Ought Does Imply Can

Miklos Kurthy[1]*, Holly Lawford-Smith[1], Paulo Sousa[2]

(Paper under review. Please do not quote without permission)

1 Department of Philosophy, University of Sheffield, United Kingdom

2 Institute of Cognition and Culture, Queen's University Belfast, United Kingdom

* Corresponding author

E-mail: mkurthy1@sheffield.ac.uk (MK)

Abstract

Most philosophers believe that a person can have an obligation only insofar as she is able to fulfil it, a principle generally referred to as “Ought Implies Can”. Arguably, this principle reflects something basic about the ordinary concept of obligation. However, in a paper published recently in this journal, Wesley Buckwalter and John Turri presented evidence for the conclusion that ordinary people in fact reject that principle. With a series of studies, they claimed to have demonstrated that, in people's judgements, obligations persist irrespective of whether those who hold them have the ability to fulfil them. We argue in this paper that due to crucial problems in their design, Buckwalter & Turri's conclusions are not warranted. We present the results of four studies demonstrating that ordinary people do judge obligations to be constrained by perceptions of inability after all. In other words, for ordinary people, “ought” does imply “can”.

Introduction

The concept of obligation seems to constitute a fundamental component of social and moral cognition [1–3]. Although there is considerable cultural variability in terms of how people understand the content, source and ground of obligations, anthropological evidence indicates that all human societies deploy the concept of obligation to organise human action and interaction [4, 5]. Obligations are deemed constraints that motivate social and moral behaviours. They are also deployed to understand and evaluate these behaviours. In particular, the non-fulfilment of an obligation is thought to constitute wrongdoing and may warrant blame. Thus, it is difficult to underestimate the importance of studying the folk concept of obligation.

According to a distinguished philosophical tradition that dates back at least to Kant, obligations are in force only when the persons holding them are able to fulfil them [6]

(although the depth of Kant's commitment to the principle is debated [7]). This idea, which has been accepted by most moral philosophers, albeit in a number of different prescriptive and descriptive guises [8–10], is now widely known as the principle that "Ought Implies Can" (henceforth, "OIC principle"). (The OIC principle is often discussed in terms of its equivalent contraposition "Inability Eliminates Obligation", and much of our discussion will follow this trend).

The OIC principle seems to be an elaboration of the apparently quotidian belief that humans cannot be constrained to do what they are unable to. More strongly, the principle may bespeak something fundamental about the ordinary concept of obligation and the inferential relations it enters into. However, Wesley Buckwalter and John Turri [11] (henceforth, B&T) have recently presented evidence for the conclusion that in fact ordinary people do not endorse the OIC principle, which is to say, do not believe that whether an agent is able to fulfil an obligation has a significant bearing on the question of whether the obligation holds. (See also [12], and discussion in [13]).

B&T conducted a series of studies to test whether ordinary people endorse the OIC principle. (B&T were also interested in probing whether people have more difficulty in perceiving inability when the source of the inability is mental rather than physical, e.g. due to clinical depression. Since this issue is tangential to the OIC principle, we leave it completely aside in this paper). In their studies, participants had to read stories in which an obligation is created, but the person initially able to fulfil the obligation is subsequently described as unable. For instance, participants in one study were asked to consider a case in which an agent ("Walter") promised to pick his friend ("Brown") up from the airport (the promise creating the obligation), but later became involved in a car accident and thereby rendered physically unable to keep the promise. Participants were then presented with the OIC probe, asking them to choose one of the following (randomly sequenced) statements:

1. Walter is obligated to pick up Brown at the airport, but Walter is not physically able to do so.
2. Walter is not obligated to pick up Brown at the airport, and Walter is not physically able to do so.
3. Walter is obligated to pick up Brown at the airport, and Walter is physically able to do so.
4. Walter is not obligated to pick up Brown at the airport, but Walter is physically able to do so.

In this and other scenarios—varying *inter alia* the source of the obligation involved (e.g., a promise or a social role), the type of inability (e.g., a physical restriction or a constraining feature of the environment), and the seriousness of the consequences of the obligation not being fulfilled (minor or fatal)—participants overwhelmingly chose the first option: "obligated, but not able". On the face of it, this choice contradicts the OIC principle, since it attributes to the individual both an obligation and the inability to fulfil it.

Moreover, to confirm that participants understood the situation as involving a literal inability to fulfil the obligation, B&T's studies included, after the OIC probe, an inability-comprehension probe, asking subjects whether the person under the obligation was literally unable to fulfil it. The great majority of participants confirmed that there was literal inability, and eliminating the few participants who denied literal inability did not change the general pattern of B&T's results. B&T concluded their paper with the claim that "commonsense moral cognition rejects the principle that ought implies can" [11].

B&T's studies testing the OIC principle also included, after the inability-comprehension probe, a blame probe, investigating whether participants would consider the individuals in their stories blameworthy for not fulfilling their obligations. (B&T's paper also had a separate study probing the relation between inability and blame without testing the OIC

principle.) They found that the great majority of participants denied that the individual is to blame in this respect, and suggested on the basis of this finding that, for ordinary people, "Blame Implies Can". It is important to note that the traditional view of the relation between blame and obligation as far as inability is concerned is that the presence of an inability undermines blame by eliminating the perception of wrongdoing—in particular, by eliminating the perception that someone did something wrong in not fulfilling her obligation because in fact the obligation was cancelled by the inability [14, 15]. Therefore, given that the above results indicate that the presence of an inability undermines blame without cancelling the obligation (and hence without eliminating wrongdoing), B&T also suggest that the traditional view of the relation between blame and obligation does not exist in ordinary cognition, and may be an invention of philosophers trying to "validate excuses" [16].

In this paper, we question B&T's claims with new evidence. We argue first that there are crucial problems with their design testing the OIC principle and report some evidence supporting our criticism. Then, we provide the rationale for an improved design, and report four new studies using this design. These studies show that the great majority of participants adhere to the OIC principle, consistent with widespread theoretical endorsement of the principle and with the idea that the principle reflects something basic about the ordinary concept of obligation. They also suggest that people take the presence of an inability to undermine blame, which, together with our OIC results, suggests that ordinary cognition is in line with the traditional view on the relation between blame, obligation, and wrongdoing.

Problems with the design

Aspects of B&T's design, in particular the way they framed the list of options of the OIC probe, make the option "obligated, but not able" the sole plausible answer, though in a trivial way that does not compromise the OIC principle.

B&T's stories are characterized by an individual under an obligation who afterwards is described as unable to fulfil the obligation. There is an obvious but trivial sense in which each story, taken as a whole, involves both an obligation and an inability. The inability creates tension with the expectation of fulfilment generated by the obligation. The option "obligated, but not able" matches this description of the story as a whole while the other options do not, since they either exclude an obligation ("not obligated, and not able") or include the ability to fulfil the obligation ("obligated, and able"; "not obligated, but able"). Moreover, the option "obligated, but not able" has an ordinary temporal reading (i.e., "obligated, but *subsequently* not able") that mirrors the temporal narrative of the story (i.e., an obligation is made salient early in the story, then later an inability is made salient). This, too, renders the option "obligated, but not able" the best description, because it captures the temporal dimension of the contrast involved in the story as a whole.

In sum, according to our interpretation when participants choose the option "obligated, but not able", they are not saying that the person is *still* under the obligation even if there is inability to fulfil it. That would be evidence against the OIC principle. Rather, they are saying that the stories involve a contrast between an obligation (made salient first) and an inability to fulfil the obligation (made salient second). This does not constitute any evidence at all against the OIC principle, because it may well turn out that subjects in B&T's studies would accept the obligation for as long as they think there is ability, and reject the obligation after the inability is made evident in the story. This would be consistent with the OIC principle.

Two predictions are suggested by this interpretation. There would be a substantial reduction of "obligated, but not able" responses in B&T's studies if one were to replace the connectives "but" and "and" in the original options with connectives that more clearly convey the main point of the OIC probe (i.e., that make participants focus on whether there is a logical relation between the concepts of obligation and ability). There would also be such a

reduction if one were simply to invert the order of the obligation and inability clauses of the original options (e.g., changing "obligated, but not able" to "not able, but obligated"), thus creating a mismatch between the order of the clauses and the temporal narrative of the story.

To test these predictions, we ran a between-subjects study using B&T's "Walter" scenario with three types of OIC probe options: the original four options described above (original condition); four options using "even if" and "because" as connectives, instead of "but" and "and" (logical relation condition); and the original four options with the order of the obligation and inability clauses inverted (inverted-order condition). The OIC-falsifying and OIC-confirming options of the original, logical relation, and inverted-order conditions were as follows (for the sake of simplicity, we leave aside the two options where Walter was described as able to fulfil his obligation):

1. (*Original*) Walter is obligated to pick up Brown at the airport, but Walter is not physically able to do so.

   (*Logical relation*) Walter is obligated to pick up Brown at the airport, even if Walter is not physically able to do so.

   (*Inverted order*) Walter is not physically able to pick up Brown at the airport, but Walter is obligated to do so.
2. (*Original*) Walter is not obligated to pick up Brown at the airport, and Walter is not physically able to do so.

   (*Logical relation*) Walter is not obligated to pick up Brown at the airport, because Walter is not physically able to do so.

   (*Inverted order*) Walter is not physically able to pick up Brown at the airport, and Walter is not obligated to do so.

The results of this study are shown in Fig 1. In the original condition, we replicated B&T's results: 88% chose "obligated, but not able", while only 12% chose "not obligated,

and not able" (N = 41). In the logical relation condition, we completely reversed B&T's results: only 5% chose "obligated, even if not able", while 88% chose "not obligated, because not able" (N = 42; 7% chose the remaining two options where Walter is described as able). Finally, in the inverted-order condition, the two relevant options were equally chosen: 47.5% chose "not able, but obligated" and 52.5% chose "not able, and not obligated" (N = 40). These results not only give strong support to our claim B&T's design does not test the OIC principle, but also suggest that the great majority of people adhere to the OIC principle when the design is improved to make the main point of the task clear.

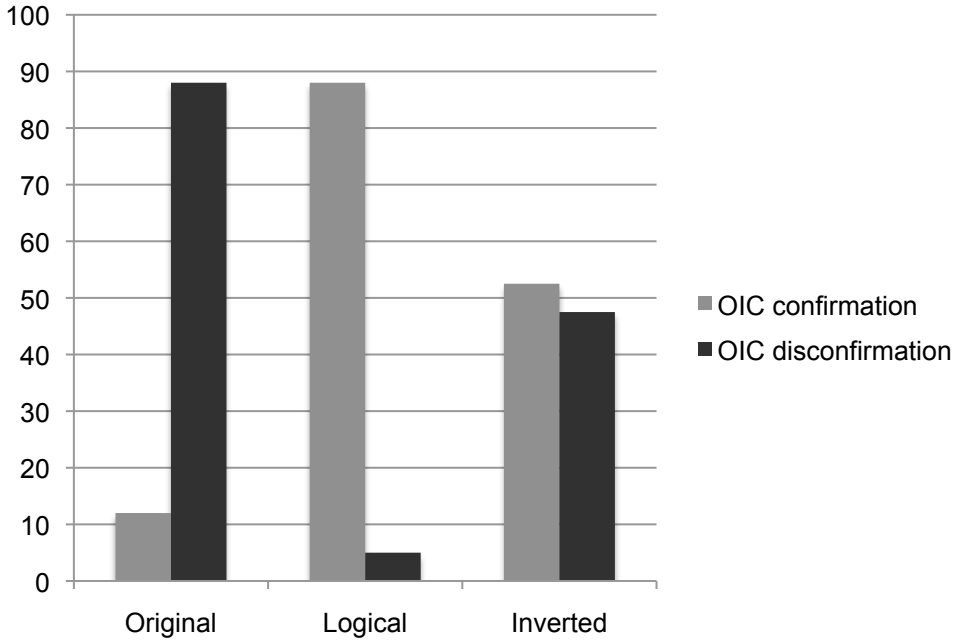

Fig 1. Percentage of responses confirming or disconfirming the OIC principle in each of the three conditions.

There is another aspect of B&T's design that may have contributed to the problem we have outlined and consequently to the choice of the "obligated, but not able" option. B&T's instructions for the OIC probe ("choose the option that best applies") implies that there is a factually correct alternative among the options, and may suggest to participants that they are

being tested on whether they interpreted the story correctly (as if the OIC probe had the same type of function as the inability-comprehension probe—the second probe of their design described earlier). If participants understood the OIC probe in this way, then rather than providing their personal opinion on the logical relation between the concepts of obligation and inability, they would simply provide the *best description* of what is involved in the story as a whole, which is plausibly the option "obligated, but not able", as discussed above.

Finally, it is important to note that none of B&T's stories explicitly state the obligation at stake in the story. In the promise scenario, the story says only that someone makes a promise; in their social-role scenarios, it says only that someone has a social role (e.g., that of a lifeguard); in another scenario, it simply describes a situation in which a small child is drowning and there is a stranger around who could easily help the child. Thus, the participant has to infer from the information given in the initial part of the story (i.e., from the fact that someone made a promise, that someone has a social role, or that someone could easily help) the existence of the corresponding obligations (i.e., the obligation to keep the promise; the obligation related to the social role; the obligation to help the drowning child). True, these inferences are somewhat obvious, and the fact that the obligations are left implicit in the stories is not a problem in itself. However, given the aforementioned problems, it may well be that at least some participants took B&T's OIC probe also as a test on whether they believe that the initial situation described in the story entailed an obligation, and chose the first option to confirm that they indeed believe that there is an obligation involved in the story.

Overview of current studies

To address the above issues and improve the design in order to test the OIC principle, we modified B&T's design in the following ways:

(i) We changed some very trivial details of the stories to make it clearer to participants that the characters in the stories are unable to fulfil their obligation, and/or to avoid misinterpretations of the story.

(ii) We changed the instructions of the OIC probe and the inability-comprehension probe to make their different purposes obvious to participants.

(iii) We positioned the inability-comprehension probe before the OIC probe, that is, just after participants read the story. And in case a participant denied that the character in the story was literally unable to fulfil their obligation, we explained to the participant that in fact the character *was* unable to do so by emphasising the relevant elements of the story; then we asked the participant to assume that there was literal inability before answering the OIC probe. (In our studies, hardly any participants disagreed that the character was literally unable to fulfil their obligation and excluding these participants from the analysis changes nothing in terms of our results and conclusions.)

(iv) We simplified the OIC probe by reducing its four options to two: one confirming the OIC principle, another disconfirming it. (Note that the two eliminated options, which say that the character in the story is able to fulfil her obligation, are completely irrelevant to testing the OIC principle.)

(v) We phrased the two options of the OIC probe in a way that makes it clearer to participants what the logical point of the OIC probe is (e.g. using the connectives "because" and "even if" instead of "and" and "but").

(vi) We included a justification probe asking participants to explain their OIC choice, in order to gain some qualitative insight into the reasons motivating participants' choices. (This step was introduced after the OIC option was irreversibly selected, so there is no reason to suppose that it could interfere with the quantitative results of the OIC probe).

The great majority of the above changes should not be controversial, as they merely clarify and/or simplify the task for the participants. Although changing the connectives of the options of the OIC probe may seem controversial (see "v"), one of our studies will give support to it (see Study 2).

Some of B&T's studies are, arguably, much less central to testing the OIC principle (e.g., Experiment 7, which tests whether the difference between moral and legal obligation is relevant to the principle). Accordingly, our studies focused on those of their studies that are most central to the OIC principle, namely, Experiments 1, 2, 4 and 5.

Our data collection methodology was similar to that employed by B&T. In all studies to be reported, participants were recruited, tested and compensated online. We used Amazon Mechanical Turk and Qualtrics as the online platforms. All participants were U.S. residents. Each participant was paid $0.50 for approximately 4 minutes of their time. Following B&T, in all studies we collected around 40 responses per condition. Participants were allowed to participate in only one of the studies (or conditions) reported in this paper.

Our research design was reviewed by the Research Ethics Committee of the School of History and Anthropology at Queen's University, Belfast, UK and by the Research Ethics Committee of the University of Sheffield, UK.

Study 1: Promise

There are subtle disagreements in the philosophy of language concerning the main function of promises (see e.g. [17, 18]) and in normative theory concerning how and why promises generate obligations (see e.g. [19–21]). However, it seems relatively uncontroversial for both philosophers and ordinary people that by making a promise one creates a socially acknowledged obligation to fulfil it. In this study, we tested the OIC principle in relation to obligations generated by promises, using the "Walter" scenario discussed previously—the

scenario of B&T's first experiment, where they found that 80% of participants chose the "obligated, but not able" option, apparently contradicting the OIC principle. In addition, we tested the principle in relation to different ordinary expressions that are commonly thought to encode the concept of obligation ("obligated", "duty", "ought"), in order to see whether there is variation in judgements as a result of these.

Method

Participants

Participants were 127 adults (60 female; 67 male; M_{age} = 33.95; *SD* = 11.54; range = 53; 97% reporting English as their first language).

Design, Materials and Procedure

After indicating informed consent, participants read the following story (divergences from B&T's wording of the story are in italics):

> Walter promised that he would pick up Brown from the airport. But on the day of Brown's flight, Walter is in a serious car accident *and is hospitalized*. As a result, Walter is not able to pick up Brown at the airport.

We added "and is hospitalized" to boost the understanding that Walter is unable to pick up Brown at the airport.

Participants were then presented with the inability-comprehension probe, whose instruction and question were as follows: "First, we would like to ask you a question to check whether you understood the story. According to the story, is the following statement true?" The statement that participants had to evaluate was: "Walter is literally unable to pick up Brown at the airport because Walter is hospitalized". If they answered "yes", they were presented with the OIC probe. If they answered "no", they were given an explanation indicating that Water is indeed unable to pick Brown up because his "injuries are so serious

that he requires hospitalization"; then they were asked to assume that this is the case before answering the OIC probe.

The instruction and question of the OIC probe were as follows: "Now, we would like to know your personal opinion about the situation. There isn't a correct answer here. Which statement best reflects your personal opinion about the situation?" Participants had to choose between two randomly sequenced statements, each confirming or disconfirming the OIC principle. In order to test the OIC principle with different ordinary expressions that encode the concept of obligation ("obligated", "duty" or "ought"), participants were randomly assigned to one of three phrasing conditions:

1. Under these circumstances, Walter is still obligated to (Walter still has a duty to / Walter still ought to) pick up Brown at the airport, even if he is unable to do so.
2. Under these circumstances, Walter is not obligated to (Walter does not have a duty to / it is not the case that Walter ought to) pick up Brown at the airport, because he is unable to do so.

After choosing one of the above statements, participants were asked to justify their choice: "Please explain why you marked this option".

Finally, participants answered a blame probe, enquiring about the degree to which they believed that Walter deserved blame for not fulfilling the obligation: "To what extent is Walter to blame for not picking up Brown?" Participants answered this probe on a seven-point scale, with "1" indicating "No blame", "4" indicating "Moderate blame", and "7" indicating "Full blame".

Results

Almost everyone (98%) agreed initially that Walter was literally unable to pick up Brown at the airport. The phrasing conditions produced no effect, χ^2 (2, 127) = .01, p = .99,

with 100%, 98% and 100% of participants choosing the option confirming the OIC principle in the "obligated", "duty" and "ought" conditions respectively. Across the phrasing conditions, 126 out of 127 participants chose the option confirming the OIC principle—goodness of fit against chance: χ^2 (1, 127) = 123.03, $p < .001$, $\varphi = 0.98$.

Blame ratings did not differ across phrasing conditions either—$F(2, 124) = 1.04$, $p = .36$. In general, blame ratings were very low ($M = 1.47$; $SD = 1.02$), with 92 of 127 participants opting for the "1" rating (i.e., "No blame").

Discussion

With our improved design, we completely reversed B&T's results using three ordinary expressions that are commonly thought to encode the concept of obligation, suggesting that there is no variation in judgement due to the examined terminological variation in this domain.

Participants' justifications suggest that, actually, there was no evidence at all against the OIC principle in our sample. Justifications of participants who chose the "not obligated" option were consistent with the principle. They often expressed that, given the inability, it would be unintelligible to attribute an obligation, or that it is self-evident that the obligation does not hold:

> "It seems silly to say that it's immoral to not keep a promise in extenuating circumstances like this."

> "It makes no sense to say he should do something he isn't able to."

> "Because he is unable to do so, it is self-explanatory."

And sometimes they explicated the OIC principle literally or in terms of its equivalent contraposition:

> "'Duty' assumes he will have the ability to implement his duty, just as a soldier is excused from duty when injured."

> "I think that the existence of a duty presupposes the ability to fulfil that duty. If it is impossible for that duty to be fulfilled, it does not exist."

> "If someone is unable to do something they can't be obligated to do it."

Now, the justification of the only participant who chose the "obligated" option suggests that, instead of denying the OIC principle, the participant simply shifted the scope of the obligation at stake:

> "Walter made an agreement with full intention of keeping it and if he cannot fulfill the agreement, notice should be sent and a proxy should be appointed to carry out the agreement as specified."

In other words, rather than maintaining that Walter is still obligated to pick up Brown at the airport even if he is unable to do so, this participant seems to be saying that even if Walter cannot pick Brown up, he is still obligated to *do something else* to improve Brown's situation. Since our scenario leaves open the possibility that Walter could still do something else in this respect, the response of this participant does not necessarily conflict with the OIC principle (this kind of justification will show up in later studies; we'll refer to it as the 'scope-shifting problem', because it involves participants' changing the scope of the obligation to include new or alternative content).

Finally, the great amount of "no blame" answers plus the overall low mean of blame ratings shows that participants think that Walter's inability eliminated his blameworthiness for not picking up Brown at the airport, which is consistent with B&T's blame results. However, contrary to B&T's results and claims, our results also suggest that participants think that the elimination of blame was linked to the fact that Walter had no related obligation under the circumstances, and, consequently, to the fact that Walter did not do anything wrong in not picking up Brown at the airport. In other words, our results are more consistent with the idea that ordinary cognition is in line with the traditional view on the relation between blame, obligation and wrongdoing.

Study 2: Playground safety worker

Social roles are normally seen as another source of obligations. In this study, we tested the OIC principle in the context of an obligation entailed by the social role of a playground safety worker. The scenario we utilized corresponds to the one of B&T's second experiment, where they found that 98% ("duty" phrasing condition) and 88% ("ought" phrasing condition) of participants chose the "obligated, but not able" option, apparently contradicting the OIC principle. In addition, we tested whether the framing of our options in terms of the connectives "even if" and "because" inadvertently biased participants towards choosing the option that confirms the OIC principle.

Method

Participants

Participants were 86 adults (40 female, 45 male, 1 "other"; M_{age} = 37.67; *SD* = 13.25; range = 53; 98% reporting English as their first language).

Design, Materials and Procedure

Participants read first the following story:

Michael is a playground safety worker. He sees some broken glass in an area where kids sometimes play barefoot. But he is stricken by a sudden *full body* paralysis *that immobilizes him to the extent that he cannot even speak*. As a result, Michael is not able to *remove* the *broken* glass.

The first two modifications of the original scenario were to boost the understanding of inability and/or to emphasize that there wasn't anything else that Michael could have done to improve the situation (e.g., ask other people to remove the broken glass), and thus to try to avoid the scope-shifting problem identified in the discussion of Study 1. The last modification replaced the verb "pick up" with the verb "remove," which more clearly describes the content of Michael's obligation in this situation.

Participants were then presented with the inability-comprehension probe, which asked them to evaluate the truth of the following statement: "Michael is literally unable to remove the broken glass from the area because he is completely immobilized." Depending on their truth evaluations, participants proceeded to the OIC probe as specified in Study 1.

The instruction and question of the OIC probe were the same as previously. Since we showed that different ordinary expressions encoding the concept of obligation do not affect the results of the OIC probe, we used only one phrasing for the statements of the probe in this study ("obligated"). However, participants were still randomly assigned to one of two conditions. In the "explicit" condition, participants had to choose between the same type of "obligated" statements of Study 1, while in the "implicit" condition these statements were presented without the inability clauses and their connectives:

1. Under these circumstances, Michael is still obligated to remove the broken glass, even if he is unable to do so (Under these circumstances, Michael is still obligated to remove the broken glass).

2. Under these circumstances, Michael is not obligated to remove the broken glass, because he is unable to do so (Under these circumstances, Michael is not obligated to remove the broken glass).

We included the implicit condition in this study because one may argue (rather implausibly in our view) that, rather than making more explicit the main point of the OIC probe, the connectives "because" and "even if" inadvertently bias participants to choose the option consistent with the OIC principle, thus distorting the results. Against this "framing" hypothesis, we predicted that there would be no effect of condition, since the fact that we asked the comprehension probe first plus the usage of "under these circumstances" and "still" already makes the main point of the OIC probe clear enough.

After answering the OIC probe, participants answered the justification probe and the blame probe, similarly to Study 1.

Results

Almost everyone (99%) accepted initially that Michael was literally unable to remove the broken glass. There was no effect of condition, χ^2 (1, 86) = .387, p = .53, with 88% and 84% of participants choosing the "not obligated" response in the explicit and implicit conditions, respectively. Thus, altogether, the overwhelming majority of participants (86%) believed that Michael did not have an obligation under the circumstances—goodness of fit against chance: χ^2 (1, 86) = 44.69, p < .001, φ = .72).

Blame ratings remained low (M = 1.79; SD = 1.41), with 59 of 86 participants opting for "No blame". A 2(condition) x 2(OIC option choice) between-subjects ANOVA on blame scores revealed a main effect of option choice, F(1, 82) = 35.6, p < .001, η_p^2 = .303, but no main effect of condition (p = .17) or interaction (p = .30). Thus, participants who chose the "obligated" option saying that Michael was obligated to remove the glass blamed him more (M = 3.67, SD = 1.67) than participants who chose the option that he was not obligated (M =

1.49, *SD* = 1.11). Accordingly, there was a significant correlation between option choice and blame ratings: $r_{pb} = .53$, $p < .01$.

Discussion

Once again, we completely reversed B&T's results. Furthermore, as we predicted, whether the OIC options involved the inability clauses and their connectives did not affect which option was chosen. This indicates that an argument according to which the effect observed in Study 1 depends on our specific framing of the options, and, in particular, on the usage of the connectives "even if" and "because", is not plausible. Indeed, our results provide further corroboration for our contention that it is B&T's design (rather than ours) that systematically distorts the results.

Justifications for "not obligated" responses were again consistent with an acceptance of the OIC principle. In contrast, the justifications of the "obligated" responses (12 in total) were more varied and, overall, less consistent with a rejection of the OIC principle. Evincing the scope-shifting problem discussed in Study 1, some participants seem to have shifted the scope of the obligation to the idea that Michael still has the obligation to do (or try to do) something else to improve the situation:

> "He has the job of playground safety worker, and he has been presented with an unsafe condition. If he can't remove the glass, he should call out to the kids to avoid the area, call out to another adult, or make some kind of effort to communicate the hazard."

> "In some way if he knows there's broken glass and no one else is notified, there needs to be a way he can communicate with someone he can or warn the kids about it."

Since these participants seem to have misinterpreted our scenario in that they still envisaged that Michael could do something else, like informing other people, to improve the situation

(or since the description of our scenario does not rule out the possibility that Michael could at least make an effort to improve the situation), their justifications are not incompatible with the OIC principle.

Some participants seem to emphasize that Michael still has the obligation to remove the glass, not at the time of his paralysis but rather as soon as he recovers:

> "Well Michael may be unable to physically remove it himself, but he is obligated to do so in the sense that he should remove it as soon as possible."

> "(…) Of course if his condition worsens or doesn't let up then he cannot act on his obligation so he won't clean up the glass, but with the knowledge he should do it, if he can."

This type of justification suggests that in fact these participants accept the OIC principle.

Many participants seem to appeal to the connection between the obligation and the nature of Michael's social role (note that the word "responsibility" is often used in the sense of obligation related to a social role [22, 23]):

> "It is still his responsibility as a playground safety worker."

> "That's his job."

> "It's his property. It's his responsibility to get it cleaned up even if he can't do it himself."

> "I believe as a worker and having knowledge makes you responsible."

From these justifications, one may take that these participants indeed reject the OIC principle—the participants seem to believe that obligations related to social roles continue to be in force independent of the circumstances, and hence seem to accept that Michael is still obligated to remove the broken glass in that situation of inability. However, it is still possible that these participants may have answered "obligated" simply to emphasize that specific and strong obligations are generally entailed by social roles, without necessarily rejecting the OIC principle: Michael has, as a playground safety worker, a defeasible obligation to remove dangerous objects from the playground; he has this obligation in general even if in this specific case it is defeated by his inability.

Finally, the large number of "no blame" answers and low mean of blame ratings, along with the positive correlation between these ratings and OIC option choices (i.e., more blame, more "obligated" response) is more consistent with the idea that ordinary cognition is in line with the traditional view of the relation between blame, obligation, and wrongdoing.

Study 3: Lifeguard

In this study, we tested the OIC principle again in the context of an obligation entailed by a social role again, this time that of a lifeguard. While studies 1 and 2 involved an "internal" inability coming from physical restrictions, this study involves an "external" inability coming from constraints of the environment like distance in space. Furthermore, while studies 1 and 2 involved relatively minor consequences like not being picked up at the airport or stepping on broken glass, this study involves a life-and-death situation. The scenario we utilized corresponds to the one in B&T's fourth experiment, where they found that 93% of participants chose the "obligated, but unable" option that apparently contradicts the OIC principle.

Method

Participants

Participants were 42 adults (11 female, 31 male; M_{age} = 38.98; *SD* = 13.13; range = 49; 98% reporting English as their first language).

Design, Materials and Procedure

Participants read the following story:

> Jessica is the only lifeguard at a remote ocean beach. Two struggling swimmers are about to drown, *and no one else is around except Jessica*. She rushes in to save them, but because of the *great* distance between the swimmers, it is physically impossible for her to rescue both swimmers. Jessica rescues one swimmer but not the other.

The main modifications of the original scenario were again introduced in order to boost the understanding of inability and/or to emphasize that there wasn't anything else that Jessica could have done to improve the situation (e.g., ask for additional help). (Other minor stylistic modifications, not indicated here, were also introduced to improve readability).

The rest of the procedure was exactly the same as in studies 1 and 2: inability-comprehension probe ("Jessica is literally unable to rescue both swimmers because they are too far apart"); OIC probe with justification probe; blame probe. In this study, there was only one OIC probe condition, with the following options:

1. Under these circumstances, Jessica is still obligated to rescue both swimmers, even if she is unable to do so.
2. Under these circumstances, Jessica is not obligated to rescue both swimmers, because she is unable to do so.

Results

Almost everyone (95%) agreed that Jessica was literally unable to save both swimmers. The great majority (79%) of participants felt that the agent was not obligated to save both swimmers—goodness of fit against chance: χ^2 (1, 42) = 13.71, $p < .001$, $\varphi = 0.57$).

Blame scores remained relatively low (M = 1.67; SD = 1.18), with 28 of 42 participants opting for "No blame". However, in contrast with the previous study, participants choosing the "obligated" option did not ascribe significantly more blame to Jessica than participants choosing the "not obligated" one: $t(40) = 1.64$, $p = .207$, $d = .49$ ("obligated": M = 2.11; SD = 1.45; "not obligated": M = 1.55; SD = 1.09). Accordingly, there was no significant correlation between option choice and blame ratings: $r_{pb} = .19$, $p = .207$.

Discussion

Yet again, in sharp contrast to B&T's findings, the "not obligated" option was clearly preferred, even in a case in which the consequences are severe (the death of a swimmer).

Moreover, again, while the justifications of the "not obligated" responses were consistent with the idea that participants indeed accept the OIC principle, the justifications of "obligated" responses (9 in total) were less consistent with a real rejection of the principle.

The great majority of "obligated" responses evinced the scope-shifting problem, in this case insisting that Jessica had a further obligation to *try to* save both swimmers:

> "Even if she thinks and it would be physically impossible, she should still make as much of an effort as possible to try to save both swimmers."

> "She should still make an attempt to do whatever she can do."

> "It is her employment obligation to at least attempt to rescue both. One at a time."

> "She should at least try to save them since we don't know if she can fail or not."

"It is her duty as a lifeguard to do the best she can with what she has. Despite her being unable to rescue both people, she has to be moral enough to try to save both."

Since our scenario does not rule out the possibility that Jessica can try to save both swimmers, these justifications are not incompatible with the OIC principle.

Again, some participants seemed to appeal to the connection between the obligation and the nature of Jessica's social role:

"The conditions of the rescue could change however her job as a lifeguard does not change"

"She was the only one there, it was her job."

As we discussed in Study 2, these justifications may indicate real rejection of the OIC principle. Alternatively, similarly to what we suggested, they may indicate that, with their "obligated" response, the participants are simply emphasising the defeasible but general obligation that is entailed by the social role of a lifeguard, without yet accepting that the obligation was in force in that specific situation—that is, without rejecting the OIC principle.

Finally, although the positive correlation between blame ratings and OIC option choices was not statistically significant, the large number of "no blame" answers and low mean of blame ratings are still more consistent with the view that ordinary cognition aligns with the traditional view of the relation between blame, obligation, and wrongdoing.

Study 4: Drowning child

Our first three studies featured obligations created either by the agent through a social action (a promise), or by the social role of the agent (safety worker, lifeguard). In this final study, we feature a case in which the obligation does not come from a promise or a social

role, but from the situation—a drowning child creating an obligation to help. The scenario corresponds to that in a particular condition ("recent") of B&T's fifth experiment, where they found that 88% of participants chose the "obligated, but unable" option that apparently contradicts the OIC principle.

Method

Participants

Participants were 41 adults (12 female, 29 male; M_{age} = 37.29; *SD* = 12.00; range = 42; 100% reporting English as their first language).

Design, Materials and Procedure

Participants first read the following story:

> Michael is relaxing in the park near a pond when he sees a small girl fall in. She is drowning and definitely will die unless someone quickly pulls her out. This part of the park is secluded and Michael is the only person around. But Michael is stricken by a sudden *full body* paralysis. As a result, Michael is not able to save the girl.

We used "full body paralysis" instead of the original "leg paralysis" on the premise that this phrasing would be perceived as more of an incapacitating condition, and also as an attempt to preclude the scope-shifting problem (in a pilot study using the scenario with "leg paralysis", a participant with an "obligated" response suggested that Michael should "at least try to crawl to save the girl").

The rest of the procedure was the same as in the previous studies: comprehension probe ("Michael is literally unable to save the small girl because he is completely paralyzed"); OIC probe with justification probe; blame probe. As in Study 3, there was only one OIC probe condition, with the following two options:

1. Under these circumstances, Michael is still obligated to save the small girl, even if he is unable to do so.
2. Under these circumstances, Michael is not obligated to save the small girl, because he is unable to do so.

Results

Almost all participants (98%) agreed that Michael was literally unable to save the girl. The great majority of participants (73%) thought that Michael was not obligated when there was an inability to fulfil the obligation—goodness of fit against chance: χ^2 (1, 41) = 8.80, $p <$ .01, $\varphi = .46$.

Although "No blame" was still the modal rating (18 out of 41 participants), blame scores were noticeably higher in this study (M = 2.73; SD = 2.1). For example, a t-test revealed that the blame scores in Study 3 and Study 4 differed significantly, $t(61) = 2.84$, $p =$ < .01, $d = 0.72$ (equality of variances not assumed). Moreover, a t-test showed that, similarly to Study 2 (but unlike in Study 3), blame scores were significantly higher for participants choosing the "obligated" option than for those choosing the "not obligated" option: $t(39) =$ 5.15, $p < .001$, $d = 1.65$ ("obligated": $M = 4.91$; $SD = 2.02$; "not obligated": $M = 1.93$; $SD =$ 1.48). Finally, there was a strong, significant correlation between statement choice and blame ratings: $r_{pb} = .636$, $p < .001$.

Discussion

We again reversed B&T's results, although, of the four studies, this one had the lowest percentage of "not obligated" responses.

However, an analysis of the justifications of "obligated" responses (11 in total) suggests that this study was beset by a major problem. About half of the participants do not seem to have maintained the assumption of literal inability when answering the OIC probe,

mostly because they took the full bodily paralysis to be a controllable emotional reaction (involving especially fear):

> "He needs to overcome his fear and save the girl."

> "You have to overcome your fear a person's life is at stake."

> "It was just an emotional reaction which he could overcome."

> "Michael is responsible to get control of himself and save the girl. He can control his emotion and reactions and needs to pull himself together."

> "He is responsible to save her even if he SEEMS unable to do it. I believe his perception of being paralyzed is not real."

If these justifications indeed correspond to the reason why participants chose the "obligated" response, then their responses are not inconsistent with the OIC principle after all.

Some participants' responses revealed the scope-shifting problem again in terms of obligation to try, which, as we already discussed, is not incompatible with the OIC principle:

> "He is obligated to at least TRY. If he can't, he can't. Maybe the water is deep and he can't swim. But he should at least try no matter what."

> "I have never heard of a sudden full body paralysis like this, and it seems like Michael should still be trying to help."

A few participants emphasized that there was a (moral) obligation in the situation:

"He had a duty to act, a moral obligation. His fear paralyzed him and he was unable to act."

"He is morally obligated to save the girl."

"Well I assume nothing has changed about the girls [sic] situation just because Michael can't move so the obligation to save her is still there, even if he can't move it still exists."

These justifications seem indeed to indicate that the participants reject the OIC principle.

The fact that the overall mean of blame ratings was a bit higher in this study (in comparison with studies 2 and 3) is not incompatible with the view that inability undermines blame, since the mean was substantially affected by the ratings of the participants with "obligated" responses that did not assume inability as shown by their justifications (with these participants eliminated from the analysis, the overall blame mean drops from "2.73" to "2.25", which is much closer to, and non-significantly different from, the overall mean of studies 2 and 3). Moreover, a large number of participants still chose the "no blame" answer. Finally, these blame ratings plus the strong correlation between blame ratings and OIC choice indicate that ordinary cognition is in line with the traditional view of the relation between blame, obligation, and wrongdoing.

General Discussion

In four studies, dealing with different types of obligations and situations of inability, we showed that the great majority of participants think that a person is not under an obligation if she is not able to fulfil it, completely reversing the results that B&T obtained with their research design (see Fig 2). Study 1 showed that the obligation to fulfil a promise is annulled when the agent is not able to fulfil it. This study also indicates that this is the case irrespective

of the particular term used to express the concept of obligation ("obligated", "duty" or "ought"). Using a different scenario, Study 2 demonstrated that these results do not depend on our particular use of connectives—rather, it is B&T's results that appear fragile in this respect, as shown in the section "Problems with the design". Studies 3 and 4 extended these findings to cases in which the consequences are more serious (the death of a person).

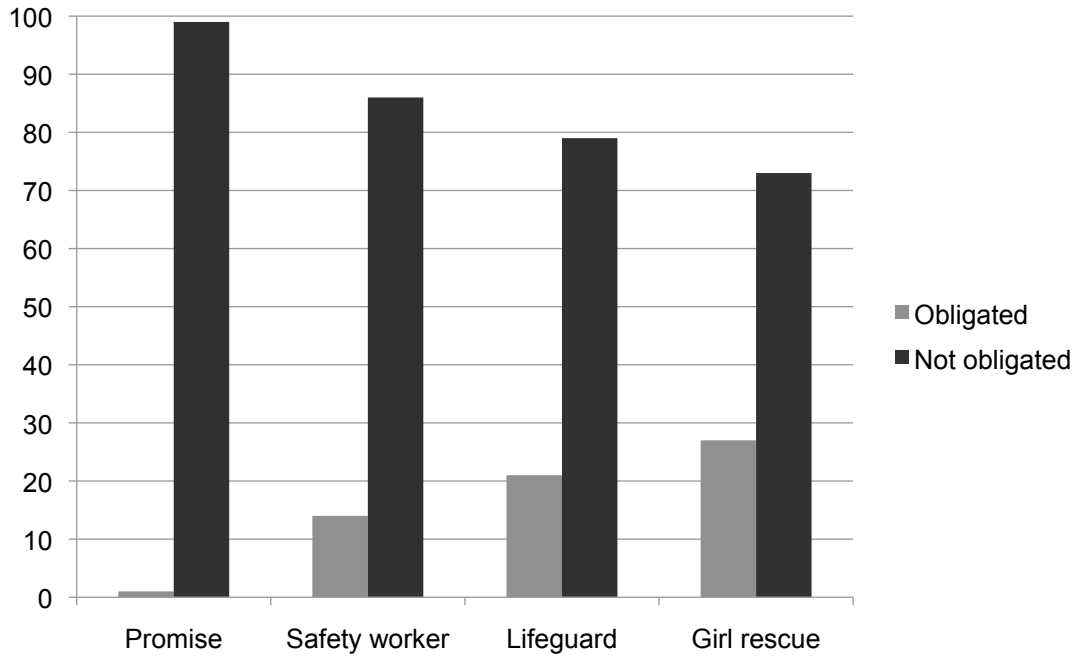

Fig 2. Percentage of responses to the OIC probe in each of the studies.

Studies 2, 3, and 4 still saw a relevant minority of participants choosing the "obligated" response, suggesting that there may be some individual variation in this domain. However, a substantial part of "obligated" responses still seems to derive from misinterpretation of the OIC probe and/or the scenarios, as evinced by justifications demonstrating the "scope-shifting" problem, which appeared across all studies, and by justifications showing that the participants did not keep the assumption of inability, which appeared in Study 4. Of course, if this is correct, it raises the question as to why there was such misinterpretation. The scope-shifting problem may be a result of participants' inclination

to blame the person specifically for not *trying* to do her best to minimize the bad consequences of the situation, something our studies did not control for. The misinterpretation of "full body paralysis" in terms of controllable emotional reaction in study 4 may have a similar explanation.

Moreover, one may raise the question of why there may have been an increase in misinterpretation between study 1 and study 4 correlated with the increase in "obligated" responses. There is a sense in which the consequence of scenario 4 (the death of a small girl) is worse than that of scenario 3 (the death of an adult), which in turn is worse than that of study 2 (the risk stepping on a broken glass), which in turn is worse than that of study 1 (not being picked up at the airport). (A small study, $N = 25$, asking participants to rate these scenarios in terms of their seriousness confirmed this hierarchy). Thus, it is also possible that this increase in seriousness may have contributed to the increase in the amount of misinterpretation from scenario 1 to 4, by pushing participants to see the situation as less determined and hence to be more hopeful about a positive outcome.

If our take on the minority responses is correct, the range of individual variation suggested by our sample is rather small—only a few participants rejected the OIC principle. This in turn raises the broader issue of why there is such a consensus on the topic—with nearly everyone accepting the OIC principle. A plausible hypothesis is that this principle is the philosophical codification of an entailment that is a core element of the set of inferential relations associated with the folk concept of obligation. More broadly the principle may reflect something about how the domains of deontic and metaphysical modals are structurally related [24]. This hypothesis is consistent with the justifications of "not-obligated" responses, as these often expressed how intuitive and self-evident the response was, or how unintelligible it would be to answer otherwise.

Turning to blame attributions, in all studies a large number of participants attributed no blame to the individual for the fact that the obligation was not fulfilled. The mean of blame ratings was low in all studies too. They were highest in study 4, but this was mainly due to the fact that some participants did not maintain the assumption of inability appropriately. Thus, overall, our results indicate that, for ordinary people, inability undermines blame, which is consistent with B&T's own results on blame.

Contrary to B&T's claim that blame attributions are unrelated to obligation attributions, the low percentage of the "obligated" responses plus the correlations between blame ratings and OIC probe choices (i.e., more blame, more "obligated" response) in our results are much more consistent with the idea that ordinary cognition is in line with the traditional view that blame reduction is related to obligation elimination via the elimination of wrongdoing.

In sum, our results show that, for ordinary people, ought does imply can, and that blameworthiness is related to both obligation and wrongdoing.

Acknowledgements

We wish to thank Stephen Laurence and Luca Barlassina for their valuable comments and thoughtful suggestions.

Author contributions

Designed and performed the studies, analysed the data, and wrote the paper: MK HLS PS.

References

1. Jackendoff R. Language, consciousness, culture: Essays of mental structure. Cambridge, MA: MIT Press; 2007.

2. Beller S. Deontic norms, deontic reasoning, and deontic conditionals. Thinking and Reasoning. 2008;14:305–341.

3. Beller S. Deontic reasoning squared. In Love BC, McRae K, Sloutsky VM, editors. Proceedings of the 30th Annual Conference of the Cognitive Science Society. Austin, TX: Cognitive Science Society; 2008.

4. Brown DE. Human Universals. New York City: McGraw-Hill; 1991.

5. Bender A, Beller S. Polynesian tapu in the 'deontic square': A cognitive concept, its linguistic expression and cultural context. In: Alterman R, Kirsh D, editors. Proceedings of the 25th Annual Conference of the Cognitive Science Society. Mahwah, NJ: Lawrence Erlbaum Associates Inc; 2003.

6. Kant I. [Critique of pure reason]. Norman Kemp Smith. London: Macmillan; 1933.

7. Stern R. Does 'ought' imply 'can'? And did Kant think it does? Utilitas. 2004;16(1):42–61. esp. 52–61 p.

8. Griffin J. The human good and the ambitions of consequentialism. In: Paul EF, Miller FD, Paul J, editors. The good life and the human good. Cambridge: Cambridge University Press; 1992.

9. Sapontzis S. 'Ought' does imply 'can'. Southern Journal of Philosophy. 1991;29(3):382–393.

10. Vranas PBM. I ought, therefore I can. Philosophical Studies. 2007;136(2):167–216.

11. Buckwalter W, Turri J. Inability and obligation in moral judgment. PLOS One. 2015;10(8): e0136589. doi:10.1371/journal.pone.0136589

12. Mizrahi M. Ought, can, and presupposition: An experimental study. Methode – Analytic Perspectives. 2015;4(6):232–243.

13. Kurthy M, Lawford-Smith H. A brief note on the ambiguity of ‘ought’. Reply to Moti Mizrahi's ‘Ought, can and presupposition: An experimental study’. Methode – Analytic Perspectives. 2015;4(6):244–249.

14. Hieronymi P. The force and fairness of blame. Philosophical Perspectives. 2004;18(1):115-148.

15. Levy N. The good, the bad, and the blameworthy. Journal of Ethics and Social Philosophy. 2005;1(2):2–16.

16. Turri J, Blouw P. Excuse validation: a study in rule-breaking. Philosophical Studies. 2015;172(3):615–34.

17. Austin JL. How to do things with words. London: Oxford University Press; 1955.

18. Searle J. Speech acts: An essay in the philosophy of language. Cambridge: Cambridge University Press; 1969.

19. Smith H. A paradox of promising. Philosophical Review. 1997;106(2):153–196.

20. Owens D. A simple theory of promising. Philosophical Review. 2006;115(1):51–77.

21. Shiffrin S. Promising, intimate relationships, and conventionalism. Philosophical Review. 2008;177(4):481–524.

22. Hart HLA. Postscript: Responsibility and retribution. In: Punishment and responsibility. Oxford: Oxford University Press; 1968.

23. Sousa P. A cognitive approach to moral responsibility—the case of a failed attempt to kill. Journal of Cognition and Culture. 2009;9:171-194.

24. Joerden JC. Deontological square, hexagon, and decagon: A deontic framework for supererogation. Logica Universalis. 2012;6:201–216.