Lying, Fast and Slow[*]

Angelo Turri [a]♣
John Turri [b]♣

[a] Independent scholar
angelo.turri@gmail.com

[b] Philosophy Department and Cognitive Science Program
University of Waterloo
200 University Avenue West
Waterloo, Ontario N2L3G1, Canada
john.turri@gmail.com

♣ Corresponding author

Abstract: Researchers have debated whether there is a relationship between a statement's truth-value and whether it counts as a lie. One view is that a statement being objectively false is essential to whether it counts as a lie; the opposing view is that a statement's objective truth-value is inessential to whether it counts as a lie. We report five behavioral experiments that use a novel range of behavioral measures to address this issue. In each case, we found evidence of a relationship. A statement's truth-value affects how quickly people judge whether it is a lie (Experiment 1). When people consider the matter carefully and are told that they might need to justify their answer, they are more likely to categorize a statement as a lie when it is false than when it is true (Experiment 2). When given options that inhibit perspective-taking, people tend to not categorize deceptively motivated statements as lies when they are true, even though they still categorize them as lies when they are false (Experiment 3). Categorizing a speaker as "lying" leads people to strongly infer that the speaker's statement is false (Experiment 4). People are more likely to spontaneously categorize a statement as a lie when it is false than when it is true (Experiment 5). We discuss four different interpretations of relevant findings to date. At present, the best supported interpretation might be that the ordinary lying concept is a prototype concept, with falsity being a centrally important element of the prototypical lie.

Keywords: lying; deception; social cognition; inference; prototypes

Word count: 6890

Introduction

Lying is an important social category and has long been studied by moral philosophers and social scientists. One question concerns the moral evaluation of lying, such as whether lying is ever justified (e.g. Augustine, 395; Lindskold & Waters, 1983; Lindskold & Han, 1986). Another, seemingly prior question concerns whether a lie has even occurred (for an extensive review, see Mahon, 2015). In other words, what counts as a lie? With respect to this question, a popular view among researchers is that a lie is a dishonest assertion (i.e. a statement believed to be false and intended to deceive). On this view, you lie if you say something that you think is false in order to deceive your audience, regardless of whether your statement is actually false. In other words, the objective *truth-value* of a deceptively motivated statement is irrelevant to whether it counts as a lie (e.g. Vrij, 2008; Fallis, 2009).

Contrary to that view, some researchers argue that a lie must be false. Three considerations help clarify the motivation for this. First, suppose we are told that a certain speaker lied to his audience about Smith's location and, moreover, that the audience believed the speaker's lie. It seems to follow without qualification that the audience has a false belief about Smith's location. But if lies could be true, then that would not follow without qualification. Second, if you are accused of lying, showing that your statement was true seems to definitively refute the accusation (Carson, 2006, p. 284). But if lies could be true, then the refutation would not be definitive. Third, since ancient times, people have noticed that statements such as "I am lying right now" are extremely puzzling, partly because although it is possibly true that one is lying right now, one

could not truthfully say that one was doing so. This is known as "the liar paradox" (for a review, see Beall & Glanzberg, 2011). But if lies could be true, then there would be no paradox: one could truthfully report that one was currently lying. All of this suggests that there is a conceptual connection between lying and falsity.

In this paper, we investigate whether patterns in ordinary social judgments support a conceptual connection between lying and falsity. Some work on ordinary lie attributions suggests that, contrary to what we suggested in the previous paragraph, there might be no connection. For instance, one study found that nearly 80% of participants judged a dishonestly motivated true statement to be a lie, whereas only 10% of participants judged an honestly motivated false statement to be a lie (Turri & Turri 2015, Experiment 1). Other work suggests that there is a weak connection. One study found that participants were slightly more likely to classify a false statement as a lie, but this was the least important of several factors affecting lie attributions (Coleman & Kay, 1981). Similar results have been observed in studies on Arabic and Spanish speakers (Cole, 1996; Hardin, 2010). Yet other work suggests that there might be a very strong connection. For example, in one study, a statement's objective truth-value was the most important factor affecting lie attributions (Strichartz & Burton, 1990).[1] In another study that used stimuli designed to disentangle perspective-taking and moral evaluation from lie attributions, a deceptively motivated false statement was judged a lie by 90% of participants, but a deceptively motivated true

_____

[1] We regret overlooking Strichartz and Burton's important study in our earlier literature review. We now regard their paper (Stricharz & Burton 1990) as the first study to explicitly consider and provide strong evidence for the hypothesis that objective truth-value is central to the ordinary lying concept.

statement was judged a lie by only 10% of participants (Turri & Turri, 2015, Experiment 3). For instance, consider a speaker who intends to deceive the authorities and says something he thinks is false but which, for reasons beyond his awareness, turns out to be true. Did he lie? If the only options are "yes" or "no," then some people who think that he did *not* lie might still answer "yes" as a way of acknowledging that *it seemed to him that he lied*, or that *he tried to lie*, or, relatedly, as a way of *blaming him for his deceptive intent*. In other words, the response options might force some participants to choose between performing a simple categorization task, on the one hand, and acknowledging the speaker's perspective or offering a moral assessment, on the other. By contrast, if the options were "he tried to lie and did lie" or "he tried to lie but didn't lie," then participants would not face such a choice, because each option acknowledges the speaker's perspective and calls attention to the morally problematic deceptive intent.

Contrary to all of those findings, other researchers, inspired by philosophical pragmatics, argue that some stimuli or procedures used in prior research pressured people into treating the statement's objective truth-value as relevant (Wiegmann, Samland & Waldmann, 2016). For instance, some argue that it is "unnatural" to suggest that lie attributions could be disentangled from perspective-taking, or that someone could try to lie but fail to do so. Perhaps people would never spontaneously distinguish between someone *trying to lie*, on the one hand, and *actually lying*, on the other. If so, then when presented with the options "he tried to lie and did lie" or "he tried to lie but didn't," participants might become confused or consult information that they ordinarily would not, such as the statement's objective truth-value. These researchers ran additional studies to support their hypothesis and concluded that objective falsity is not part of the ordinary

concept of lying (Wiegmann, Samland & Waldmann, 2016).

In light of the conflicting results reviewed above, further work is needed to fully understand the potential role of truth-value in the ordinary concept of lying and lie attributions. Equally importantly, it is worth seeking additional evidence to help evaluate several competing theoretical accounts of the available evidence. One possibility, arguably the most popular in the literature, is that the ordinary lying concept is a prototype concept (Coleman & Kay 1981; Cole, 1996; Hardin, 2010). On this account, objective falsity need not be, strictly speaking, required for lying, because prototypes can be characterized by central tendencies rather than necessary (or sufficient) conditions. Nevertheless, objective falsity is essential to the ordinary lying concept because the prototypical lie is false. In this respect, objective falsity could be to lying as flight is to birds: flight is essential to our concept of bird but it is not strictly necessary, as demonstrated by the fact that we count penguins and ostriches as birds. A second possibility is that the ordinary lying concept does require objective falsity, but this is sometimes obscured by factors interfering with people's performance on lie attributions, such as perspective-taking, moral evaluation, or other forms of task substitution (Turri & Turri 2015). On this account, objective falsity is essential to the ordinary lying concept because it is a requirement. A third possibility is that objective falsity is inessential to the ordinary lying concept, but this is sometimes obscured by factors interfering with people's performance, such as task demand or pragmatic reinterpretation (Wiegmann, Samland & Waldmann, 2016). A fourth possibility is that there are multiple ordinary lying concepts, that falsity is essential to some but not others, and that perhaps none of the existing findings are due to interference with people's performance (Turri & Turri, 2015; Peterson, Peter-

5

son & Seeto, 1983; Machery, 2009, pp. 34-5).

Accordingly, we conducted five new experiments to advance understanding of these important issues by using a wider variety of tasks and dependent measures than featured in previous research. We tested for a relationship between lying and truth-value in five different ways. In Experiment 1 we tested whether truth-value affects how quickly participants judge whether a statement is a lie. In Experiment 2 we tested whether truth-value affects whether participants attribute a lie when they consider the matter slowly and are asked to justify their answer. Some participants' free-form justifications suggested that perspective-taking was artificially inflating the rate of lie attribution. So in Experiment 3 we tested whether participants categorize deceptively motivated true statements as lies when they are given options that inhibit perspective-taking. In Experiment 4 we tested whether describing a speaker as "lying" leads participants to infer that the speaker's statement is false. Finally, in Experiment 5 we tested whether truth-value affects how frequently participants spontaneously attribute lies. In the General Discussion, in light of the findings, we evaluate the four broad theoretical possibilities identified above.

Experiment 1

This experiment investigates whether a deceptively motivated statement's truth-value affects how quickly people judge whether it is a lie. If a statement's truth-value is conceptually related to whether it counts as a lie, then we would expect reaction-times to be faster when deciding whether a false statement is a lie.

Method

Participants

Fifty-three participants were recruited and tested online using Amazon Mechanical Turk and Qualtrics and compensated $4.00 for approximately 18 minutes of their time. Fifty-one participants completed the experiment (aged 18-61, mean age = 34 years, SD = 10.8; 14 female; 90% reporting English as a native language). The same basic recruitment procedures were used in all experiments reported here. Repeat participation was prevented (by AMT Worker ID) within and across experiments.

Materials and Procedure

Participants first completed a lexical decision task to familiarize them with giving speeded responses (See the supplemental file for further information on the lexical decision task and analyses.) Participants then performed a narrative evaluation task. This task began with 6 warmup trials (order randomized), followed by 16 experimental and 15 filler trials (order randomized). A trial started with participants reading a brief text about a situation. After participants read and understood the text, they pressed the spacebar to proceed. The next screen contained a partial sentence (horizontally centered) missing its final word, which remained on the screen for 3000 ms. Next a fixation cross (horizontally centered) appeared for 1500 ms, followed immediately by a word (horizontally centered) that completed the sentence. The task was to decide ("as quickly as you can") whether the completed sentence accurately described the situation. Participants

pressed the f-key to answer "no" and the j-key to answer "yes." A trial timed-out after 5 seconds. The next trial started immediately after the participant responded or the trial timed-out.

The experimental trials took the form of a 2 (Truth Value: false, true) × 2 (Quality: negative, positive) within-subjects design. The Truth Value factor manipulated whether the agent in the story said something false or true. The test sentence that participants evaluated on experimental trials always pertained to whether the agent lied. The Quality factor manipulated whether the sentence was negative (the agent "didn't lie") or positive (the agent "did lie"). If the test statement was negative, answering "no" was scored as a lie attribution, whereas if it was positive, answering "yes" was scored as a lie attribution.

Here is an example illustrating all four treatments for a single scenario (coffee break):

(Text) Samantha's co-worker Allen recently started taking an extra coffee break in the morning. One morning the boss visits Samantha's unit and asks, "Where is Allen?" Samantha thinks that Allen is taking an extra coffee break, so she says, "He's meeting a client." In fact, Allen is [out on a coffee break/meeting a client]. So what Samantha said was [false/true]. (Test sentence) Samantha [lied/didn't lie].

Participants completed all four treatments for four separate scenarios (= 16 experimental trials). A supplemental file includes all stimuli used in this experiment. Here is an example of a filler trial:

A pedestrian accidentally drops her phone on the sidewalk in front of a bakery.

Later that same day, some children run by, step on the phone, and crack it.

(Probe) The phone was broken by the [children/pedestrian/baker]

Results

We removed data from four participants who seem to have either answered randomly or reversed the response keys. These participants scored 0% correct on high-frequency word trials whereas, overall, average accuracy on high-frequency trials was 88%; these participants also attributed a lie on 0% of false positive trials whereas, overall, a lie was attributed on 89% of such trials. We did not remove any other responses. Thus the analyses reported below included 47 participants.

For each participant, we calculated (i) the average (mean) rate of lie attribution in the four cells of the experimental design and (ii) the average (mean) reaction-time (RT) in the four cells. We then analyzed average rates of lie attribution and RT from the experimental trials with 2×2 repeated measures analyses of variance. Each analysis used Truth Value (false, true) and Quality (negative, positive) as within-subjects factors.
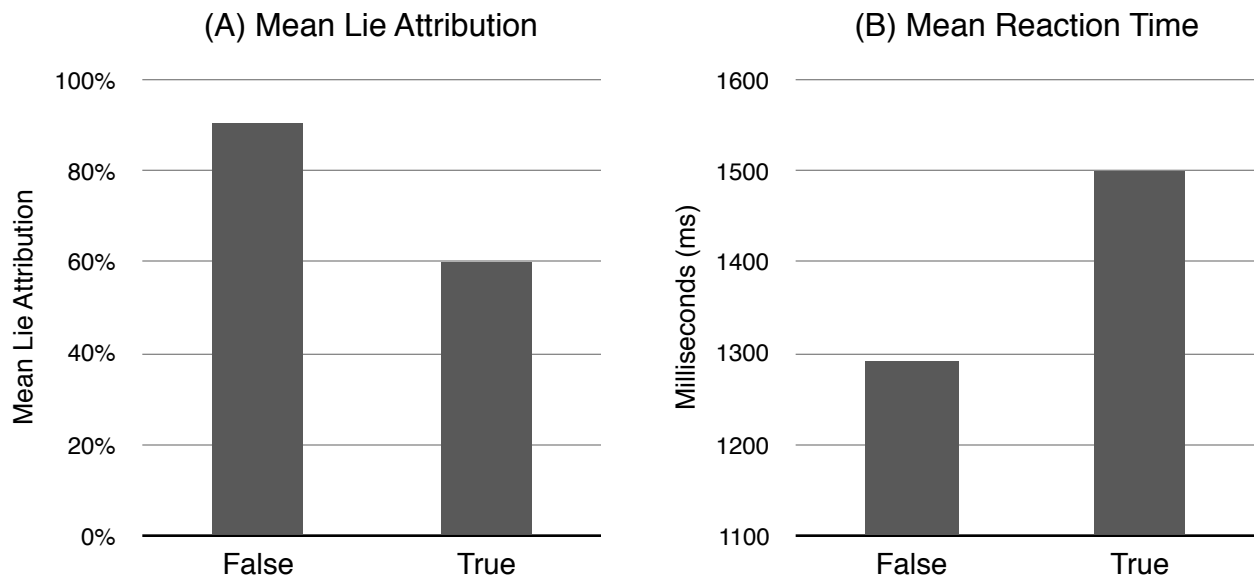

Figure 1. Experiment 1. Panel A: mean rate of lie attribution in the two truth-value conditions (false, true). Panel B: mean reaction time for lie attributions in the two conditions.

Beginning with average lie attribution, there was a very large main effect of Truth Value, $F(1, 46) = 33.26$, $p < .001$, $\eta_p^2 = .420$, no effect of Quality, $F(1, 46) = 0.44$, $p = .510$, n.s., $\eta_p^2 = .009$, and no interaction, $F(1, 46) = 0.02$, $p = .895$, n.s., $\eta_p^2 < .001$. (See Figure 1.) Lie attribution was significantly higher when the agent made a false statement (90%) than when the agent made a true statement (59%). Moving on to average RTs, there was a main effect of Truth Value, $F(1, 46) = 9.59$, $p = .003$, $\eta_p^2 = .172$, a main effect of Quality, $F(1, 46) = 6.27$, $p = .016$, $\eta_p^2 = .120$, and no interaction, $F(1, 46) = 1.93$, $p = .172$, n.s., $\eta_p^2 = .040$. Average RT was significantly slower when the claim was negative (1493 ms) than when it was positive (1296 ms). Similarly, average RT was significantly slower when the agent made a true statement (1498 ms) than when the agent made a false statement (1292 ms).

Discussion

The results from this experiment provide evidence that a statement's truth value is conceptually related to whether it counts as a lie. This conclusion is supported in two ways. First, it is supported by the rate at which people attributed lies. People were more likely to attribute a lie when the agent made a false statement than when the agent made a true statement. This difference was not only statistically significant but also large. Second, it is supported by the time it took people to judge whether an agent lied. People were significantly faster to make a judgment when the agent made a false statement than when the agent made a true statement.

Two issues might be raised about the finding that people are significantly less likely to attribute a lie when the agent's statement is true. On the one hand, perhaps the fact that people

10

were under time-pressure caused them to make more mistakes. On the other hand, perhaps the fact that we explicitly described the statement as "true" or "false" pressured people into basing their answer on that information, which they would not do otherwise. The next experiment addresses both of these issues.

Experiment 2

In Experiment 1, we encouraged people to evaluate lie attributions as fast as they could. In this experiment, we encouraged people to take much longer and reflect carefully before judging whether someone lied. Participants also explained their answers. We refrained from explicitly describing an agent's statement as "true" or "false" in the scenario. If a statement's truth-value is conceptually related to whether it counts as a lie, then we would expect people to be more likely to attribute a lie when the agent makes a false statement rather than a true one.

Method

Participants

One hundred twenty-five participants (aged 18-72, mean age = 34 years, SD = 11.1; 39 female; 94% reporting English as a native language) were tested and compensated $0.75 for approximately 3 minutes of their time.

Materials and Procedure

Participants were randomly assigned to read one of four scenarios in a 2 (Scenario: break, pool) × 2 (Truth Value: false, true) between-subjects design, after which they judged whether an agent lied and explained their answer. We used the break and pool scenarios from Experiment 1; the only difference was that we eliminated the final sentence indicating that the agent's statement was "true" or "false." Here is the text for the pool scenario:

> (Pool) Karen's neighbor Anthony just got back from vacation. Anthony asks Karen, "Did your children swim in my pool while I was gone?" Karen thinks that her children did swim in his pool, so she says, "No, they did not swim in your pool." In fact, they [did/did not] swim in the pool.

Right below the scenario were these instructions:

> Instructions: On the next page of the survey, you will be asked a question about the scenario described above. When the question is revealed to you, please take at least thirty seconds to carefully formulate your answer. (You can take longer than thirty seconds, if you wish.) After thirty seconds, you'll be able to advance and submit your answer. After you submit your answer, you might be asked to justify the answer you gave.

Participants then advanced to a new screen. The scenario remained atop the screen and below it was a question. Here is the question for the pool scenario:

> Question: In the scenario described above, did Karen lie?

Participants were unable to advance past this screen for at least thirty seconds; the advance but-

ton appeared only after thirty seconds, at which point participants could click it when they chose to. The next screen included the story atop and a lie attribution below it:

Did Karen lie? [Yes/No]

Response options to the lie attribution were rotated randomly. Participants then advanced to a new screen that included a single question,

Why did you choose that answer?

and a text box beneath.

Results

Analysis of variance revealed that the time participants spent reflecting on the scenario and question ("reflection time") was unaffected by Scenario, $F(1, 121) = 2.55$, $p = .113$, n.s., Truth Value, $F(1, 121) = 0.37$, $p = .546$, n.s., or their interaction, $F(1, 121) < .01$, $p = .933$, n.s. Mean reflection time was 37 seconds (SD = 9.85) and reflection time ranged from 31 to 95 seconds.

We conducted a binary logistic regression to determine whether lie attributions were predicted by Scenario, Truth Value, or their interaction. (See Table 1.) Only Truth Value significantly predicted lie attributions: 58% of participants attributed a lie in true conditions compared to 94% in false conditions. (See Figure 2.) Changing the agent's statement from true to false increased the odds of a lie attribution by a factor of 13.13. The rate of lie attribution did not differ from what was observed in Experiment 1 for either the false conditions, binomial test, $p = .464$ (test value = .90), or the true conditions, binomial test, $p = .978$ (test value = .59).
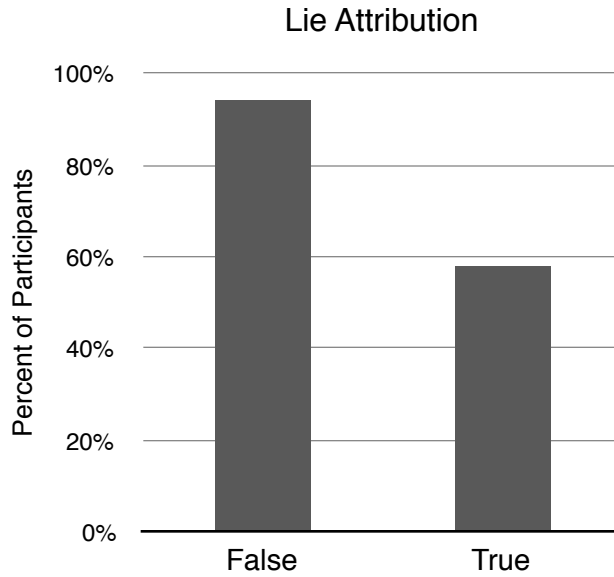
13

Figure 2. Experiment 2. The percent of participants attributing a lie in the two truth-value condi‐
tions (false, true). Results collapse across scenario.

Table 1. Experiment 2. Logistic regression predicting lie attribution.

| | B | SE | Wald | df | p | Odds Ratio | Odds Ratio 95% CI | |
							LLCI	ULCI
Scenario	0.38	0.52	0.53	1	.466	1.46	0.53	4.02
Truth Value	2.58	0.82	9.93	1	.002	13.13	2.65	65.08
Scenario × Truth Value	-0.41	1.16	0.13	1	.722	0.66	0.07	6.38
Constant	0.13	0.37	0.13	1	.715	1.14		

Note: $\chi^2(3, n = 125) = 24.18$, $p < .001$. Reference class for Scenario: pool. Reference class for Truth Val‐
ue: true.

Discussion

The results from this experiment provide further evidence that a statement's truth value is con‐

ceptually related to whether it counts as a lie. On a task that encouraged careful consideration of

the matter, people were more likely to attribute a lie when the agent made a false statement than

14

when the agent made a true statement. This difference was not only statistically significant but also large: changing an agent's deceptively motivated statement from true to false increased the odds of a lie attribution by a factor of 13.

We did not formally code and analyze participants' explanations for their responses, but since they help motivate our approach to subsequent experiments, we will briefly mention a few examples and relate them to feedback received by other researchers. Some of our participants indicated that lying requires saying something false. For instance, one participant denied that a deceptively motivated true statement was a lie because, "technically, what she [i.e. the speaker] said was true." Others distinguished *thinking that one is lying* from *actually lying*. For instance, one participant wrote, "She believed she was lying, but she was in fact telling the truth." Other participants indicated that their lie attributions were based on perspective-taking, based on attributing deceptive intent, or not a characterization of the agent's literal statement to her interlocutor. For instance, some participants called a true statement a "lie" because the speaker "intended to lie," "lied to herself," or "lied about knowing." These examples cohere with feedback received by previous researchers during debriefing. For instance, one of Strichartz and Burton's participants remarked, "I could have answered some of those stories the other way. I wasn't sure whether to answer in terms of what really happened, or what [the speaker] believed" (Strichartz & Burton, 1990, p. 217). The next experiment builds on participants' explanations to investigate whether the rate of lie attribution for true statements is inflated due to perspective-taking.

Experiment 3

In Experiment 2, when asked to explain their judgments, many participants cited an appearance/reality distinction for lying or the speaker's perspective on the situation. Following our participants' lead, in this experiment we tested the effect of response options that respect the appearance/reality distinction and disentangle the agent's perspective from what actually occurs. More specifically, we compared the rate of lie attribution when the choice was between answering that the agent "lied" or "did not lie" (*plain* options), or between answering that the agent "thinks she lied and actually did lie" or "thinks she lied but actually did not lie" (*contrast* options). If the rate of lie attribution in earlier experiments was inflated due to perspective-taking, then we should expect the contrast options to lower the rate of lie attribution.

Method

Participants

One hundred sixty-nine participants (aged 21-68, mean age = 35 years, SD = 11.2; 64 female; 95% reporting English as a native language) were tested and compensated $0.50 for approximately 2 minutes of their time.

Materials and Procedure

Participants were randomly assigned to read one of four scenarios in a 2 (Option: plain, contrast) × 2 (Truth Value: false, true) between-subjects design, after which they judged whether the agent

lied and explained their answer. We used the pool scenarios from Experiment 2. The Option fac-

tor manipulated whether the answer options were "Karen lied" and "Karen didn't lie" (plain) or

"Karen thinks she lied and did actually lie" and "Karen thinks she lied but did not actually

lie" (contrast). The contrast options allow for an appearance/reality distinction. In particular, they

allow for a contrast between the agent's perspective and reality. The Truth Value factor manipu-

lated whether Karen's statement was false or true. After reading the story, participants were

asked, "Which option better describes the case?" and presented with one of the response-option

pairs described above (order rotated randomly). Participants then advanced to a new screen and

were asked to explain their answer in the same way as in Experiment 2.

Results

Binary logistic regression revealed that lie attribution was significantly predicted by Option, by

Truth Value, but not by their interaction. (See Table 2 and Figure 3.) Changing the response op-

tions from contrast to plain increased the odds of a lie attribution by a factor of 5.64. Changing

the agent's statement from true to false increased the odds of a lie attribution by a factor of 56.36.

Follow-up pairwise comparisons using Fisher's exact test revealed that the rate of lie attribution

did not differ between the plain false (95%) and contrast false (95%) conditions, $p = 1$, n.s.; it

was significantly higher in the plain true condition (67%) than in the contrast true condition

(26%), $p < .001$, Cramer's $V = .406$; it was significantly higher in the plain false condition than

in the plain true condition, $p < .001$, Cramer's $V = .367$; and it was significantly higher in the

contrast false condition than in the contrast true condition, $p < .001$, Cramer's $V = .707$. The rate

17

of lie attribution significantly exceeded chance rates (= 50%) in the plain false condition, $\chi^2(1, n = 43) = 35.37$, $p < .001$, in the contrast false condition, $\chi^2(1, n = 42) = 34.38$, $p < .001$, and in the plain true condition, $\chi^2(1, n = 42) = 4.67$, $p = .031$, but it was significantly below chance rates in the contrast true condition, $\chi^2(1, n = 42) = 9.52$, $p = .002$.
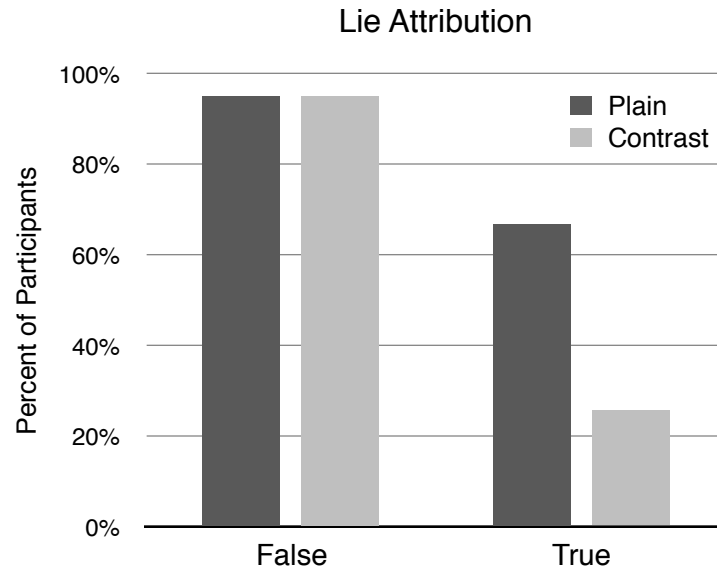

Figure 3. Experiment 3. Rate of lie attribution in the four conditions.

Table 2. Experiment 3. Logistic regression predicting lie attribution.

	B	SE	Wald	df	p	Odds Ratio	Odds Ratio 95% CI	
							LLCI	ULCI
Option	1.73	0.48	12.98	1	<.001	5.64	2.20	14.44
Truth Value	4.03	0.81	25.08	1	<.001	56.36	11.63	273.08
Option × Truth Value	-1.71	1.13	2.27	1	.132	0.18	0.02	1.67
Constant	-1.04	0.35	8.72	1	.003	0.36		

Note: $\chi^2(3, n = 169) = 69.48$, $p < .001$. Reference class for Option: contrast. Reference class for Truth Value: true.

Discussion

The results from this experiment provide evidence that the rate of lie attribution for true statements was inflated due to perspective-taking in earlier experiments, which in turn provides further evidence that a statement's truth-value is conceptually related to whether it counts as a lie. The rate of lie attribution was much lower when participants chose between response options that allowed them to distinguish between an agent's perspective and what really happened. Importantly, this happened only when the agent's deceptively motivated statement was true. When the agent's deceptively motivated statement was false, the possibility of distinguishing appearance from reality did not lower lie attributions, which remained at ceiling.

In this experiment, the options distinguishing appearance from reality were formulated based on participants' own open-ended explanations of their reasoning in earlier experiments. Based on those explanations, we asked participants to choose between answering that an agent "thinks she lied and actually did lie" or "thinks she lied but actually did not lie." But this is not the only way to contrast appearance and reality. Researchers have previously tested lie attribu-

tions when the options were that the agent "tried to tell a lie but failed to tell a lie" or "tried to tell a lie and succeeded in telling a lie"; researchers have also tested lie attributions when the options were "he tried to lie and actually did lie" or "he tried to lie but only thinks he lied" (Turri & Turri, 2015, Experiments 2-3). In each case, the results were substantially similar to the present results: when evaluating a deceptively motivated true statement and given the option to distinguish appearance from reality, the vast majority of participants did not attribute a lie.

Experiment 4

In Experiments 1 and 2, participants rated lie attributions, either very quickly or slowly, based on information given in the scenario. In this experiment, we told participants that an agent was "replying," "not lying," or "lying" as part of the scenario, and then we investigated what they would infer about the statement's truth-value. To accomplish this, we asked participants to rate their agreement with the content of the agent's statement. Participants also rated lie attributions. If a statement's truth-value is conceptually related to whether it counts as a lie, then we would expect two results. First, participants who were told that the agent lied will infer that the statement's content is false. Second, participants' own lie attributions will be strongly negatively correlated with their evaluations of the statement's truth value.

Method

Participants

One hundred twenty-three participants (aged 20-65, mean age = 35 years, SD = 11; 56 female; 93% reporting English as a native language) were tested and compensated $0.50 for approximately 2 minutes of their time.

Materials and Procedure

Participants were randomly assigned to one of three conditions (reply, no lie, lie) in a between-subjects design. Each participant read a brief scenario and responded to two test statements. The basic scenario was the same as the "coffee break" scenario used in earlier studies. The conditions differed in how the agent's speech act was described: she was described as "replying," "not lying," or "lying" in response to a question. Here is the text for the scenarios:

> Samantha's co-worker Allen recently started taking an extra coffee break on some mornings, down in the cafe in the lobby. On other mornings, Allen meets with clients in the board room. This morning the boss visits Samantha's unit and asks, "Where is Allen?" Samantha is [replying/not lying/lying] to the boss when she says, "Allen is in the board room meeting a client."

After reading the scenario, participants rated their agreement with two test statements:

1. Allen is in the board room. (truth evaluation)

2. Samantha was lying about Allen's location. (lie attribution)

Responses were collected on a standard 7-point Likert scale, 1 (strongly disagree) – 7 (strongly agree), left-to-right on the participant's screen. The truth evaluation appeared immediately under the story. Participants then went to a new screen and responded to the lie attribution; the story did not appear on this screen and participants could not return to the previous screen.

Results

Analyses of variance revealed a very large effect of condition (reply, no lie, lie) on response to the truth evaluation, $F(2, 120) = 77.30$, $p < .001$, $\eta p^2 = .563$, and the lie attribution, $F(2, 120) = 64.28$, $p < .001$, $\eta p^2 = .517$. (See Figure 4 and Table 3.) Response to both dependent measures in the reply and no-lie conditions was substantially similar, whereas responses in the lie condition differed from the other two conditions on both measures. Overall, responses to the truth evaluation and lie attribution were very strongly negatively correlated, $r = -.875$, $n = 123$, $p < .001$, indicating that to the extent that participants judged that the agent lied, they inferred that the agent's statement was false.
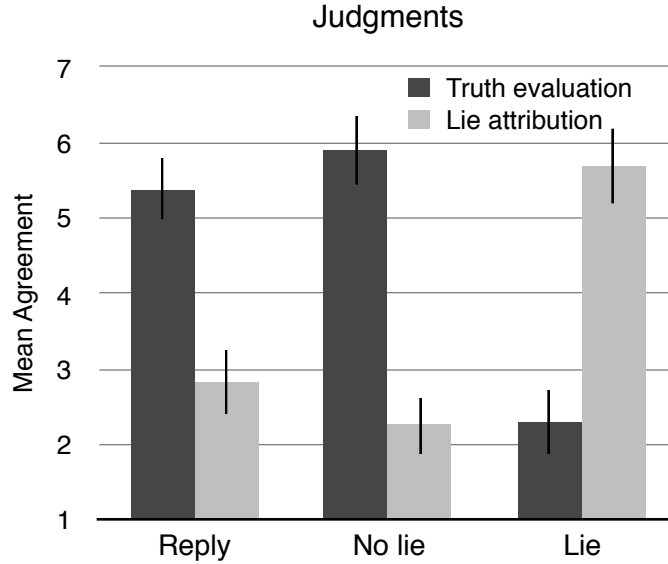
Figure 4. Experiment 4. Mean response to the truth evaluation and lie attribution in the three conditions. Scales ran 1 (strongly disagree) – 7 (strongly agree). Error bars represent boot-strapped 95% confidence intervals.

Table 3. Experiment 4. Descriptive statistics for response to the two test statements in the three conditions.

Measure	Reply (N = 41)				No lie (N = 40)				Lie (N = 42)			
	Mo	Md	M	SD	Mo	Md	M	SD	Mo	Md	M	SD
truth-value	6	6	5.39	1.34	7	6	5.90	1.41	1	2	2.29	1.54
lied	2	3	2.83	1.43	2	2	2.25	1.24	7	6	5.69	1.72

Discussion

Describing an agent as "lying" led people to strongly conclude that the statement's content was false, and people's own judgments about whether the speaker lied were strongly negatively cor-related with their own assessment of whether the statement was true. These findings provide fur-ther evidence that a statement's truth value is conceptually related to whether it counts as a lie.

23

Experiment 5

In Experiments 1-4, we asked participants to rate lie attributions or manipulated whether we told them that a statement was a lie. It is a theoretical possibility that, somehow, the results were due to our bringing up "lying," one way or another. In this experiment, we asked participants to perform a very different task: if they had to describe an agent's deceptively motivated statement in just one word, what word would it be? We manipulated whether the agent's statement was true or false. People's performance on this spontaneous description task cannot be due our bringing up "lying." If a statement's truth-value is conceptually related to whether it counts as a lie, then we would expect people to spontaneously describe the statement as a lie more frequently when it is false.

Method

Participants

One hundred three participants (aged 19-69, mean age = 34 years, SD = 10.3; 50 female; 98% reporting English as a native language) were tested and compensated $0.50 for 1-2 minutes of their time.

Materials and Procedure

Participants were randomly assigned to one of two conditions (false, true) in a between-subjects design. Each participant read a brief scenario and completed a simple descriptive task. The basic

scenario was the same as the "coffee break" scenario used in Experiments 1 and 2. The conditions differed in whether the agent in the scenario said something false or true:

Samantha's co-worker Allen recently started taking an extra coffee break in the morning. One morning the boss visits Samantha's unit and asks, "Where is Allen?" Samantha thinks that Allen is taking an extra coffee break, so she says, "He's meeting a client." In fact, Allen is [out on a coffee break/meeting a client]. So what Samantha said was [false/true].

Participants were then instructed,

If you had to describe Samantha's statement in just one word, what word would you use? Please enter it in the text box below. (Remember: just one word.)

Responses were entered in a text box immediately below the instructions.

Results

We coded responses as a lie attribution if and only if they used "lie" or a cognate, including "lied," "lying," and "liar," or an obvious misspelling of such a word, including "lier" and "lyer." Binary logistic regression revealed a significant association between assignment to condition and lie attributions: 20% of participants attributed a lie in the true condition compared to 58% in the false condition. (See Table 1.) Changing the agent's statement from true to false increased the odds of a lie attribution by a factor of 5.59.

Table 4. Experiment 5. Logistic regression predicting lie attribution.

	B	SE	Wald	df	p	Odds Ratio	Odds Ratio 95% CI	
							LLCI	ULCI
Condition	1.72	0.45	14.58	1	<.001	5.59	2.31	13.53
Constant	-1.41	0.35	16.01		<.001	0.24		

Note: $\chi^2(1, n = 103) = 16.28$, $p < .001$. Reference class for Condition: true.

Discussion

The results from a spontaneous description task provide further evidence that a statement's truth-value is conceptually related to whether it counts as a lie. Changing an agent's deceptively motivated statement from true to false significantly increased the odds (by a factor of 5.59) that participants would spontaneously describe the statement as a lie.

General Discussion

The results from five experiments advance knowledge of lie attributions in ordinary social cognition. We found that a statement's truth-value affects how quickly people judge whether it is a lie (Experiment 1), people are more likely to categorize false statements as lies (Experiments 1–3), people spontaneously cite truth-value and perspective-taking when justifying their answers about lying (Experiments 2-3), people tend to not categorize deceptively motivated true statements as lies (Experiment 3), people infer that a liar's statement is false (Experiment 4), and people are more likely to spontaneously categorize false statements as lies (Experiment 5).

Does the evidence to date support the hypothesis that there is a conceptual connection be-

tween lying and saying something false? We think that it clearly does, especially if the ordinary concept of lying is a prototype concept, as several researchers have previously argued (Coleman & Kay, 1981; Cole, 1996; Hardin, 2010; Strichartz & Burton, 1990). On a prototype account, the concept of lying summarizes the central tendency of lies (Wittgenstein 1953; Rosch & Mervis 1975). The central tendency consists of a cluster of properties that a prototypical lie has, none of which is strictly necessary for a lie.

Consider the hypothesis that falsity is essential to the ordinary concept of lying because the prototypical lie is false. Prototype concepts have several features that support this hypothesis. First, people are more likely to categorize typical instances as a category member (Hampton, 1995). This would explain the findings from several studies where people are more likely to categorize false statements as lies. Second, people categorize typical category members faster than atypical ones (Rosch, 1975). This would explain our finding that people are faster to judge whether false statements are lies. Third, people tend to learn typical category members prior to learning atypical ones (Mervis & Pani, 1980; Mervis, Catlin & Rosch, 1975). This would explain the developmental finding that from ages 4 to 7, children base their lie attributions almost exclusively on a statement's truth-value, and that the influence of truth-value continues through adulthood (Strichartz & Burton, 1990). When combined with the finding that children tend to assume that others know what they themselves know (Birch & Bloom, 2003; Birch & Bloom 2007), this supports the hypothesis that the prototypical lie is asserting something you know is false (see Turri, 2016, ch. 2; Benton, in press). Fourth, typicality affects how strongly people draw conclusions from information (Rips, 1975; Osherson, Smith, Wilkie, López & Shafir, 1990). This could

explain our finding that when informed that a statement is a lie, people strongly infer that it is false (see also Turri & Turri, 2016, Experiment 4). It could also explain our finding that people are more likely to spontaneously describe false statements as lies, given the plausible assumption that their "best" description expressed the inference they felt most confident in. Finally, because prototypical features are not necessary for category membership, a prototype theory of lying is consistent with the finding that people sometimes classify true statements as lies.

A bolder interpretation of the evidence to date is that lying requires saying something false, but this is sometimes obscured by factors interfering with people's lie attributions. For example, in one recent study, the vast majority of participants judged a dishonestly motivated *true* statement to be a lie (Turri & Turri 2015, Experiment 1). Surely this suggests that the — or, at least, *an* — ordinary concept of lying does not require falsity. Nevertheless, researchers observed that this might have been due to confounds interfering with people's responses. For example, when forced to agree or disagree with the claim that a dishonest speaker "lied," people might agree in order to acknowledge the speaker's perspective or avoid suggesting that they condone his deceptive intent. (On the related phenomena of "blame validation" and "excuse validation, see Alicke, 1992 and Turri & Blouw, 2015.) In a follow-up study, researchers tried to avoid these potential confounds by asking people to select between the two options, "He tried to lie and actually did lie" and "He tried to lie but only thinks he lied," which match one another in attributing deceptive intent but differ in whether a lie occurred. Very few participants attributed a lie in this context, thus completely reversing the pattern of attribution. Moreover, in the present studies, when participants were given an opportunity to explain their answers, some participants' explanations

suggested that perspective-taking and other factors could have artificially inflated the rate of lie attribution for deceptively motivated true statements. Following our participants' lead, we tested different response options that inhibited perspective-taking and found that this greatly reduced the rate of lie attribution and, indeed, led the vast majority of participants to deny that the agent lied (Experiment 3). Perhaps perspective-taking, blame validation, excuse validation or other mechanisms can fully explain why people sometimes count true statements as lies.

The inverse of the previous interpretation is also bold: a statement's truth value is irrelevant to whether it is a lie, but this is sometimes obscured by factors interfering with people's lie attributions. For example, researchers have argued that people base their lie attributions on truth-value only because of task demand connected to the way lie attributions were collected (Wiegmann et al., 2016). According to this line of reasoning, participants in some earlier studies were pressured into "unnaturally" attending to truth-value because they were asked about the statement's truth-value in the same context that they were asked whether the agent lied, or because of "complex" "two-part" probes. The "complex" probes involved asking participants to choose between "He tried to lie but only thinks he lied" and "He tried to lie and actually did lie." However, the present findings rule out this concern because some participants spontaneously cited truth-value or drew an appearance/reality distinction to justify their judgments about lying, similar to the feedback received by other researchers (Strichartz & Burton, 1990). Moreover, it is very unlikely that task demand can explain all the ways that truth-value affects lie attributions. For example, we see no straightforward way for such factors to explain why truth-value affects reaction-times for lie attributions, why people strongly infer that the content of a lie is false, why a statement's

truth-value strongly affects whether people spontaneously describe it as a lie, or why truth-value figures centrally in the early development of the concept of lying.

Another possible interpretation of existing findings involves positing multiple ordinary lying concepts. On this approach, one lying concept requires making a false statement, whereas another one does not. For instance, suppose that one concept requires objective falsity whereas another requires only a deceptively motivated statement. Depending on the situation, people could flexibly choose to apply one or the other concept when judging whether someone "lied." This pluralist hypothesis receives support from the fact that in some studies switching a statement from false to true pushes lie attributions from ceiling to floor, whereas in other studies the vast majority of participants judge a deceptively motivated true statement to be a lie. The hypothesis that participants choose to competently apply one concept or the other is arguably simpler or more charitable than the hypothesis that their judgments are seriously distorted in one situation or the other. A potential concern about positing multiple ordinary concepts is that it is difficult to strongly differentiate this possibility from the prototype theory, which would allow people to relax one or another of the prototypical features while still competently applying the concept. Of course, this concern does not call into question the theoretical coherence distinguishing multiple concepts from a flexible prototype, but rather highlights the challenge of designing studies that empirically discriminate between the two accounts.

An important limitation of the present research is that it was limited to anglophone residents of the United States. And while some findings on lie attributions have been replicated cross-culturally (Coleman & Kay, 1981; Cole, 1996; Hardin, 2010), other research shows cross-cultural

differences in how people judge lies (e.g. Fu, Lee, Cameron & Xu, 2001; Fu, Xu, Cameron, Heyman & Lee, 2007). Accordingly, the conclusions we have drawn here should be understood to pertain specifically to the ordinary lying concept, or concepts, of American anglophones. Generalizing beyond that to lying concepts in other cultures or languages — or, more ambitiously, to features of a potential species-typical lying concept — will require further research. Similarly, our results do not speak to potential effects due to cultural variability, such as those associated with geographical region or socioeconomic status, among American anglophones.

References

Alicke, M. D. (1992). Culpable causation. Journal of Personality and Social Psychology, 63(3), 368–378.

Augustine. (395AD). On lying. In P. Schaff & K. Knight (Eds.), Nicene and post-Nicene Fathers, First Series (Vol. 3). Christian Literature Publishing. Retrieved from http://www.newadvent.org/fathers/1312.htm

Beall, J., & Glanzberg, M. (2011). The liar paradox. Stanford Encyclopedia of Philosophy. Re-

trieved August 2015, from http://plato.stanford.edu/entries/liar-paradox/

Benton, M. A. (in press). Lying, belief, and knowledge. In J. Meibauer (Ed.), The Oxford hand-
book of lying. Oxford: Oxford University Press.

Birch, S. A. J., & Bloom, P. (2003). An asymmetric bias in mental-state attribution. Psychologi-
cal Science, 14(3), 283–286.

Birch, S. A. J., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs.
Psychological Science, 18(5), 382–386. http://doi.org/10.1111/j.1467-9280.2007.01909.x

Carson, T. L. (2006). The definition of lying. Noûs, 40(2), 284–306.

Cole, S. (1996). Semantic prototypes and the pragmatics of Lie across cultures. The LACUS Fo-
rum 23, 475–483.

Coleman, L., & Kay, P. (1981). Prototype semantics: the English word lie. Language, 57(1), 26–
44.

Fallis, D. (2009). What is lying? Journal of Philosophy, 106(1), 29–56.

Fu, G., Lee, K., Cameron, C. A., & Xu, F. (2001). Chinese and Canadian adults' categorization
and evaluation of lie-and truth-telling about prosocial and antisocial behaviors. Journal of
Cross-Cultural ….

Fu, G., Xu, F., Cameron, C. A., Heyman, G., & Lee, K. (2007). Cross-cultural differences in
children's choices, categorizations, and evaluations of truths and lies. Developmental
Psychology, 43(2), 278–293. http://doi.org/10.1037/0012-1649.43.2.278

Hampton, J. A. (1995). Testing the prototype theory of concepts. Journal of Memory and Lan-
guage, 34(5), 686–708.

Hardin, K. J. (2010). The Spanish notion of lie: revisiting Coleman and Kay. Journal of Pragmatics, 42(12), 3199–3213. http://doi.org/10.1016/j.pragma.2010.07.006

Lindskold, S., & Han, G. (1986). Intent and the Judgment of Lies. The Journal of Social Psychology, 126(1), 129–130. http://doi.org/10.1080/00224545.1986.9713581

Machery, E. (2009). Doing without concepts. Oxford: Oxford University Press.

Mahon, J. E. (2015). The definition of lying and deception. Stanford Encyclopedia of Philosophy, https://plato.stanford.edu/entries/lying-definition/.

Mervis, C. B., & Pani, J. R. (1980). Acquisition of basic object categories. Cognitive Psychology, 12(4), 496–522.

Mervis, C. B., Catlin, J., & Rosch, E. (1975). Development of the structure of color categories. Developmental Psychology.

Osherson, D. N., Smith, E. E., Wilkie, O., & Lopez, A. (1990). Category-based induction. Psychological Review, 97(2), 185–200.

Peterson, C. C., Peterson, J. L., & Seeto, D. (1983). Developmental changes in ideas about lying. Child Development, 54(6), 1529–1535.

Rips, L. J. (1975). Inductive judgments about natural categories. Journal of Verbal Learning and Verbal Behavior.

Rosch, E. (1975). Cognitive representations of semantic categories. Journal of Experimental Psychology: General, 104(3), 192. http://doi.org/10.1037/0096-3445.104.3.192

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. Cognitive Psychology, 7(4), 573–605.

Strichartz, A. F., & Burton, R. V. (1990). Lies and truth: A study of the development of the concept. Child Development, 61(1), 211–220.

Turri, J., & Blouw, P. (2015). Excuse validation: a study in rule-breaking. Philosophical Studies, 172(3), 615–634. http://doi.org/10.1007/s11098-014-0322-z

Turri, A., & Turri, J. (2015). The truth about lying. Cognition, 138(C), 161–168.

Turri, A., & Turri, J. (2016). Lying, uptake, assertion, and intent. International Review of Pragmatics, 8(2), 314–333. http://doi.org/10.1163/18773109-00802006

Turri, J. (2016). Knowledge and the norm of assertion: an essay in philosophical science. Open Book Publishers.

Vrij, A. (2008). Detecting lies and deceit: pitfalls and opportunities (2nd ed.). Hoboken, NJ: John Wiley & Sons.

Wiegmann, A., Samland, J., & Waldmann, M. R. (2016). Lying despite telling the truth. Cognition, 150(C), 37–42. http://doi.org/10.1016/j.cognition.2016.01.017

Wittgenstein, L. (1953). Philosophical investigations. (G. E. M. Anscombe, Trans.). Oxford: Basil Blackwell.

Lying Fast and Slow
Supplemental File

Lexical Decision Task

The lexical decision task involved deciding whether a string of letters flashed on the screen was a meaningful word in the English language. Participants judged 9 low frequency words, 9 high frequency words, and 9 pronounceable non-words selected from lists provided by the English Lexicon Project Web Site (Loftis, 2014). This task familiarized participants with giving speeded responses and served to validate the current procedures for gathering reaction-time data. In particular, it allowed us to verify that the current procedures reproduced prior findings whereby people respond faster and more accurately to high frequency words than to low frequency words. Prior research has already shown that Amazon Mechanical Turk and Qualtrics software can be used to successfully replicate reaction-time findings first discovered in lab settings (Zwaan & Pecher, 2012; Barnhoorn, Haasnoot, Bocanegra & van Steenbergen, 2014).

The advertisement for recruitment explicitly told recruits twice that only those using a computer with a keyboard should volunteer to participate, because keystrokes would be used to record responses. Because of the way the survey was coded in Qualtrics, it was not possible to record answers using a touch screen.

Participants were instructed that it was important for them to "answer fast and accurately." A trial started with a horizontally centered fixation cross for 1500 ms, followed immediately by a new screen with a word or nonword. Participants pressed the f-key to answer "no" and the j-key

to answer "yes." The next trial started 500 ms after the response. The order of all words and non-words was randomized.

We performed paired-sample t-tests on accuracy rates and reaction times, with word frequency (high, low) as the within subjects factor. (See Table 1.) The results replicated findings from previous research: participants were more accurate and faster for high frequency words. (See Zwaan & Pecher, 2012 for references and similar validation studies administered online using Amazon Mechanical Turk and Qualtrics.)

Table 1. Lexical decision task. Paired-samples t-tests on mean accuracy rates and reaction times for the lexical decision task, comparing high frequency and low frequency words.

	High frequency		Low frequency							
Measure	M	SD	M	SD	t	df	p	MD	95% CI	d
accuracy (%)	88.42	25.80	71.57	18.55	5.54	46	<.001	16.85	10.72, 22.96	0.85
reaction time (ms)	728	180	941	294	-6.91	46	<.001	-213	-275, -15	1.16

Instructions to Participants

You'll be performing a task that requires you to answer correctly as quickly as you can. In other words, for present purposes, it is important that you answer *fast and accurately*. ¶2 We'll start with a word recognition task. You'll first see a "fixation" mark appear on the screen, followed by a screen with a string of letters. Your job is to accurately decide, *as quickly as you can*, whether the string of letters forms a meaningful word in the English language. In short, is it a real word?

---

2 Indicates a paragraph on the participant's screen.

¶ Press the **f-key** to answer "no". Press the **j-key** to answer "yes".

Word lists

Table 2. Words used for the lexical decision task.

High frequency	Low frequency	Non-words
from	tot	aal
wrote	ilk	yop
other	yap	dut
from	hubs	nove
people	avid	thit
which	moat	glap
all	cubic	eroch
the	comma	plask
would	onset	inler

Narrative Evaluation Task

Instructions to Participants

Now you'll perform a slightly different task. As with the previous task, it is important that you answer fast and accurately. ¶ You'll first read a brief text about a situation. After you've read and understood the text, press the space-bar to proceed. ¶ After you press the space bar, you'll see the beginning of a sentence. It is not yet a complete sentence because the last word is missing. You'll then see a "fixation" mark appear on the screen, followed by a screen with the last word of the

sentence. ¶ Your job is to accurately decide, *as quickly as you can*, whether the sentence accurately describes the situation you just read about. ¶ Press the **f-key** to answer "no". Press the **j-key** to answer "yes".

Stimuli

Warmup Trials

(Text) A park ranger is patrolling the park's north side. The ranger leaves some debris near a camper's tent. This violates the rules set by the park manager. (Probe) The ranger left [debris/tents/rulebooks].

(Text) A politician is campaigning in a critical area. The politician meets with local officials who organized the event. Leaders from industry are also there. (Probe) The event was organized by [local officials/the politician/industry leaders].

Filler Trials

(Text) A bank posts its operating hours. The corporate headquarters decides the hours. The manager post the hours weekly on instructions from headquarters. (Probe) The bank's hours are decided by the [headquarters/manager/government]

(Text) An executive ordered a transaction. Several employees are carrying out the transaction based on the company's standard operating procedure. (Probe) The transaction was ordered by the [executive/employees/union].

(Text) A freelancer is designing a new car in his spare time after work. Based on the design,

an entrepreneur later builds the car and sells it to the public. (Probe) The car was designed by a [freelancer/entrepreneur/retiree].

(Text) A biologist makes a map of a previously unknown region. He sends the map to a publisher, who prints the map and puts it up for sale. (Probe) The map was made by a [biologist/ publisher/chemist].

(Text) A pedestrian accidentally drops her phone on the sidewalk in front of a bakery. Later that same day, some children run by, step on the phone, and crack it. (Probe) The phone was broken by the [children/pedestrian/baker].

Experimental Trials

(Text) Jacob's friend Mary recently posted classified government documents on the internet. Soon some federal agents visit Jacob and ask, "Where is Mary?" Jacob thinks that Mary is at the grocery store, so he tells the agents, "She is at her brother's house." In fact, Mary is at [the grocery store/her brother's house]. So what Jacob said was [false/true]. (Probe) Jacob lied [didn't lie].

(Text) Samantha's co-worker Allen recently started taking an extra coffee break in the morning. One morning the boss visits Samantha's unit and asks, "Where is Allen?" Samantha thinks that Allen is taking an extra coffee break, so she says, "He's meeting a client." In fact, Allen is [out on a coffee break/meeting a client]. So what Samantha said was [false/true]. (Probe) Samantha lied [didn't lie].

(Text) Karen's neighbor Anthony just got back from vacation. Anthony asks Karen, "Did

your children swim in my pool while I was gone?" Karen thinks that her children did swim in his pool, so she says, "No, they did not swim in your pool." In fact, they [did/did not] swim in the pool. So what Karen said was [false/true]. (Probe) Karen lied [didn't lie].

(Text) Eric and Lisa are co-workers in the sales department. Lisa was supposed to follow up with the client they met last week. Today the boss asks Eric, "Did Lisa remember to follow up with the client?" Eric thinks that she forgot to follow up, so he says, "Yes, she remembered." In fact, she [forgot/remembered]. So what Eric said was [false/true]. (Probe) Eric lied [didn't lie].

References (for Supplemental File)

Barnhoorn, J. S., Haasnoot, E., Bocanegra, B. R., & van Steenbergen, H. (2014). QRTEngine: An easy solution for running online reaction time experiments using Qualtrics. Behavior Research Methods. http://doi.org/10.3758/s13428-014-0530-7

Loftis, B. 2014. English Lexicon Project Website. <http://elexicon.wustl.edu>, accessed August 2015.

Zwaan, R. A., & Pecher, D. (2012). Revisiting mental simulation in language comprehension: six replication attempts. PLOS One, 7(12), 1–10.