

How “Ought” Exceeds but Implies “Can”: Description and Encouragement in Moral Judgment*

John Turri

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

Abstract: This paper tests a theory about the relationship between two important topics in moral philosophy and psychology. One topic is the function of normative language, specifically claims that one “ought” to do something. Do these claims function to describe moral responsibilities, encourage specific behavior, or both? The other topic is the relationship between saying that one “ought” to do something and one’s ability to do it. In what respect, if any, does what one “ought” to do exceed what one “can” do? The theory tested here has two parts: (1) “ought” claims function to both describe responsibilities and encourage people to fulfill them (*the dual-function hypothesis*); (2) the two functions relate differently to ability, because the encouragement function is limited by the person’s ability, but the descriptive function is not (*the interaction hypothesis*). If this theory is correct, then in one respect “ought implies can” is false because people have responsibilities that exceed their abilities. But in another respect “ought implies can” is legitimate because it is not worthwhile to encourage people to do things that exceed their ability. Results from two behavioral experiments support the theory that “ought” exceeds but implies “can.” Results from a third experiment provide further evidence regarding an “ought” claim’s primary function and how contextual features can affect the interpretation of its functions.

Keywords: responsibility; ability; blame; moral psychology; speech acts

Word count: 6535

* This is the penultimate version of a paper forthcoming in *Cognition*. Please cite the final, published version if possible.

Introduction

Normative claims about what people “ought” to (or, more typically, “should”) do are common and important in human life. It is no surprise, then, that researchers in a variety of disciplines, including philosophy, have studied them extensively. This paper deals with a topic at the intersection of two central theoretical questions regarding “ought” claims.

The first question concerns the relationship between what people *ought* to do and what they are *able* to do. One view is that if someone ought to do something, then he is able to do it (e.g. Kant 1793; Moore 1922). This view is so common in philosophy that it even has a slogan: “ought implies can.” The contrary view is that a person’s ability does not always constrain what he ought to do. This view does not have a slogan but one possibility is “ought exceeds can.” The second question concerns the linguistic function of “ought” claims. One theory is that “ought” claims (purport to) describe normative facts about an agent’s responsibilities or obligations. A natural corollary of this view is that (sincere) “ought” claims express beliefs or knowledge, which represent those normative facts (e.g. Mackie 1977; Boyd 1988; Huemer 2005). Another theory is that “ought” claims function as encouragement toward certain behavior. A natural corollary of this view is that “ought” claims express an emotion, command, preference, or other mental state that is not truth-evaluable, but which would be satisfied by the agent acting in the indicated way (e.g. Stevenson 1937; Carnap 1937; Gibbard 1990). A third, hybrid theory is that “ought” claims function to both describe responsibilities and encourage their fulfillment (e.g. Stevenson 1944; Hare 1952; Barker 2000).

In contemporary discussions, these two questions are related in an interesting way. More specifically, the hybrid theory of the function of “ought” claims might help explain how “ought implies can” is both wrong and right. (As far as I am aware, Sinnott-Armstrong 1984 contains the first authoritative statement of this sort of proposal; see also Pigden 1990.) Briefly, “ought implies can” is false because *having a responsibility* does not entail being able to fulfill it, but it seems correct because *encouraging someone to fulfill a responsibility* suggests a commitment to their being able to fulfill it. We can partially unpack that brief statement as follows:

On theoretical and empirical grounds, it has been persuasively argued that “ought implies can” is false when “implies” is understood as *conceptual entailment*. On theoretical grounds, it is false because it is inconsistent with the logic of excuses and the felicity of some apologies, threatens to trivialize many moral obligations, has sustained fruitless debates, and falsely entails that disabled people lack full moral agency (for a review, see Buckwalter ms). On empirical grounds, in a wide range of experimental studies, the vast majority of competent adult speakers attribute moral responsibilities in tandem with inability to fulfill them (Buckwalter & Turri 2014; Buckwalter & Turri 2015; Mizrahi 2015; Turri 2015; Chituc, Henne, Sinnott-Armstrong & De Brigard 2016; Henne, Chituc, De Brigard & Sinnott-Armstrong 2016; Turri 2016a; Turri 2016b; Turri ms). Some of the clearest and most widely studied cases involve promises (Sinnott-Armstrong 1984). Promising generates a responsibility that is not preempted or canceled by inability to fulfill it. If you promise to deliver a package by 4pm, then you have a responsibility to deliver the package by 4pm. If it turns out that you are unable to deliver the package on time, you still have a responsibility to do so — you ought to do something that you cannot do. Note that this

pertains to “ought” in the descriptive sense of having a responsibility. Consider, by contrast, its function to encourage behavior. It would typically be pointless to encourage you to, say, deliver the package by 4pm if it is acknowledged that you cannot do so — as many commentators have noted, it defies expectations, seems linguistically odd, or is otherwise infelicitous (see Pigden 1990 for a rewarding and informed review of sources dating back over 500 years).

One hypothesis, then, is that “ought” statements can be interpreted as both descriptions of responsibilities and as encouragement toward their fulfillment. Their descriptive function is compatible with inability, but their encouragement function is not. When I describe you as having a responsibility to fulfill your promise, my remark is perfectly consistent with your being unable to do so. In this respect, “ought” exceeds “can.” But when I encourage you to fulfill your promise, the natural interpretation of my remark requires attributing to me the belief that you can do so. In this respect, “ought” implies “can.”

As noted above, this theory is not original with me. And while many contributors have recognized that empirical evidence is clearly relevant to understanding the function of linguistic expressions (e.g. Pigden 1990; Huemer 2005: 18), to date the empirical evidence has consisted principally of introspective and social observations. Introspective and social observation is a good place to start when studying familiar social activities, including the use of normative language. But we should not draw firm conclusions from reported patterns in a single researcher’s experience. The researcher’s experience might be idiosyncratic, or the researcher’s theoretical commitments might cause him or her to selectively keep track of, or misinterpret, observations. Similar remarks apply to a small group of researchers’ experience. One way to avoid these prob-

lems is to test observational claims against the judgments of a large sample of competent, theoretically-uncommitted language users experienced in the use of normative language. In this spirit, I conducted two experimental studies to begin rigorously testing the theory, described above, that “ought” exceeds but implies “can.”

More specifically, I conducted two experiments that begin testing two principal hypotheses implicated in the theory. The *dual-function hypothesis* states that “ought” claims have (at least) two functions: description and encouragement. The *interaction hypothesis* states that those two functions relate differently to an agent’s ability or inability to perform the relevant action. In the first experiment, participants read a story in which an agent promises to do something, and a speaker tells the agent that he “ought” to do what he promised. I manipulated whether the agent was able or unable to fulfill the promise. Participants recorded judgments about whether the speaker was describing an obligation that the agent had, whether the speaker was encouraging the agent to fulfill the obligation, and whether the agent was able to fulfill it. Three key questions here are whether people actually do interpret the speaker as both describing and encouraging, whether the ability/inability manipulation affects people’s interpretation of those two speech acts, and whether the manipulation differentially affects the interpretation of description compared to encouragement.

In the second experiment, participants read a story about an agent who promises to do something. I manipulated the time at which the agent was no longer able to fulfill the promise. Participants then constructed timeline graphs to represent the time in the story during which several things were true. Among these things were whether the agent has a responsibility to fulfill

the promise, whether it is worthwhile to encourage the agent to fulfill the promise, and whether the agent is able to do so. The key questions here are whether people represent responsibilities as outlasting the ability to fulfill them, and whether they think it is worthwhile to encourage someone to do something after he is no longer able to do it.

The results of such studies could potentially support several different conclusions. On one hand, the results could undermine the dual-function hypothesis by demonstrating that people interpret “ought” claims as describing responsibilities but not as encouraging their fulfillment, or vice versa. On another hand, the results could support the dual-function hypothesis but undermine the interaction hypothesis. That is, it could turn out that people interpret “ought” claims as both describing responsibilities and encouraging their fulfillment, but both functions are similarly affected by the manipulation of ability/inability. For instance, people might view inability as equally irrelevant to whether responsibilities exist and whether it is worthwhile to encourage their fulfillment. On yet another hand, the results could support both the dual function hypothesis and the interaction hypothesis. In this case, people would view “ought” claims as having both functions but whereas inability is irrelevant to its descriptive function, it undercuts its encouragement function. It is only this combination of results that would strongly support the theory that “ought” exceeds but implies “can.”

I conducted one final experiment to provide further information regarding related theoretical issues about the function of normative language, in particular whether an “ought” claim has a primary function and how contextual features affect the interpretation of its functions. In this third experiment, participants again read a story about an agent who promises to do something

and someone else remarks that this “ought” to be done. I manipulated whether the “ought” claim was made in the present or past tense, and whether it was addressed to the person who made the promise or to a third party. Participants recorded judgments about whether the speaker was describing a responsibility, encouraging its fulfillment, and casting blame. The key question here is whether any of the three functions is consistently present, thereby providing some evidence that it is a good candidate for an “ought” claim’s primary function.

Experiment 1

The purpose of this experiment is to investigate the potential for speakers of “ought” claims to be interpreted as *describing moral obligations* and as *encouraging certain behavior*, and whether the interpretation is affected by acknowledged ability or inability to produce the relevant outcome. Broadly speaking, one possible outcome is that there are no meaningful differences in an “ought” claim’s potential to function as a description or as encouragement. Both functions might always be attributed to speakers of “ought” claims, or the two functions might sometimes both be attributed and other times both denied. Another possibility is that the two functions are consistently viewed differently. Perhaps “ought” claims are always more readily interpreted as describing obligations than as encouraging behavior, or vice versa. Or, more complexly, the difference between the two functions might vary depending on whether it is able or unable to produce certain outcomes.

Method

Participants

One hundred and twenty participants were tested (aged 19-73 years, mean age = 36 years; 41 female; 96% reporting English as a native language). Participants were U.S. residents, recruited and tested online using Amazon Mechanical Turk (AMT) and Qualtrics, and compensated \$0.40 for approximately 1-2 minutes of their time. The same recruitment procedures were used for all experiments reported in this paper. Repeat participation was prevented within and across experiments (by AMT worker ID).

Materials and Procedure

Participants were randomly assigned to one of three conditions (control, impossible, possible) in a between-subjects design. All participants read a simple story, responded to four statements, then completed a brief demographic questionnaire. The story was adapted from previous research on related questions and featured an agent who makes a promise and subsequently is unable to fulfill the promise (Sinnott-Armstrong 1984; Buckwalter & Turri 2015).

The story for all conditions began with this paragraph about the agent, Alex:

In Alex's possession is a package. The package contains hand-signed documents for an international real estate deal that must be physically delivered by 4pm. If they aren't physically delivered by 4pm, the clerk's office will close and the deal will definitely be cancelled. There is no leeway: if the package is even the slight-

est bit late, it might as well not have been delivered at all.

The story for the control condition contained one more short paragraph, which states that Alex promised to deliver the package and introduces another person, Alex's sister Julia.

Alex promised to deliver the package by 4pm. Alex's sister, Julia, is with him.

She says, "Alex, you ought to deliver that package by 4pm."

Instead of that one additional short paragraph included in the control condition, the story for the impossible and possible conditions added two other paragraphs. These two paragraphs similarly describe Alex's promise and introduce Julia, but they contain additional information about Alex's ability or inability to deliver the package on time, which is reiterated by Julia when she speaks to Alex (impossible/possible manipulation in brackets):

Alex promised to deliver the package by 4pm. Given current traffic conditions, if he leaves now, it is physically [impossible/possible] for him to deliver the package by 4pm. It is literally [no longer/just barely] possible that he can make it by 4pm. ¶¹ Alex's sister, Julia, is with him. She says, "Alex, you ought to deliver that package by 4pm. At this point, it is literally [no longer/just barely] possible for you to do so."

The purpose of the details in the story's opening paragraph is to unambiguously fix certain aspects of Alex's situation, especially the time during which a specific, familiar action or outcome would still be possible. (Pilot testing confirmed that the details effectively accomplished this.) The purpose of introducing Julia is so that participants can record judgments about some-

¹ Indicates a paragraph break on the participant's screen.

one's communication with Alex. The purpose of the control condition is to provide a baseline for interpreting people's judgments about Julia's speech acts. The purpose of the impossible/possible manipulation is to assess whether it affects people's judgments about Julia's speech acts.

After reading the story, participants rated their agreement or disagreement with four statements (order randomized; here and throughout, all labels are for expository convenience and were not seen by participants):

1. Julia describes a moral obligation that Alex has to go deliver the package by 4pm. (description)
2. Julia is encouraging Alex to go deliver the package by 4pm. (encouragement)
3. If Alex leaves immediately, he can still deliver the package by 4pm. (ability)
4. The real estate deal can still happen if the package arrives late. (deal)

The purpose of the description and encouragement statements is to assess whether people interpreted Julia as performing those speech acts. The purpose of the ability statement was to assess whether people accurately perceived Alex's abilities in the various conditions. The deal statement was included as a further check that people correctly understood the case. Responses were collected using a standard 7-point Likert scale, 1 ("strongly disagree") – 7 ("strongly agree"), left-to-right on the participant's screen.

Results

Participants understood relevant details of the case, as revealed by their response to the deal and ability statements. Mean response to the deal statement was very low ($M = 1.53$, $SD = 1.18$) and

did not differ significantly across conditions, $F(2, 117) = 2.32, p = .103$. Mean response to the ability statement varied appropriately across the conditions, $F(2, 117) = 137, p < .001, \eta_p^2 = .701$, with participants agreeing in the control ($M = 5.49, SD = 1.21$) and possible ($M = 5.76, SD = 1.32$) conditions and disagreeing in the impossible condition ($M = 1.82, SD = 1.01$). Participant age had an unpredicted effect on response to the description statement, $Beta = .292, p = .001$. Older participants were more likely to interpret an “ought” claim as a description of moral obligation. No other demographic effects were detected.

In order to assess how people interpreted the speaker’s speech acts and whether acknowledged ability or inability affected the interpretation, I conducted a mixed design analysis of variance with assignment to condition as a between-subjects factor and response to the description and encouragement statements as two levels of a within-subjects factor. Because of the effect of participant age noted above, I included participant age as a covariate. The analysis revealed a strong interaction between assignment to condition and response to the description and encouragement statements, $Wilks' \Lambda = .769, F(2, 116) = 17.40, p < .001, \eta_p^2 = .231$. (See Figure 1.)

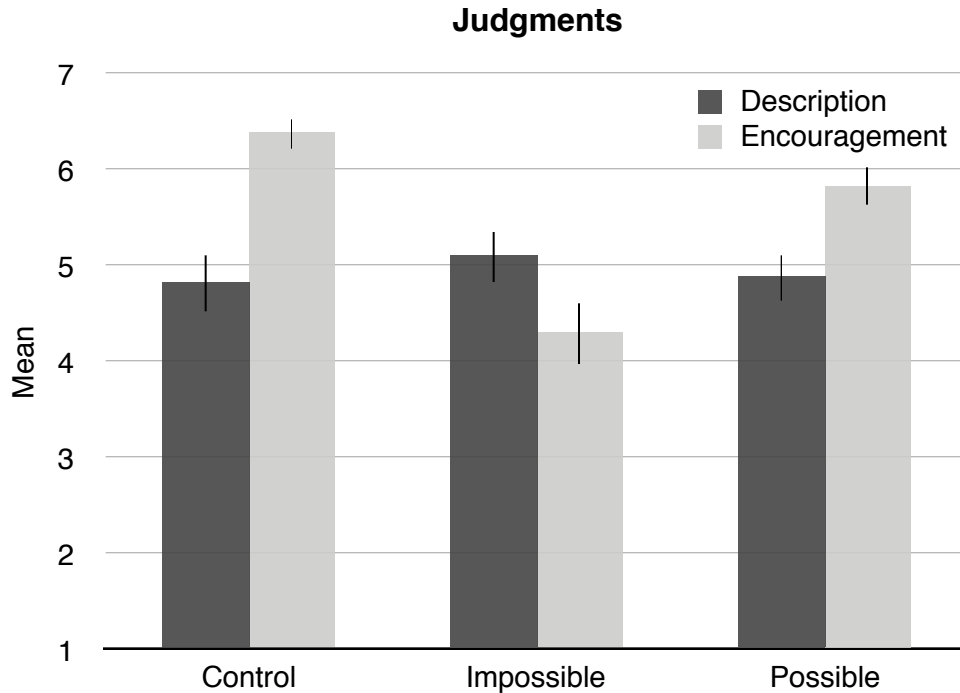


Figure 1. Experiment 1. Mean response to the description statement (the speaker is describing a moral obligation) and the encouragement statement (the speaker is encouraging a specific sort of behavior). Scales ran 1 (“strongly disagree”) – 7 (“strongly agree”). Error bars represent +/- SEM.

To explore this interaction, I conducted paired samples t-tests on the description and encouragement statements in the three conditions separately. In the control condition, mean response to the description statement ($M = 4.82$, $SD = 1.81$) was significantly lower than mean response to the encouragement statement ($M = 6.38$, $SD = 0.94$), $t(38) = -5.06$, $p < .001$, $d = 0.86$. Similarly, in the possible condition, mean response to the description statement ($M = 4.88$, $SD = 1.55$) was lower than mean response to the encouragement statement ($M = 5.83$, $SD = 1.22$), $t(40) = -3.61$, $p < .001$, $d = 0.57$. By contrast, in the impossible condition, mean response to the description statement ($M = 5.08$, $SD = 1.61$) was *higher* than mean response to the encourage-

ment statement ($M = 4.33$, $SD = 1.99$), $t(39) = 2.38$, $p = .023$, $d = 0.38$.

Separate analyses of variance revealed that assignment to condition did not affect participant response to the description statement, $F < 1$, $p = .773$, but assignment to condition strongly affected response to the encouragement statement, $F(2, 117) = 21.29$, $p < .001$, $\eta_p^2 = .267$. Pairwise comparisons using independent samples t-tests revealed that mean response to the encouragement statement was higher in the control condition than in the possible condition, $t(78) = 2.27$, $p = .026$, $d = 0.51$, higher in the control condition than in the impossible condition, $t(77) = 5.86$, $p < .001$, $d = 1.34$, and higher in the possible condition than in the impossible condition, $t(79) = 4.11$, $p < .001$, $d = 0.92$.

Discussion

This experiment investigated the potential for speakers of “ought” claims to be interpreted as describing moral obligations and as encouraging their fulfillment, and whether the interpretation is affected by an agent’s ability or inability to fulfill the obligation. People interpreted the descriptive and encouragement functions differently and, critically, the nature of this difference depended on whether the agent was able or unable to perform the relevant action. In a control condition where neither ability nor inability was mentioned, people interpreted the agent as able to perform the relevant action, and while they tended to interpret the speaker as both describing a moral obligation and encouraging its fulfillment, they more strongly viewed it as encouragement (as indicated by a significantly higher mean agreement with the encouragement statement). A similar pattern was observed in another condition where the agent was able to fulfill the obligation. By

contrast, in a condition where the agent was unable to do so, participants continued to interpret the speaker as describing a moral obligation, but they were significantly less likely to view the speaker as encouraging its fulfillment.

Overall, then, “ought” claims were consistently interpreted as describing moral obligation, regardless of whether the agent was able or unable to fulfill the obligation. By contrast, how strongly the speaker was interpreted as encouraging its fulfillment depended on whether the agent was able or unable to do so. In particular, when the agent was unable, people were significantly less likely to interpret the speaker as encouraging fulfillment of the obligation.

I observed one unpredicted demographic effect on the interpretation of “ought” claims: older participants were more likely to view them as descriptions of moral obligations. Perhaps this is because older participants are more familiar with the phrase “moral obligation.” A search using Google Books Ngram Viewer (<https://books.google.com/ngrams>) provides some support for this hypothesis. “Obligation” and “responsibility” are synonyms (*Oxford American Writer’s Thesaurus*). In the Google Books database, from 1800 until approximately 1995, the phrase “moral obligation” occurred more frequently than the phrase “moral responsibility.” But since then, the reverse is true, with “moral responsibility” occurring more frequently than “moral obligation.” This is probably related to a more general trend over the twentieth century whereby “responsibility” went from occurring roughly as frequently as “obligation” to occurring roughly three times more frequently. Motivated by these observations, in the next experiment I phrased matters in terms of responsibility instead of obligation.

Experiment 2

The purpose of this experiment was to further investigate people views of ability, responsibility, and encouragement. Whereas Experiment 1 studied these issues in the context of interpreting a third-party's speech, this experiment had participants assess matters directly and used a measure other than verbal reports. In order to do this, I asked participants to complete a timeline graphing task to represent the period during which certain propositions were true in a story.

Method

Participants

One hundred and eighteen new participants were tested (aged 18-70 years, mean age = 33 years; 38 female; 99% reporting English as a native language). Participants were compensated \$1.50 for approximately 5 minutes of their time. Data were excluded from four participants who completed the test phase in less than 10 seconds and two participants who had participated in Experiment 1.

Materials and Procedure

Participants were randomly assigned to one of two conditions (drive, email). All participants began by completing a familiarization phase, in order to accustom them to completing timeline graphs and receive feedback on how to perform the task correctly. A supplemental file includes the material used for the familiarization phase. During the testing phase, participants read a short scenario and then graphed bars under several timelines. These bars represent the period during

which statements are true in the scenario.

The testing phase began with these instructions:

Drag the bars to represent the period during which the statements are true in the scenario described below. Remember:

- If a statement is never true in the story, then leave bar blank (the default setting).
- If a statement remains true throughout the whole story, then extend the bar all the way to the right.
- If a statement starts out true but then stops being true, then extend the bar to whichever time in the story it stops being true.

Beneath the instructions was the text of a scenario, which was based on the one used in Experiment 1, except this time it featured only Alex (rather than Alex and his sister). The principal difference between the story in the two conditions is the method by which Alex can deliver the document. The purpose of this manipulation is to change the time during which it is true that Alex can still deliver the document by 4pm, while keeping the two conditions otherwise very closely matched. In the drive condition, Alex possesses a physical, hand-signed document that he promised to deliver by 4pm. The only way he can do this is to drive the document there using his car; in order to deliver it there on time, he must start driving by 3pm. In the email condition, Alex possesses a scanned electronic document that he promised to electronically deliver by 4pm. The only way he can do this is to email the document using his smartphone; in order to deliver it on time, he must email it by 4pm. A supplemental file contains the complete text of the stories used

in both conditions.

Below the story appeared five statements, each underneath a timeline, in randomized order. Participants could tap or drag to graph a bar for each statement's timeline separately. The timelines ran from 1pm to 5pm, left-to-right on the participant's screen, and were anchored with each full hour ("1pm," "2pm," ... , "5pm"). (Figure 2 shows how timeline graphs initially appeared on a participant's screen.) The graphs were adjustable down to one hundredth of an increment (i.e. of an hour), which, from the user's standpoint, seemingly provides continuous variability. Four of the five statements differed slightly across the two conditions, while the fifth was the same. These were the statements (for 1-4, the drive version is listed first)

1. ability

Alex can still deliver the document by 4pm.

Alex can still electronically deliver (email) the document by 4pm.

2. description

Alex still has a responsibility to deliver the document by 4 pm.

Alex has a responsibility to electronically deliver (email) the document by 4 pm.

3. encouragement

It is still worthwhile to encourage Alex to go deliver the document by 4pm.

It is still worthwhile to encourage Alex to email the document by 4pm.

4. ownership

Alex owns a car.

Alex owns a smartphone

5. open

The clerk's office is still open.

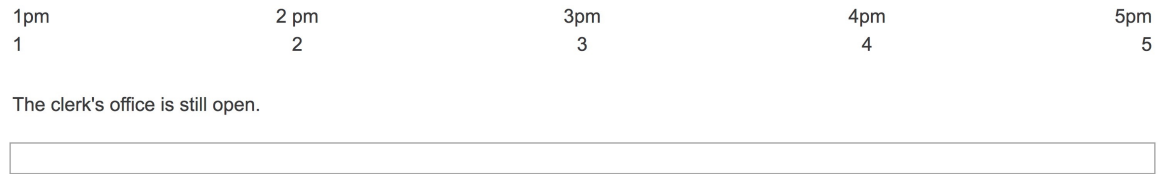


Figure 2. Experiment 2. Example of blank timeline graphs.

Results

Participants understood relevant details of the case, as revealed by their timeline graphs for the ownership, open, and ability statements. (See Figure 3.) Mean estimate on the ownership timeline (“Alex owns a car/smartphone”) ($M = 4.88$, $SD = 0.66$) closely approximated the correct value (= 5) and did not differ across conditions, $t(116) = 0.22$, $p = .828$. Mean estimate on the open timeline (“The clerk’s office is still open”) ($M = 3.96$, $SD = 0.35$) also closely approximated the correct value (= 4) and did not differ across conditions, $t(116) = 0.04$, $p = .970$. Mean estimate on the ability timeline (“Alex can still” deliver the document on time) differed appropriately across the conditions (drive, $M = 3.27$, $SD = 0.68$; email, $M = 4.00$, $SD = 0.04$), $t(116) =$

-8.18, $p < .001$, $d = 1.52$.

In order to assess how people estimated the truth of the description and encouragement statements, I conducted a mixed design analysis of variance with assignment to condition as a between-subjects factor and performance on the two timeline graphs as two levels of a within-subjects factor. The analysis revealed a strong interaction between assignment to condition and performance on the two timeline graphs, Wilk's Lambda = .746, $F(1, 116) = 39.45$, $p < .001$, $\eta_p^2 = .254$. To explore this interaction, I conducted paired samples t-tests on the two timeline graphs in the two conditions separately. In the drive condition, mean estimate on the description timeline ($M = 4.15$, $SD = 0.71$) was significantly higher than mean estimate on the encouragement timeline ($M = 3.29$, $SD = 0.76$), $t(58) = 7.50$, $p < .001$, $d = 0.99$. But in the email condition, mean estimate on the description ($M = 4.06$, $SD = 0.22$) and encouragement ($M = 3.98$, $SD = 0.24$) timelines did not differ significantly, $t(58) = 1.78$, $p = .080$. Independent samples t-tests revealed that mean estimate on the encouragement timeline was significantly higher in the email condition than in the drive condition, $t(116) = 6.58$, $p < .001$, $d = 1.22$, but mean estimate on the description timeline in the two conditions did not differ, $t(116) = -0.88$, $p = .380$.

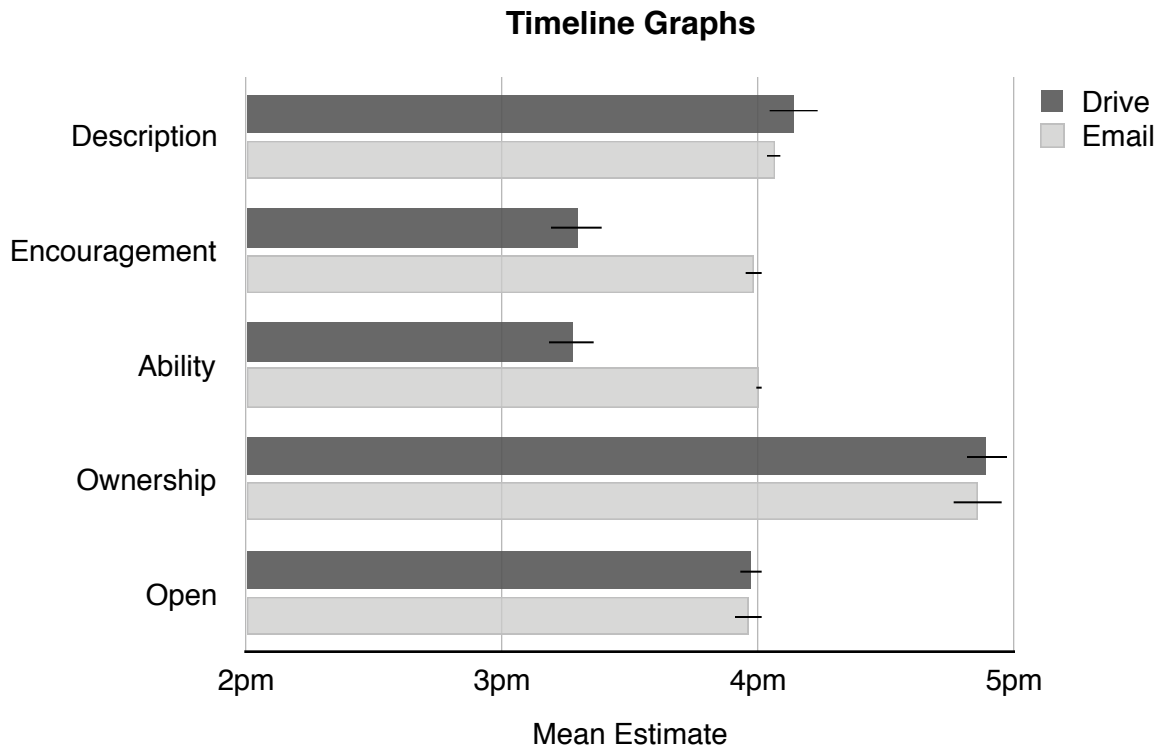


Figure 3. Experiment 2. Results from a timeline graph task for five statements. Participants constructed bar graphs representing the time during which each statement was true in the scenario. Error bars represent +/- SEM.

Discussion

This experiment investigated participants' views about an agent's ability, his responsibilities, and whether it makes sense to encourage him to fulfill responsibilities. People's own graphical estimations of the truth of various propositions revealed that they tend to view the agent as having a responsibility to perform an action significantly beyond the time he was able to perform it. By contrast, people did not think it was sensible to encourage the agent to perform an action beyond the time he was able to perform it. This suggests that the descriptive and encouragement functions of "ought" statements relate differently to an agent's ability to perform.

Experiment 3

The purpose of this experiment is to investigate whether two additional factors affect an “ought” claim’s potential to function as describing, encouraging, or blaming. The factors are the claim’s tense (present or past) and the party to whom it is addressed (to the person who made the promise, or to a third party). Investigating this could be informative regarding an “ought” claim’s primary function.

Method

Participants

Two hundred and two participants were tested (aged 20-72 years, mean age = 37 years; 77 female; 98% reporting English as a native language).

Materials and Procedure

Participants were randomly assigned to one of four conditions in a 2 (tense: present, past) × 2 (person: second, third) design. All participants read a simple story and responded to three test statements. The story was a simplified version of the one used in Experiment 1, with Alex and Julia. The tense factor manipulated whether Alex promised to deliver the package today or yesterday. The person factor manipulated whether Julia’s “ought” statement is addressed to Alex (she speaks to him in the second person), or to a third party (she speaks to someone else about Alex in the third person). A supplemental file contains all stimuli used in this experiment. Partic-

ipants then rated three statements (order randomized) regarding different speech acts Julia might be performing (tense differences in brackets):

1. Julia is describing a moral obligation that Alex [has/had] to deliver the package by 4pm [today/yesterday]. (description)
2. Julia is encouraging Alex to deliver the package by 4pm [today/yesterday]. (encouragement)
3. Julia is blaming Alex for not delivering the package by 4pm [today/yesterday]. (blame)

Responses were collected using a standard 7-point Likert scale, 1 (“strongly disagree”) – 7 (“strongly agree”). Participants then went to a new screen and answered two comprehension questions from memory.

Results

I excluded data from 18 participants who failed a comprehension question, but including them yields the same basic pattern of results reported below. In order to assess how people interpreted the speaker’s speech acts and whether tense or the person addressed affected the interpretation, I conducted a mixed design analysis of variance with tense and person as between-subjects factors and response to the three test statements as levels of a within-subjects factor. The analysis revealed a strong three-way interaction between tense, person, and statement evaluated, Wilk’s Lambda = .786, $F(2, 179) = 24.43$, $p < .001$, $\eta_p^2 = .214$. (See Figure 4.)

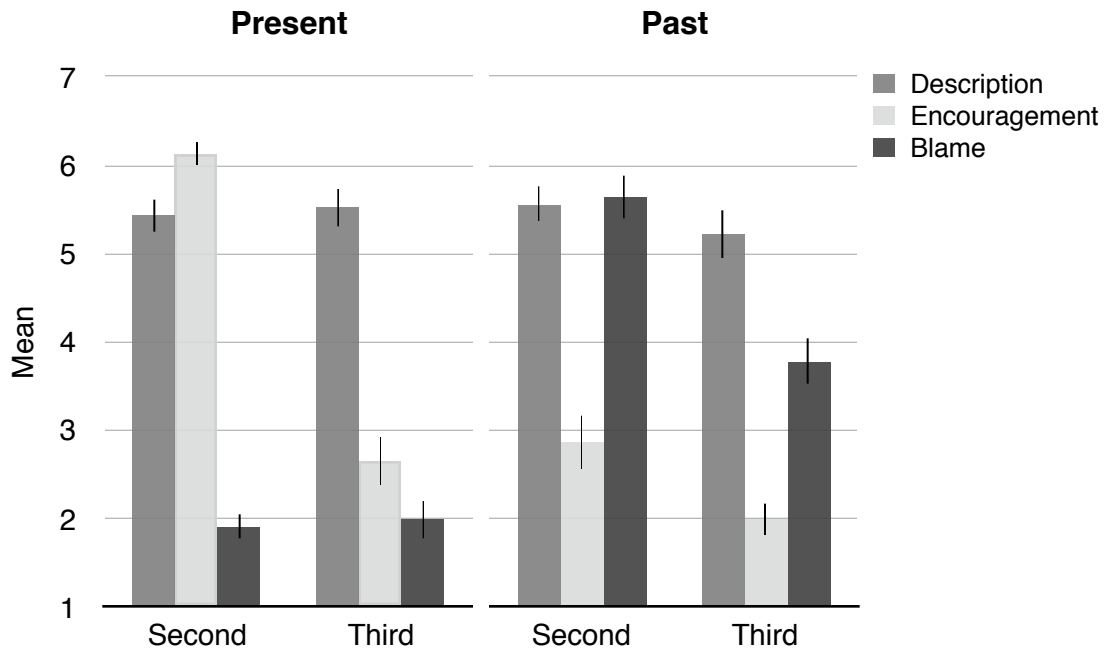


Figure 4. Experiment 3. Mean response to the three speech-act evaluations (within subjects: *description*, *encouragement*, *blame*) in the four conditions (between subjects: *present* or *past* tense crossed with *second-person* or *third-person* address). Scales ran 1 (“strongly disagree”) – 7 (“strongly agree”). Error bars represent +/- SEM.

To explore the interaction, I conducted a two-way analysis of variance for each test statement with tense and person as between-subjects factors. (See Table 1.) These analyses revealed no effect on response to the description statement and significant interaction effects on response to the encouragement and blame statements. One sample t-tests revealed that mean response to the description statement was above the midpoint in all four conditions. (See Table 2.) By contrast, mean response was above the midpoint in only one of the four conditions for the encouragement statement (second-person present condition) and the blame statement (second-person past condition). Moreover, for the encouragement and blame statements, mean response was below the midpoint in all other conditions, except for the third-person past condition, where mean response to the blame statement did not differ significantly from the neutral midpoint (= 4). In-

spection of the histogram for the blame statement in this condition revealed a conspicuous bimodal distribution, which I return to briefly in the discussion section below.

Table 1. Analysis of variance for the three dependent measures.

Measure	Factor											
	Tense				Person				Tense x Person			
	F	df	p	η_p^2	F	df	p	η_p^2	F	df	p	η_p^2
description	0.16	1, 180	.689	.001	0.26	1, 180	.608	.001	1.01	1, 180	.317	.006
encouragement	72.36	1, 180	<.001	.287	89.22	1, 180	<.001	.331	32.53	1, 180	<.001	.153
blame	158.44	1, 180	<.001	.468	16.72	1, 180	<.001	.085	20.05	1, 180	<.001	.100

Table 2. One sample t-tests on the three test statements in the four conditions (test value = 4).

Person	Tense									
	Present					Past				
	t	df	p	MD	d	t	df	p	MD	d
Second										
description	7.09	46	<.001	1.43	1.04	7.88	42	<.001	1.56	1.20
encouragement	15.79	46	<.001	2.13	2.31	-3.64	42	<.001	-1.14	0.56
blame	-15.40	46	<.001	-2.09	2.25	7.25	42	<.001	1.65	1.10
Third										
description	7.14	44	<.001	1.53	1.06	4.84	48	<.001	1.22	0.69
encouragement	-5.02	44	<.001	-1.36	0.75	-11.13	48	<.001	-2.00	1.59
blame	-9.60	44	<.001	-2.00	1.43	-0.82	48	.417	-0.22	0.11

Replicating a finding from the control condition of Experiment 1, in the second-person present condition, mean response was significantly higher for the encouragement statement ($M = 6.13$, $SD = 0.92$) than for the description statement ($M = 5.43$, $SD = 1.38$), $t(46) = 2.74$, $p = .009$, $d = 0.46$, in a context where both were significantly above midpoint.

Discussion

This experiment investigated the potential for “ought” claims to be interpreted as description, encouragement, and blame. The results show that they can function in each of these ways, but the functions related differently to the claim’s tense (present or past) and the person to whom the statement is addressed (in the second person or third person). “Ought” claims were consistently interpreted as describing moral responsibilities, regardless of the tense or addressee. By contrast, people tended to interpret the claim as encouragement only when it was in the present tense and addressed to the person it was about, and they tended to interpret it as blame only when it was in the past tense and addressed to the person it was about. Overall, the results provide some evidence for the hypothesis that description is an “ought” claim’s primary function, which I will take up further in the General Discussion.

In the third-person past-tense condition, mean response to the blame statement did not differ from the neutral midpoint, and this was underlain by a bimodal distribution. One potential explanation for this is that some participants interpreted “ought” in a moral sense (which is my primary focus here), whereas others interpreted it in an epistemic sense. “Ought” can express a prediction based on indirect evidence, as it would in a case where someone says, “They left two hours ago, and it takes only ninety minutes to get there, so they ought to have arrived by now.” When Julia says to a third party, “Alex ought to have delivered the package yesterday,” it is not hard to hear that in the epistemic sense, so some participants might opt for that reading instead of a harsher reader involving blame.

General Discussion

Three experiments advanced our understanding of the functions of normative language. People view claims about what an agent “ought” to do as having dual potential, namely, the potential to *describe moral obligations* and the potential to *encourage agents to perform actions* (Experiment 1). But these two functions relate differently to an agent’s ability or inability to perform. When an agent is unable to perform an action, people are significantly less likely to interpret someone saying he “ought to do it” as encouraging the agent to do it; by contrast, an agent’s inability to perform did not make people less likely to interpret the “ought” claim as describing a moral obligation that the agent has (Experiment 1). Utilizing a novel timeline graphing task, I also found that people view an agent’s responsibility as significantly outlasting his ability to fulfill it, whereas people did not view it as worthwhile to encourage him to fulfill it beyond the time he was able to do so (Experiment 2). Aside from functioning to describe and encourage, in some contexts “ought” claims can also function to cast blame (Experiment 3). Overall, these findings shed light on the function and interpretation of normative language, specifically “ought” claims, which are important in everyday human life and society.

The findings are also theoretically significant. More specifically, they support the theory that *ought* exceeds but implies *can*. According to this theory, “ought” statements have (at least) two functions, *describing* moral responsibilities and *encouraging* behavior, and these functions relate differently to an agent’s abilities. The descriptive function is not constrained by ability: people can have moral responsibilities that they are unable to fulfill. But the encouragement

function is constrained by ability: it is not worthwhile to encourage people to do things that they are unable to do. Aside from now being empirically supported by the present studies, this theory can help explain persistent disagreement in philosophy over the “ought implies can” principle. Some philosophers say that the principle is obviously true, whereas others say it is obviously false. One explanation of this disagreement is that one side’s intuitions track the encouragement function, whereas the other side’s intuitions track the descriptive function.

Whereas many philosophers have accepted that “ought” claims function to describe, encourage, and blame, some have speculated that the encouragement function is “primary” (for a review, see Van Roojen 2013). The present results provide some evidence against this view. When an “ought” claim’s addressee and tense were manipulated in a single design, it consistently functioned to describe moral responsibilities (Experiment 3). By contrast, it functioned as encouragement only when addressed to the person it was about, in the present tense. Similarly, it clearly functioned to cast blame only when addressed to the person it was about, in the past tense. By the same token, the present findings provide some support for the hypothesis that the descriptive function is primary. The descriptive function was stably present across all the different contexts tested, demonstrating robustness against manipulations that strongly affected the other functions. This is what we would expect from an expression’s primary function. Moreover, given two further assumptions, the descriptive function’s primacy could also explain why “ought” claims can function to encourage and blame. On the one hand, if people tend to be morally motivated, then a claim whose primary function is to describe their moral responsibilities can also function secondarily to encourage specific behavior. On the other hand, if failing to

fulfill a moral responsibility is typically blameworthy, then describing an unfulfilled moral responsibility could be a way to indirectly cast blame. Of course, factors other than ability, addressee, and tense might affect an “ought” claim’s function. For instance, perhaps it is affected by the type or severity of outcomes hanging in the balance. So perhaps further work will reveal sufficient complexity and variability to undermine the hypothesis that the descriptive function is primary. Further work on this question is needed before drawing firm conclusions.

The present findings also relate to ongoing debates regarding the relationship between moral responsibility and ability in commonsense morality. Researchers have repeatedly found that people attribute moral responsibilities in tandem with the inability to fulfill them (see the Introduction for references). In many studies, 80-90% of participants attributed responsibility in tandem with inability. In response, some critics allege that these findings are misleading due to temporal ambiguity in interpreting participant responses (e.g. Kurthy, Lawford-Smith & Sousa 2017; see also Streumer 2003). Perhaps participants systematically index “responsibility” and “ability” judgments to different points in the scenario and in a way that invalidates a straightforward interpretation of their responses. For instance, participants might agree with responsibility attributions if *at any point in the scenario* the agent had a responsibility, while agreeing with the ability attribution if *at any point in the scenario* the agent was able to perform. If so, then there might be no point in the scenario where participants attribute both responsibility and inability. This objection is unpersuasive partly because it is ad hoc and partly because researchers have tested cases in which it is made clear that at no point was the agent able to perform the relevant action (e.g. Buckwalter & Turri 2015: Experiment 5). Additionally, the novel timeline graphing

task, used in Experiment 2, provides a completely different and conclusive response to the objection. Instead of providing verbal answers that could potentially be implicitly indexed to different points in the scenario, participants graphed timelines to visually represent the time during which a statement was still true. In the relevant condition, their “responsibility” graphs significantly exceeded their “ability” graphs. In other words, people drew simple pictures showing that, in their view, some responsibilities outlast the ability to fulfill them.

Future research could extend the present findings in several ways. For instance, I studied only scenarios involving promising. This has the advantage of being a familiar and frequent source of moral responsibility and pedagogy, being a frequent topic in the theoretical literature on moral responsibility, and being used in prior experimental research on moral judgments. But it has the disadvantage of missing potentially important differences for other common sources of moral responsibility. Perhaps responsibilities arising from being a parent or political leader, on the one hand, and responsibilities arising from making promises or other explicit commitments, on the other hand, relate differently to inability. Further research is needed to evaluate this possibility.

Acknowledgments — For helpful feedback and discussion, I thank Wesley Buckwalter, Ori Friedman, Walter Sinnott-Armstrong, Angelo Turri, and Sarah Turri. Thanks also to Steven Slovic and three anonymous reviewers for *Cognition*. This research was supported by the Social Sciences and Humanities Research Council of Canada, the Ontario Ministry of Economic Development and Innovation, and the Canada Research Chairs program.

References

- Barker, S. J. (2000). Is value content a component of conventional implicature? *Analysis*, 60(3), 268–279.
- Boyd, R. (1988). How to be a moral realist. In G. Sayre-McCord (Ed.), *Essays on moral realism* (pp. 181–228). Ithaca, NY: Cornell University Press.
- Buckwalter, W. (ms). Theoretical refutation of “ought implies can.” University of Waterloo.
- Buckwalter, W., & Turri, J. (2014). Inability and obligation: compelling counterexamples to “ought implies can.” Presented at the Buffalo Experimental Philosophy Conference, Buffalo, NY.
- Buckwalter, W., & Turri, J. (2015). Inability and Obligation in Moral Judgment. *PLOS One*, 10(8), e0136589. <http://doi.org/10.1371/journal.pone.0136589.g004>
- Carnap, R. (1937). *Philosophy and logical syntax*. London: Kegan Paul, Trench, Trubner & Co.
- Chituc, V., Henne, P., Sinnott-Armstrong, W., & De Brigard, F. (2016). Blame, not ability, impacts moral 'ought' judgments for impossible actions: Toward an empirical refutation of 'ought' implies “can.” *Cognition*, 150(C), 20–25. <http://doi.org/10.1016/j.cognition.2016.01.013>
- Gibbard, A. (1990). *Wise choices, apt feelings*. Cambridge, Mass.: Harvard University Press.
- Hare, R. M. (1952). *The language of morals*. Oxford: Oxford University Press.
- Henne, P., Chituc, V., De Brigard, F., & Sinnott-Armstrong, W. (2016). An Empirical Refutation of ‘Ought’ Implies ‘Can’. *Analysis*, 76(3), 283–290. <http://doi.org/10.1093/analys/an->

w041

- Huemer, M. (2005). *Ethical intuitionism*. New York: Palgrave Macmillan.
- Kant, I. (2009). *Religion within the bounds of bare reason*. (W. S. Pluhar, Trans.). Indianapolis: Hackett.
- Kurthy, M., Lawford-Smith, H., & Sousa, P. (2017). Does ought imply can? *PLOS One*, 12(4), e0175206. <http://doi.org/10.1371/journal.pone.0175206.s001>
- Mackie, J. L. (1977). *Ethics: inventing right and wrong*. New York: Penguin.
- Mizrahi, M. (2015). Ought, can, and presupposition: an experimental study. *Methode*, 4(6), 232–243.
- Moore, G. E. (1959). *Philosophical papers*. New York: Collier Books.
- Pigden, C. R. (1990). Ought implies can: Erasmus, Luther, and R.M. Hare. *Sophia*, 29(1), 2–30.
- Sinnott-Armstrong, W. (1984). 'Ought' conversationally implies "can." *Philosophical Review*, 93(2), 249–261.
- Stevenson, C. L. (1937). The emotive meaning of ethical terms. *Mind*, 46(1), 14–31.
- Stevenson, C. L. (1944). *Ethics and language*. New Haven: Yale University Press.
- Streumer, B. (2003). Does 'ought' conversationally implicate 'can'? *European Journal of Philosophy*, 11(2), 219–228.
- Turri, J. (2015). *Compatibilism and incompatibilism in social cognition*. Presented at the Society for Philosophy and Psychology, Duke University.
- Turri, J. (2016). *Compatibilism and incompatibilism in social cognition*. *Cognitive Science*. <http://doi.org/10.1111/cogs.12372>

Turri, J. (2017). Compatibilism can be natural. *Consciousness and Cognition*, 51, 68–81. <http://doi.org/10.1016/j.concog.2017.01.018>

Turri, J. (ms). Pathways from inability to blamelessness in moral judgment. University of Waterloo.