sentences. Studies 2 and 3 explore this by systematically varying the contexts seen before each trial. All three of the studies presented here use a sentence verification task paradigm, in which participants are asked to decide whether a sentence is true or false based on the accompanying picture. The critical manipulation in each experiment consisted of the presentation of a visual context, which was designed to set up or manipulate participants' expectations about the picture in each trial.

Study 1: Effect of context on negation processing

Study 1 replicates the finding that negative sentences take adults longer to evaluate than positive sentences when presented without any context (Clark & Chase, 1972; Carpenter & Just, 1975; Just & Carpenter, 1971, 1976), and further examines the effect of visual contexts on the processing of negative sentences (Wason, 1965). The study was conducted using Amazon's Mechanical Turk website, an online crowd-sourcing platform for conducting survey research, and was programmed using Javascript, allowing us to collect reaction times for participants' responses.

CITE OTHER PAPER ABOUT RT ON TURK. Participants viewed positive and negative sentences such as "Bob has/has no apples" along with pictures of people who were either holding the target item (e.g. "apples") or holding nothing, and were asked to evaluate whether the sentences were true or false. Half of the participants first viewed a "context screen" which showed e.g. three boys standing with apples. By replicating a known, robust finding from the literature on Mechanical Turk, this study allowed us to determine whether online platforms are appropriate for collecting sensitive measures such as reaction time (should I take this sentence out now that another study has been published - or should I keep it in, since there still hasn't been a lot of work done in this way?). In addition, previous work examining context effects on the processing of negative sentences required participants to actively engage with the context, either by describing the stimuli (Wason, 1965) or by reading sentences (Dale & Duran, 2011; Glenberg et al., 1999; Lüdtke & Kaup, 2006). This study explores whether a passively-viewed visual context is sufficient for facilitating the processing of negative sentences. (my feeling about this right now is that if we're submitting this to psych science, I need the motivation for this study to be catchier. But the truth is, this study in itself isn't very novel - it is the last two that are interesting. How should I introduce this without making this research sound "incremental")

Method

(note - I think we talked about changing this setup, so that I present the method for all three studies right up front, and then present the results for all three studies. I haven't done this here but I think it is probably a good idea, both in terms of space and also because I think it is more engaging)

Participants. 100 participants were recruited to participate in an online experiment through Amazon's Mechanical Turk website. Participants ranged in age from 18-65, with the majority of participants being between 18 and 25 (45 participants). 63 participants were male and 37 were female. Participation was restricted to individuals in the United States, and participants who indicated that their primary language was something other than English were excluded from

Sentence type

	POSITIVE	NEGATIVE
TRUE	Bob has apples	Bob has no apples
FALSE	Bob has apples	Bob has no apples

Truth value

Figure 1. Examples of the four trial types: True Positive, False Positive, True Negative, False Negative.

analysis. Participants were paid 30 cents to participate in the study, which took approximately 5 minutes to complete.

Stimuli. 28 trial items were created in which a character was shown either in possession of two of the same common, recognizable objects (such as apples, buckets, toy cars, cookies, flowers, kites, umbrellas, etc.), or standing alone. On each trial, beneath the picture, was a sentence of the form "[NAME] has/has no [ITEM]". Half of the sentences were positive and half were negative, and they were paired with pictures such that half were true and half were false. Lexical negation was used to make the negative sentences match the positive sentences as closely as possible. The decision to have characters hold two items, rather than one, was made because the negative "Bob
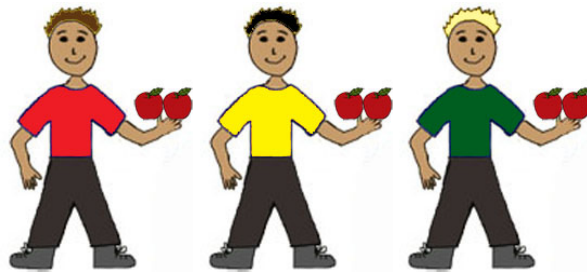
Look at these boys!

Figure 2. An example of the context that participants in the context condition viewed before each trial

has no apples" was judged by the authors as more felicitous than "Bob has no apple".

The experiment was fully crossed, with participants receiving 7 true positive, 7 false positive, 7 true negative and 7 false negative sentences over the course of the study. Trials were presented to participants in a randomized order.

Examples of the four trial types (true positive, false positive, true negative and false negative) can be seen in figure 1.

Participants were randomly assigned to either the "No Context" condition or the "Context" condition. Participants in the "No Context" condition saw a blank screen before each trial, while participants in the "Context" condition viewed a context slide before each trial. The context slide showed a picture of three characters, each holding two items. The characters in the context condition all differed from the trial character and from each other in hair colors and shirt colors. Beneath the three people was the sentence "Look at these [boys/girls]!" An example of a context slide can be seen in figure 2.

Procedure. Participants were first presented with an instructions screen which described the task, followed by 8 practice trials in which participants judged whether a sentence about the color of a character's clothing was true or false. During the practice trials, if participants answered a

question incorrectly a pop-up window informed them that they were incorrect and instructed them to try again. At the end of the practice, another instructions screen reminded them of the procedure and told them to press space to start the "game".

Before each trial, participants either viewed a blank screen or a context screen, depending on the condition they were assigned to. In both cases the experiment automatically advanced after 3 seconds. They then saw a picture and a sentence about that picture, and they were asked to select whether the sentence was true or false by pressing "Q" for false or "P" for true. Participants were told to answer and quickly and accurately as possible. Once participants made their selection by pressing "P" or "Q", they were automatically brought to the next trial after a 500 ms pause. Participants were instructed to use their left hand to select the "Q" key and their right hand to press the "P" key, to ensure that they would be able to respond as quickly as possible. Reaction times, measured as the time from when the pictures/sentence were presented to the moment when the participant made their response, were recorded by the computer

At the end of the experiment, participants were taken to a screen that collected demographic information. Participants were required to enter their gender, age, and native language, and then were invited to provide optional comments about the task.

Results

Study 1 examined the effect of a visual context on the processing of negative sentences. Previous research led us to expect a main effect of sentence type, with negative sentences incurring greater RTs than positive sentences, as well as a main effect of truth value, with false sentences incurring greater RTs than true sentences, and a sentence type*truth value interaction in the no context condition (Clark & Chase, 1972). We hypothesized that participants in the context condition, who first viewed a visual context setting up an expectation (e.g. that boys have apples), would show faster responses to negative sentences than participants in the no context condition.

	No Context		Context	
Trial Type	**Mean**	**SD**	**Mean**	**SD**
True Positive	1513	618	1481	692
False Positive	1713	634	1564	582
True Negative	1805	552	1555	645
False Negative	1724	561	1594	622

Table 1

Means and standard deviations of reaction times in milliseconds for each of the four trial types, for each of the two context conditions.
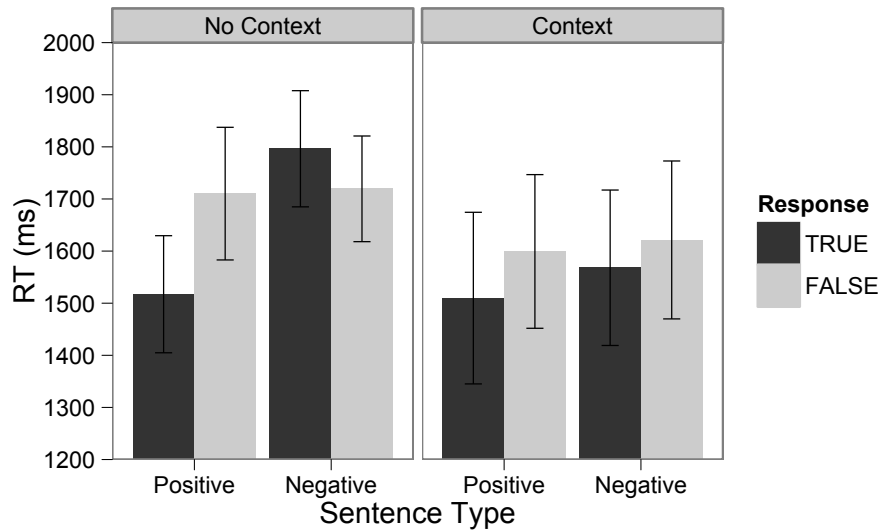
Figure 3. Reaction times for each trial type, across different conditions. Error bars represent the 95% confidence intervals.

Fixed Effects	Coef.	Std. Error	t value
Sentence type(negative)	9.566	47.620	0.201
Response(true)	-200.187	42.552	-4.704
Context condition(context)	-123.582	97.790	-1.264
Trial number	-7.838	1.300	-6.029
Sentence type(negative)*Context condition(context)	21.774	65.565	0.332
Sentence type(negative)*Response(true)	264.017	61.924	4.264
Context condition(context)*Response(true)	118.445	58.622	2.020
Sentence type(negative)*Context condition(context)*Response(true)	-231.469	85.424	-2.710

Table 2

Coefficient estimates from a linear effects model estimating the effects of context condition, sentence type, and response on reaction time, accounting for random effects of participant and item.

Six participants who listed a language other than English as their native language were excluded from analysis. Seven additional participants were excluded for having participated in a previous iteration of the experiment. Four participants were excluded for having an overall accuracy of below 80%. Thus, data from a total of 83 participants were analyzed.

Only correct trials were analyzed, due to the well-documented speed-accuracy tradeoff found in measures of reaction time. Trials with RTs greater than 3 standard deviations from the mean were also excluded from the analysis. Mean RTs and standard deviations for each of the four trial types are presented in table 1. A graph of these data can be seen in figure 3. Looking just at the data for participants in the "no context" condition, the effects seen here replicate previous work on the processing of negative sentences (Clark & Chase, 1972; Carpenter & Just, 1975; Just & Carpenter, 1971, 1976), with negative sentences slower than true sentences, false sentences slower

than true sentences, and an interaction between sentence type and response. However, comparing this to the data from participants in the context condition, there appears to have been a facilitating effect of context, such that participants who first viewed the context slides had lower RTs. This effect was strongest for true negative sentences.

A maximal linear mixed effects analysis was conducted to examine the effects of sentence type, response (truth value), and context on log(RT). Sentence type (positive or negative), response (true or false), and context condition (context or no context) and their interaction as were included as fixed effects, and the model accounted for random intercepts and slopes across participants and items. (note: is this the best way to report this? I noticed in Alex stiller's paper he reports two separate models, one for the experimental group and one for the control. If i split this analysis up into two models, one for no context and one for context, would that make my results easier to interpret?)

Discussion

Study 1 provides further evidence for the effects of context on the processing of negative sentences. Several previous studies have shown that embedding negative sentences in a supportive linguistic context facilitates the processing of negative sentences (Dale & Duran, 2011; Glenberg et al., 1999; Lüdtke & Kaup, 2006). In this study, we presented participants with a visual context that was designed to set up an expectation for what the trial picture might look like. For example, if participants viewed a context showing three boys holding apples, this sets up an expectation that boys have apples. We found that viewing these contexts before each trial lead to faster processing of negative sentences, with true negative sentences showing the greatest effect.

To understand why context had the greatest effect on true negative sentences, consider what a true negative trial looks like to a participant in the no context condition. These are trials in which the participant has no expectation about what the character might be holding, because no context was provided to set up such an expectation. The participant would then see a picture of a boy, not

holding anything, with the sentence "Bob has no apples". It seems obvious that a sentence such as this would cause participants to falter, since there is no reason for "apples" to be mentioned at all. However, when a participant first views a context such as the one pictured in figure ??, they can form an expectation based on this context that boys hold apples. Now, when the trial shows a boy with no apples, a sentence such as "Bob has no apples" makes sense and can be processed rapidly, because the image violates an expectation that boys should have apples.

Study 1 contributes to a body of evidence suggesting that negative sentences are more felicitous when they are used to negate an expectation that participants might hold, and that such expectations can be set up by an appropriate context. In study 2, we examine how participants responses are affected by changing the strength of the expectations set up by the context.

Study 2: 3-person contexts

While study 1 provides evidence that negative sentences - in particular, true negative sentences - are facilitated by the presence of a context, it does not explain how or why context has this effect. In order to further examine the effect of context on negation, we quantitatively manipulated the strength of the context, so that some participants saw contexts in which none of the characters were holding items (i.e. all three were empty-handed), some saw contexts in which 1/3 of the characters had items, some saw contexts in which 2/3 of the characters had items, and some saw contexts in which all of the characters had item (as in the "Context" condition in study 1). If one thinks of the contexts as giving participants a glimpse into the "world" that each trial exists in, the context gives the participant a small sample of the base rate of what the characters in this "world" look like. By manipulating the base rate of e.g. boys who have apples, we can change peoples' expectations about what the trial character will look like. If the differences in reaction times between the no context and the context condition in study 1 really are due to the difference in the expectations set up by the context, we should be able to further manipulate reaction times by systematically varying the strength of the expectation set up by the context.

Recent work by Frank et al. (Frank, Goodman, Lai, & Tenenbaum, 2009; Frank & Goodman, 2012) has attempted to quantify the pragmatic inferences that adults make when playing simple "language games". The assumption underlying the predictions made by these authors is that speakers are informative - that is, they will produce utterances that will pick out smaller subsets of the context, leaving as little ambiguity as possible for the listener. Regarding the present study, this should predict a linear effect of context on RTs to negative sentences, since "Bob has no apples" is not very informative at all when the context contains three boys with no apples - it could apply to any of the boys that the participant has seen. However, as the context changes and more boys in the context *have* apples, the informativeness of the statement "Bob has no apples" increases, because it applies to an increasingly smaller subset of the context. Study 2 tests this prediction by varying the contexts that different participants see before each trial. (Note: I wasn't sure what predictions I should make here. I think originally we did have a prediction that we would find a linear effect. Do I stick with this as our original hypothesis, or do I frame this study based on what we know now - that it is actually more like a u-shaped relationhip?)

Method

Participants. 200 participants were recruited to participate in an online experiment through Amazon's Mechanical Turk website. Participants ranged in age from 18-70, with the majority of participants being between 18 and 25 (84 participants). 129 participants were male and 71 were female. Participation was restricted to individuals in the United States, and participants who indicated that their primary language was something other than English were excluded from analysis. Participants were paid 30 cents to participate in the study, which took approximately 5 minutes to complete.

Stimuli. Study 2 used the same 28 trial items and sentence types as those used in Study . A between-subjects "context" factor determined what type of context participants saw. In the "none" context condition, participants saw three characters without any objects. In the "one" context

| Trial Type	None		One		Two		Three	
	Mean	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
True Positive	1519	578	1412	612	1456	543	1487	592
False Positive	1685	668	1526	567	1489	475	1610	566
True Negative	1697	593	1629	666	1505	498	1600	560
False Negative	1692	604	1543	602	1528	525	1737	672

Table 3

Means and standard deviations of reaction times in milliseconds for each of the four trial types, across each of the four contexts.

condition, participants saw two characters with nothing and once character holding two of the expected object. In the "two" context condition, participants saw one character with nothing and two characters holding the expected objects. In the "all" context condition, participants saw three characters holding the expected objects (identical to the "Context" condition in Study 2). Beneath the three characters was the sentence "Look at these [boys/girls]!" Examples of the None, One, and Two context conditions can be seen in figure 4. Trial stimuli were identical to the stimuli used in Studies 1 and 2.
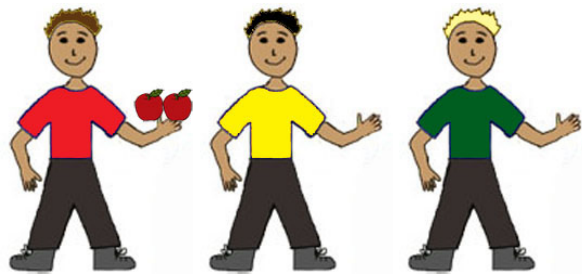
Stimuli.

Procedure. The procedure was identical to the procedure in Study 2, with participants randomly assigned to one of the four context conditions.
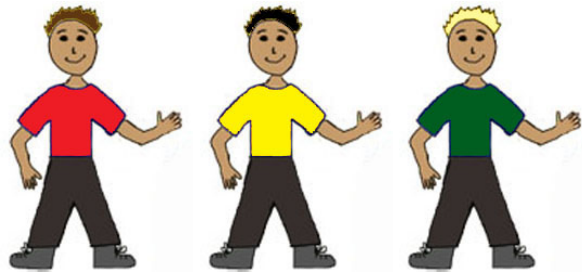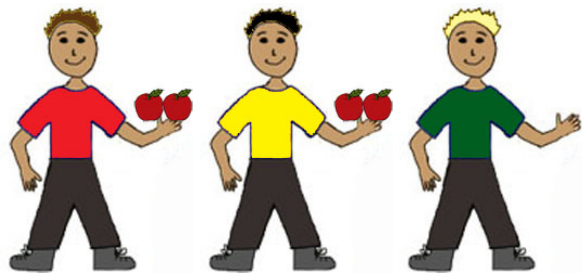
Look at these boys!

Look at these boys!

Look at these boys!

Figure 4. Examples of the none context (top), the one context (middle), and the two context (bottom) from study 2.
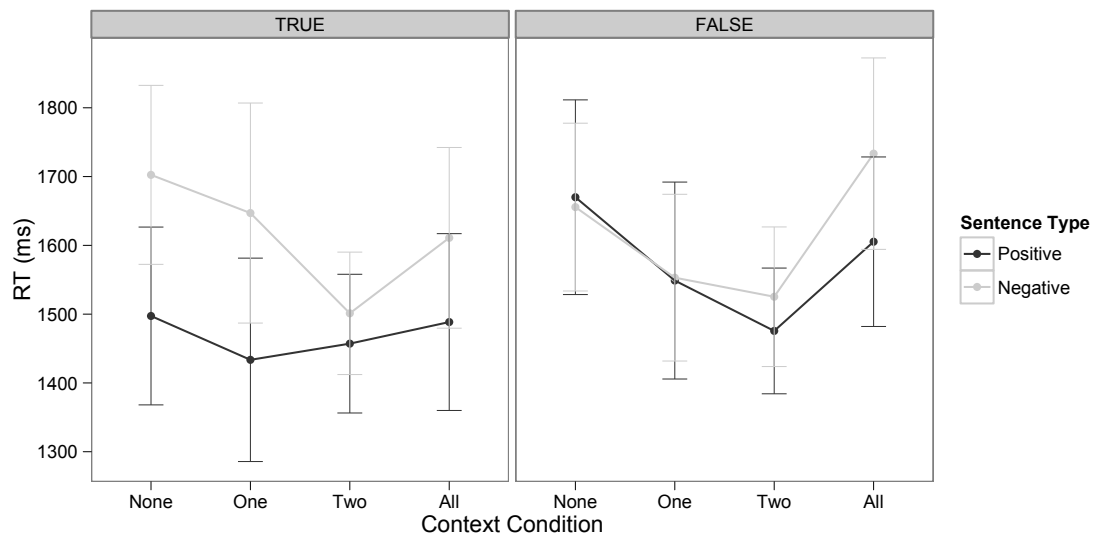
Figure 5. Reaction times for each trial type, across different conditions.

Results

Study 2 examined the effect of manipulating the context in ways that should change participants' expectations about what each trial should look like. We predicted that there would be an inverse linear effect of context, such that as the number of e.g. boys with apples in the context increased, reaction times to evaluate negative sentences would decrease.

Seven participants who listed a language other than English as their native language were excluded from analysis. Fifteen additional participants were excluded for having participated in a previous iteration of the experiment. Eleven participants were excluded for having an overall accuracy of below 80%. Thus, data from a total of 167 participants were analyzed.

Only correct trials were analyzed, due to the well-documented speed-accuracy tradeoff found in measures of reaction time. Trials with RTs greater than 3 standard deviations from the mean were also excluded from the analysis. Mean RTs and standard deviations for each of the four trial types are presented in table 3. A graph of RTs for each trial type across conditions is presented

Fixed Effects	Coef.	Std. Error	t value
Context(quadratic)	-3.8301	9.0699	-0.42
Sentence type(negative)	-6.1572	30.6565	-0.20
Response(true)	-120.6406	28.1636	-4.28
Trial number	-6.7562	0.8036	-8.419
Context(quadratic)*Sentence type(negative)	15.3235	5.3769	2.85
Context(quadratic)*Response(true)	6.1625	5.6339	1.09
Sentence type(negative)*Response(true)	178.5123	42.8571	4.17
Context(quadratic)*Sentence type(negative)*Response(true)	-26.3395	8.0837	-3.26

Table 4

Coefficient estimates from a linear effects model estimating the effects of context condition (quadratic), sentence type, and response on reaction time, accounting for random effects of participant and item.

in figure 5.

Looking at figure 5, there appears to be a quadratic effect of context on RT, particularly for true negative sentences, such that the "two" context has the greatest effect of increasing RTs for these sentences, and RTs for True Negative sentences appear to decrease slightly in the "all" context condition. With context condition coded as a continuous variable, we tested a maximal mixed effects model which examined the effect of sentence type, response, the quadratic effect of context, and the interaction between the three on reaction time. Effects of participants and items were included as random effects. In this model, the interaction between context as a quadratic variable and negative sentence types was positive (t=2.85), as was the three-way interaction between a quadratic effect of context, negative sentence types, and true sentences (t=-3.26) (note:

does this make sense?), suggesting that the pattern seen in figure 5 is statistically significant. Estimated coefficients for this model are presented in table 4.

Discussion

Study 2 was designed to test how the effect of context changes when the strength of the context is quantitatively manipulated. We hypothesized that the effect of context would increase linearly as the strength of the context increased (i.e. the number of boys with apples in the context). We found that our manipulation of the context actually produced a quadratic effect, such that a context in which participants saw 2 out of 3 boys holding apples produced the fastest reaction times, while the context in which all 3 boys were holding apples showed slightly increased reaction times.

The fact that we found a quadratic relationship between context and reaction times to respond to negative sentences was unexpected, but can also be explained in terms of the pragmatics of negationl. In our original hypothesis, we were considering the surprisal of the *sentence*, "Bob has no apples". However, if you also consider the surprisal of seeing a certain *picture*, the quadratic relationship seen in study 2 begins to make more sense. In the "None" context, seeing the sentence "Bob has no apples" is always surprising, because there is no reason to talk about a lack of apples - none of the boys have apples, so attributing this fact to Bob is not informative and it is not relevant to the context. In the "one" context, the sentence "Bob has no apples" becomes slightly more relevant, because apples have been introduced to the context, but since the majority of the boys don't have apples, it is still uninformative. However in the "two" condition, now the majority of the boys have apples, making Bob's lack of apples highly relevant and also informative. This is true in the "all" condition as well, but in the "all" condition you never see a boy without apples in the context - therefore, it seems that the slight increase in RTs in the "all" context could be due to a small amount of surprise at seeing a boy with nothing, since the expectation is that boys will have apples. This expectation makes negation in the "all" condition

felicitous, but could cause a slight increase in RT as the participants encode this unexpected picture. Thus, the "two" condition, which sets up an expectation that the majority of boys have apples while also allowing for other possibilities (i.e. boys with nothing), has the greatest effect on facilitating the processing of negative sentences.

Study 3: 4-person contexts

Study 3 was designed to further examine the possibility of a U-shaped relationship between the base rate set up by the context and reaction times in response to negative sentences, as was suggested by study 2. This was done by expanding the possible contexts, so that each context showed 4 boys/girls and there were 5 possible

Method

Participants. 400 participants were recruited to participate in an online experiment through Amazon's Mechanical Turk website. Participants ranged in age from 18 to over 65, with the majority of participants being between 18 and 25 years old (184 participants). 205 participants were male and 195 were female. Participation was restricted to individuals in the United States, and participants who indicated that their primary language was something other than English were excluded from analysis. Participants were paid 40 cents to participate in the study, which took approximately 7 minutes to complete.

Stimuli. In study 3, the number of items was increased to 48. As in study 2, a between-subjects "context" factor determined what type of context participants saw. The contexts were the same as those in study 2, except that each context contained 4 boys and there were therefore 5 context conditions (none, one, two, three, and four).

Procedure. The procedure was identical to the procedure in studies 1 and 2, with participants randomly assigned to one of the five context conditions.
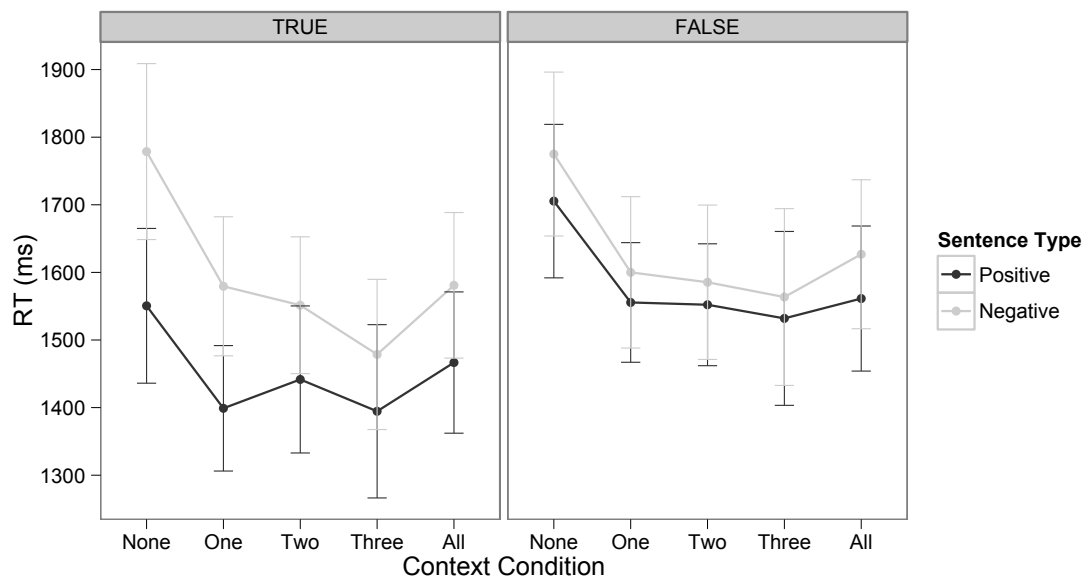
Figure 6. Reaction times for each trial type, across different conditions.

Results

Study 3 explored the U-shaped effect of context on reaction time seen in study 2 by expanding the contexts to include 4 characters and thus creating 5 different context conditions. We predicted a similar pattern as that seen in study 2, with reaction times decreasing as the number of characters with the target item in the context increased, but with the "all" context showing increased reaction times.

?? participants who listed a language other than English as their native language were excluded from analysis. ?? additional participants were excluded for having participated in a previous iteration of the experiment. ??participants were excluded for having an overall accuracy of below 80%. Thus, data from a total of ?? participants were analyzed (sorry I didn't have these numbers on-hand - I'll fill these in ASAP, I thought I'd written them down but I'll have to run the R script again to check them).

Fixed Effects	Coef.	Std. Error	t value
Context(quadratic)	-5.1314	3.9488	-1.30
Sentence type(negative)	51.5396	22.4416	2.30
Response(true)	-147.4821	18.7467	-7.87
Trial number	-3.4087	0.2772	-12.30
Context(quadratic)*Sentence type(negative)	0.1794	2.1696	0.08
Context(quadratic)*Response(true)	3.3698	2.1134	1.59
Sentence type(negative)*Response(true)	124.0178	28.8466	4.30
Context(quadratic)*Sentence type(negative)*Response(true)	-5.8065	2.9911	-1.94

Table 5

Coefficient estimates from a linear effects model estimating the effects of context condition (quadratic), sentence type, and response on reaction time, accounting for random effects of participant and item.

Only correct trials were analyzed, due to the well-documented speed-accuracy tradeoff found in measures of reaction time. Trials with RTs greater than 3 standard deviations from the mean were also excluded from the analysis. A graph of RTs for each trial type across conditions is presented in figure ??. (I don't have a table of means for this study - I can make one easily, but I'm undecided on how useful it is. I made the ones for studies 1 and 2 for my FYP, which is why I included them here, but do you think they are just a waste of space and I should only include the graphs?)

Figure 6 shows a similar U-shaped function as the one seen in study 2 - reaction times decrease as the number of boys/girls with the target item increases, but increase again in the "all condition". As before, a maximal mixed effects model was tested, examining the effect of sentence

type, response, the quadratic effect of context, and the interaction between the three on reaction time with items and participants as random effects. The three-way interaction between a quadratic effect of context, negative sentence types, and true sentences was (marginally??) significant (t=-1.94). Estimated coefficients for this model are presented in table 5.

A meta-analysis of the data was conducted by pooling together results of both studies, and plotting the results. In these figures (figure 7 and figure 8), the context condition was plotted on the x-axis as the proportion of characters in the context with the target item, such that 0 = "none" context condition and 1 = "all" context condition. Figure 7 shows the standardized difference scores, calculated as (true negative RTs - true positive RTs) / (true negative RTs + true positive RTs) for each subject, then aggregated and standardized within each experiment. Figure 8 shows the standardized RTs for true positive and true negative sentences in each study. This graph shows that the U-shaped relationship seen in figure 7 are driven primarily by changes in responses to true negative sentences, while responses to true positive sentences stay fairly consistent across contexts. (I know I need to say more here about this analysis, but I'm not sure what. I don't have any statistical analysis to report here - should I?)

General Discussion

This paper explores the relationship between the processing of negative sentences, and the contexts in which these sentences are presented. Previous work has shown that although it takes longer to evaluate negative sentences than positive sentences when test sentences are presented with no context (Carpenter & Just, 1975; Just & Carpenter, 1971, 1976; Clark & Chase, 1972), the presence of a verbal context (Dale & Duran, 2011; Glenberg et al., 1999; Lüdtke & Kaup, 2006) or the process of describing a visual context (Wason, 1965) can facilitate the processing of negative sentences. This suggests that a pragmatic account of negative sentences is warranted, such that negative sentences are licensed in contexts which set up an expectation which is then violated.
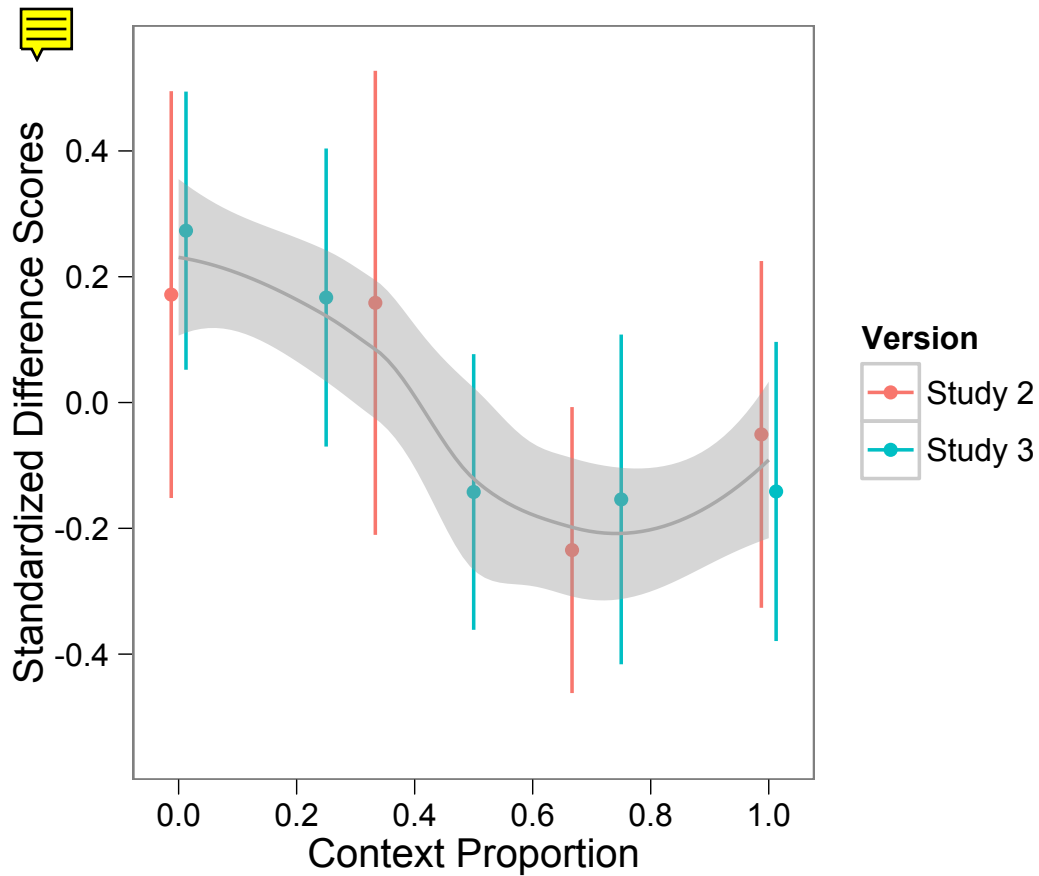
Figure 7. A graph of standardized difference scores (true negative - true positive / true negative + true positive) from studies 2 and 3

However, the extent to which expectation modulates the processing of negative sentences has not been systematically examined until now.

In the studies presented here, we quantitatively manipulated the expectation set up by the context by changing the proportion of people in the context who held a certain target item. We found a quadratic relationship between the proportion of people in the context with the target item and the reaction times to respond to negative sentences. This effect was most striking in response to true negative sentences, in which participants saw a person holding nothing and saw a sentence
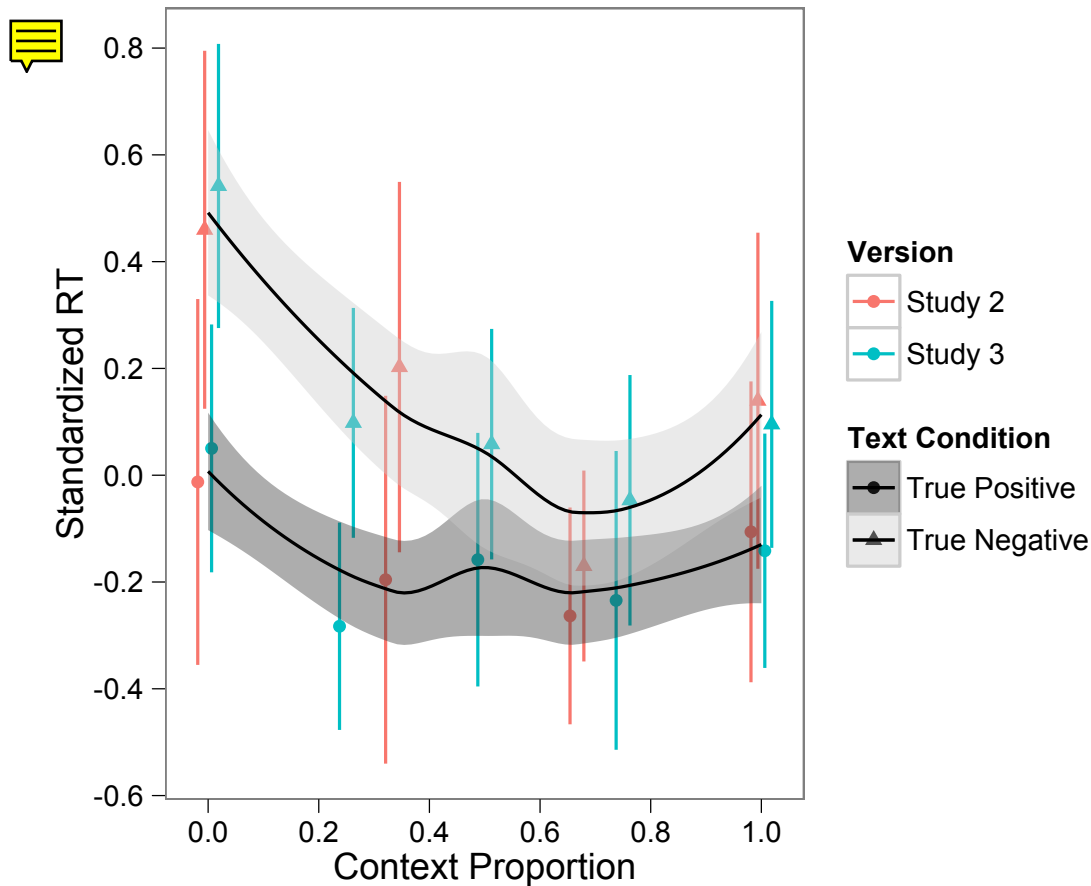
Figure 8. A graph of standardized reaction times for true positive and true negative sentences, from studies 2 and 3.

such as "Bob has no apples". True positive sentences showed little, if any, effect of context.

To understand the quadratic effect of context, we must consider how the context alters a participant's expectations about the picture that may appear as well as the likelihood of a negative term being used to describe that pictures. In our original predictions, we considered only the effects of expectation on the likelihood that a negative would be used to describe the picture; that is, the stronger the expectation that e.g. boys have apples, the more likely a negative will be used to describe a picture of a boy with no apples. However, a linear prediction does not account for the

surprisal incurred by seeing a boy with no apples when you've received a context that suggests that *all* boys have apples. If this is taken into account, the relationship between context and reaction time seen in studies 2 and 3 makes sense - reaction times decrease as the proportion of people with the target item increases, but reaction times to negative sentence increase in the "all" context due to surprise at seeing an unexpected boy or girl with nothing. (I'm just repeating myself now, and not saying anything interesting, so I'm going to stop and reconsider what I want to be saying here as I wrap up this paper...)

References

Bloom, L. (1970). Language development: Form and function in emerging grammars.

Bloom, L. (1993). *Language development from two to three*. Cambridge Univ Pr.

Carpenter, P., & Just, M. (1975). Sentence comprehension: A psycholinguistic processing model of verification. *Psychological Review*, *82*(1), 45.

Clark, H., & Chase, W. (1972). On the process of comparing sentences against pictures. *Cognitive psychology*, *3*(3), 472–517.

Dahl, Ö. (2010). Typology of negation. *The Expression of Negation, ed. Laurence R. Horn*, 9–38.

Dale, R., & Duran, N. (2011). The cognitive dynamics of negated sentence verification. *Cognitive science*.

De Villiers, J., & Tager Flusberg, H. (1975). Some facts one simply cannot deny. *Journal of Child Language*.

Frank, M., & Goodman, N. (2012). Predicting pragmatic reasoning in language games. *Science*, *336*, 998.

Frank, M., Goodman, N., Lai, P., & Tenenbaum, J. (2009). Informative communication in word production and word learning. In *Proceedings of the 31st annual conference of the cognitive science society* (pp. 206–211).

Glenberg, A., Robertson, D., Jansen, J., & Johnson-Glenberg, M. (1999). Not propositions. *Journal of Cognitive Systems Research*, *1*, 19-33.

Grice, H. (1975). Logic and conversation. *1975*, 41–58.

Just, M., & Carpenter, P. (1971). Comprehension of negation with quantification. *Journal of Verbal Learning and Verbal Behavior*, *10*, 244-253.

Just, M., & Carpenter, P. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, *8*, 441-480.

Lüdtke, J., & Kaup, B. (2006). Context effects when reading negative and affirmative sentences. In *Proceedings of the 28th annual conference of the cognitive science society* (pp. 1735–1740).

Pea, R. (1980). The development of negation in early child language. *The social foundations of language and thought*, 156–186.

Wason, P. (1965). The contexts of plausible denial. *Journal of Verbal Learning & Verbal Behavior*.

Watson, J. (1979). Referential description by children in negative form. *British Journal of Psychology*, *70*, 199-204.