I Don’t Know

Christopher Walker
chris.walker@ischool.berkeley.edu
August 22, 2015
https://github.com/cw25/i_dont_know_experiment

"The only true wisdom is in knowing you know nothing." -- Socrates

INTRODUCTION

Admitting that we do not know the answer is hard. We are swimming in information, promised that it can educate us and inform our decisions. Certainty is valued, but, of course, one person cannot know everything. There will always be questions to which the only truly correct answer is: “I don’t know.”

People frequently navigate around unknown answers by guessing, essentially choosing not to admit to the things they do not know. Ideally, with so much data around us, we should be guessing less since we know more. But does the presence of data give us a false sense of certainty? Are we less likely to admit we don’t know when we have data to lean upon? What if we don’t know anything about where the data came from?

This experiment seeks answers to those questions. We presented participants with questions whose only true answer is “I don’t know.” We then measured how the presence of data about previous responses affected their willingness to admit that they did not know.

EXPERIMENTAL DESIGN

Our primary hypothesis for this experiment was:

H1: When presented with summary statistics about previous responses to a question, people will be less likely to answer that question with “I don’t know.”

The outcome we sought to measure was the number of times each participant responded with “I don’t know” or an equivalent phrase. To properly measure the rate at which people admitted they did not know, we had to first generate questions that had no factually correct answer. If the question had a knowable answer, participants who actually knew it would not be faced with the choice that we wanted them to face.

As our impossible questions, we planned to ask participants to look at pictures of ten fake animals and identify them. The intent was to find pictures of real animals that had been altered, but not altered so drastically that they were obviously fake. We had to maintain the illusion that the animals were real so the participant would be faced with the decision of either guessing the animal's identity or admitting that they did not know it.

After an extensive search of Google Images, it became clear that such images were very difficult to find. Most of them were obviously jokes, and were too absurd to be useful. Because they were so scarce, we only found two usable fake animal images. In fact, one of the images used in the final study was believed to be fake until after the pilot study was conducted, when it was revealed to be a hairless bear.

Faced with so few options, and without the skills necessary to create new images, we made the decision to use pictures of very obscure animals to fill out the remaining eight questions. This choice introduced a potential bias. Participants who knew the identities of the animals would presumably never answer "I don't know." This problem affected both treatment and control since they were both shown the same set of animals. In cases where the participant knew any of the animals, there was less opportunity for answering "I don't know." This could create underestimation of the dependent variable since the maximum measurable value of the outcome would then shrink from 10 by the number of animals known to the participant.

By choosing images of animals that are not commonly known, we hoped to reduce the likelihood that such conditions would arise. For this, we leaned on our own perceptions and knowledge of animals, which admittedly might be flawed as well. To further hedge against this problem, we used our pilot studies as an opportunity to identify images that were at higher risk for positive identification. Two pilot studies were conducted, each with 20 participants recruited through Amazon Mechanical Turk. Some minor changes were made to the layout of the survey (i.e. removing some unnecessary question labels), but otherwise, the animal images and their accompanying statistics were the only changes made as a result of the pilot.

After evaluating the responses to the first pilot, the four animal images (saiga antelope, proboscis monkey, African wild dog, and tapir) that had the highest rate of positive recognition were replaced. Responses to the second pilot study were evaluated to ensure that the rates of positive identification for those questions had dropped. The second set of images was ultimately used in the final study (see Appendix A).

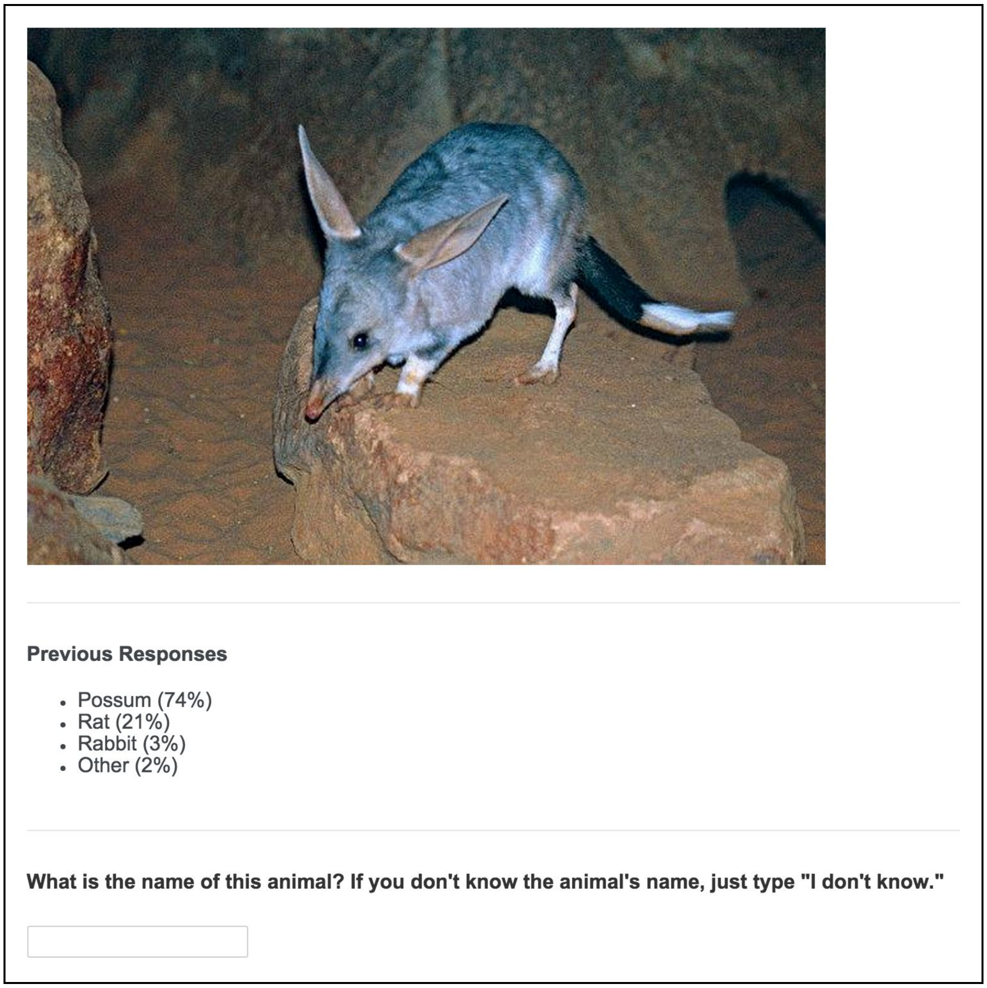
Control group participants were shown the animal picture, and below it saw the prompt: *What is the name of this animal? If you don't know the animal's name, just type "I don't know."* Treatment group participants were shown the same animal and same prompt, but between the image and the prompt, they were also shown fabricated summary statistics about previous responses (see Figure 1).

The summary statistics were labeled as “Previous Responses” in an effort to use the most generic labeling possible. We considered using a label that indicated that the responses came from an expert source, but decided against it. Our goal was to demonstrate how the presence of data in general affect “I don’t know” responses, rather than how expert opinions affect them. We felt the “Previous Responses” verbiage was more in line with that goal, and would be more generalizable to other scenarios.

Two levels of treatment were employed -- moderate and strong. In the moderate treatment, the previous response with the highest confidence was between 55% and 62%. In the strong treatment, the previous response with the highest confidence was between 71% and 80%. The differing levels of treatment were used to test a second hypothesis:

H2: Seeing previous responses that express a high degree of confidence in the top answer has a greater effect than seeing previous responses that were less confident in that same answer.

Figure 1. A screenshot of one of the prompts shown to participants under the strong treatment.



Previous Responses

- Possum (74%)
- Rat (21%)
- Rabbit (3%)
- Other (2%)

What is the name of this animal? If you don't know the animal's name, just type "I don't know."

Participants were provided with a web site link (http://goo.gl/dyGfp1) that forwarded them to a randomization script (http://people.ischool.berkeley.edu/~chris.walker/w241/index.php). The randomization script generated a random number between 100000000 and 999999999. If the number was odd, the participant was forwarded to the control group version of the survey. If the number was even, another random number within the same range was generated. If that second random number was odd, the participant was forwarded to the moderate treatment survey. If it was even, they were sent to the strong treatment survey. All three surveys were identical aside from the presence or absence of the previous response summaries.

The introductory page provided basic instructions and prompted the participant for their gender, age, and whether they held a Bachelor's degree or higher. These values were captured for use in the covariate analysis phase. The additional hypotheses were:

H3: Males receiving the treatment will be less likely than females to answer "I don't know."

H4: Holders of 4-year college degrees receiving the treatment will be even more likely to answer "I don't know" than those who do not hold degrees.

H5: Participants over 30 years old who receive the treatment will be more likely to answer "I don't know" than those aged 30 or younger.

The ten pages asking the animal identification questions were served to each participant in random order to eliminate any possibility of the sequencing of the images introducing bias. After the animal identification questions, a manipulation check question was presented, asking the participant how helpful they found the summary statistics to be (including an option to indicate that no statistics were present). The final page of the survey provided a debriefing message and offered participants who were recruited into the study via Mechanical Turk a randomly generated participation code that they could redeem for payment.

RECRUITMENT AND DECEPTION

Participants were recruited into the study through multiple avenues. The author of the study posted personal appeals to Facebook, LinkedIn, Twitter, Slack, and two private email lists to which he is a member (see Appendix B). Facebook and LinkedIn participants were also encouraged to share the post in an effort to maximize exposure to potential participants.

In another effort to maximize the sample size and statistical power for the experiment, a block of 1000 HITs on Amazon Mechanical Turk was also purchased. Turk participants were offered $0.50 apiece to complete the survey until either 1000 responses had been captured or the allotted time for the study ran out.

The survey was left open for one week (August 2, 2015 through August 9, 2015), during which a total of 589 responses were captured. 201 participants entered through Mechanical Turk and the remaining 388 entered via the links shared over social media.

Some deception was necessary in the recruitment phase so as not to bias the results of the experiment. The recruitment message made no mention of the true purpose of the study. Instead, the survey was referred to only as an "animal identification" study. By referring to it in this way, we ensured that no undue attention was given to the presence or lack of summary statistics, nor to the availability of the option to answer "I don't know." The debriefing message found on the last page of the survey did provide a more truthful indication of what was being studied:

> *The experiment you just participated in is not actually about identifying animals! We're sorry for the deception, but we are actually studying whether the presence of summary statistics about previous responses has an effect on the way people answer our questions. Some people will be shown the summary statistics and others won't, but either way, your contribution to this study is greatly appreciated!*
>
> *In order for us to do the most successful experiment possible, we ask that you please do not share any information about this study with others! Thank you!*

The author of the study judged that such a deception caused minimal harm to participants and did not violate any of the relevant ethical guidelines.

RESULTS

In order to evaluate the treatment effect, the response data first had to be cleansed and interpreted. We sought to measure the number of responses for each participant that were equivalent to "I don't know." To do this, we created a comprehensive list of responses for every animal identification question and extracted every response where the participant indicated that they did not know the answer. Most of these were simple variants of "I don't know" such as typographical errors or abbreviations like "IDK." There were also numerous responses where a silly response was given along with an admission that the participant did not know. For example, one participant responded with "opossum with super sonic hearing (I don't know)." This response and others like it were treated as valid "I don't know" responses because the participant actively and intentionally acknowledged that they did not know. (A full list of the equivalent responses can be found in Appendix C).

Among the responses that were collected (*N=589*), 24 were cases where the first page (the covariate data) was completed, but no animal identification questions were answered. We had no indication as to why these participants attrited out of the study. Because we could not know whether there was some aspect of the treatment that was causing the attrition, we were forced to assume that there was and rely on the calculation of the complier average causal effect (CACE) to estimate the effect size of our treatment.

We also had the additional wrinkle that 42 participants provided at least one animal identification response, but failed to complete the full set of ten. This raised the question of how to score missing responses. One approach might be to treat it as though it were not equivalent to “I don’t know.” However, doing so might bias the results against cases where the participant would have responded with “I don’t know” if only they had completed the survey. To resolve this, we decided to err on the side of caution and produce a highly conservative estimate of the effect by calculating extreme value bounds.

To calculate those bounds, we used both scoring strategies and produced two versions of the outcome variable. Using those outcomes, we produced two separate regression models and two sets of effect estimates. Both models employed 2-stage linear regression to obtain their estimates, and both indicated highly significant CACEs. All of the covariates that we captured were also included in the right-hand sides of all models, but none yielded a significant effect.

Table 1. Results of the two regression models used to evaluate the primary hypothesis. The first model treats all unanswered animal identification questions as though the participant did not answer “I don’t know.” The second model treats blank responses as “I don’t know.”

	Blanks as Non-IDK	**Blanks as IDK**
(Intercept)	7.04*** (0.35)	7.62*** (0.31)
Treatment: Prev. Responses Seen	-3.00*** (0.28)	-2.39*** (0.24)
Male	-0.30 (0.28)	-0.36 (0.25)
Bachelor’s Degree Held	0.37 (0.28)	0.18 (0.25)
Age Over Thirty	-0.02 (0.29)	0.21 (0.26)
Observations	565	565
R^2	0.10	0.18

*** *$p < 0.01$*

Using these models, we computed both 95% confidence intervals, [-3.54, -2.45] and [-2.87, -1.91]. To give us the best possible likelihood of bracketing the true treatment effect, we took the conservative approach of taking the most extreme and least extreme quantiles from these two confidence intervals to produce the extreme value bounds estimate of the CACE: [-3.54, -1.91]. In doing this, we are essentially pushing out one side of the 95% confidence interval for each of the models. This widening of both intervals implies that the resulting interval represents a confidence level greater than 95%, and as such, could probably be narrowed. However, in order to err on the side of caution, we maintain the estimated CACE of [-3.54, -1.91]. This confirms our primary hypothesis, that the presence of summary statistics does indeed cause participants to answer “I don’t know” less frequently.

To evaluate Hypothesis 2, we constructed another regression model to investigate the difference between the moderate and strong treatments. We opted to take the conservative approach again, and used the model that included blank responses as equivalent to “I don’t know” because it produced the less extreme treatment effect estimates in our analysis of the primary hypothesis.

Table 2. Results of the regression model used to explore Hypothesis 2. Moderate and strong treatment effects are estimated separately.

(Intercept)	7.61*** (0.31)
Treatment: Strong	-2.56*** (0.34)
Treatment: Moderate	-2.24*** (0.30)
Male	-0.35 (0.25)
Bachelor’s Degree Held	0.19 (0.25)
Age Over Thirty	0.21 (0.26)
Observations	565
R^2	0.18

The two CACE estimates are not far apart, and their 95% confidence intervals showed substantial overlap. The moderate treatment produced a confidence interval of [-1.67, -2.82], and the strong treatment confidence interval was [-1.96, -3.16]. Given the overlap in the two, we suspected that there was no significant difference between the treatments. A Student’s t-test of the difference in means between the moderate and strong treatments bore this out, *t(265.17) = -0.91, p = 0.37*. Therefore, we fail to reject the null hypothesis that there is no difference between the moderate and strong treatment effects.

All three of the remaining hypotheses that investigated interaction effects of the covariates failed to produce significant results. 2-stage regression models were used in all three cases to look for differential effects of gender, possession of a Bachelor's or higher degree, and age over 30 years among those assigned to treatment. The gender interaction coefficient check produced a p-value of 0.61, the degree interaction coefficient check produced a p-value of 0.26, and the age over 30 interaction coefficient check had a p-value of 0.73. In all cases, we fail to reject the null hypothesis that there was no differential effect on the covariate under treatment.

MANIPULATION CHECK

To verify successful delivery of the treatment, a manipulation check was added as the last question of the survey. Participants were asked: *Were the statistics that you were shown during this study helpful in identifying the various animals?* Three responses were available: *Yes, they were helpful*; *No, they were not helpful*; and *I did not see any statistics*. These responses were condensed into a single dummy variable to indicate whether or not the participant had seen any statistics. That value was tested to check for a difference in means between treatment and control, and yielded a statistically significant result, $t(311.22) = -48.48, p < 0.01$. We therefore reject the null hypothesis that there was no difference in the delivery rate of the treatment between treatment and control, indicating that treatment was successfully delivered.

LIMITATIONS

While the primary hypothesis test did yield a significant result with a practically significant CACE, there were also elements of the experimental design that could have been improved upon. First, the use of the obscure, but real animal images did not properly set the stage for each participant to face our intended choice (guess or answer "I don't know"). The existence of a factually true answer for 8 of the 10 questions had a narrowing effect on the space of possible outcome scores. If the participant knew the actual identity, they would presumably be biased toward providing that answer. Also, if they had once known the animal's identity, they might be inclined to guess based on their memory, regardless of whether it was accurate. In future versions of this experiment, it would be preferable to obtain a complete set of falsified animal images to correct this.

When selecting fake images, it is also important that any real animals used in the image compositing process not be identifiable in the final image. This issue was apparent with the "wombatbat" image used in this experiment. 56 participants named the animal as either wombat or bat, having recognized the wombat body or bat face present in the picture.

Table 3. Positive identification rates for each of the animal images used in the experiment. Responses with misspellings included as correct. Percentages exclude the 24 participants who attrited before providing any responses to animal identification questions.

True Animal Name	# Positive Ident.	% Positive Ident.	Notes
Bongo Antelope	14	2.48%	Including “bongo” and “antelope” as correct responses
Jerboa	14	2.48%	
Bilby	12	2.12%	
“Wombatbat” (fake)	1 (57)	0.18% (10.09%)	46 participants answered “wombat” 10 participants answered “bat”
Axolotl	27	4.78%	
Axolotl	39	6.90%	
“Haggis” (fake)	4	0.71%	
Bear (hairless)	102	18.05%	Including all responses containing “bear”
Frilled Shark	30	5.31%	Including all responses containing “shark”
Nudibranch (or Blue Dragon)	9	1.59%	
Overall	**252 (308)**	**4.46% (5.45%)**	

Participants for this experiment were recruited from sources that had varying levels of risk for communication spillover. For example, Mechanical Turkers frequently share information about available jobs with one another through message boards online. If any of the Turkers took offense at the deception employed in the survey or found the task especially amusing, they might be motivated to discuss it with other Turkers. This might lead to biased results for those who took the survey after being involved in those discussions. This same principle applies for participants who were recruited via social media outlets. We attempted to combat this by including a message on the final debriefing page of the survey:

> *In order for us to do the most successful experiment possible, we ask that you please do not share any information about this study with others!*

We believe that ultimately the risks posed by spillover are low. Even with foreknowledge about the survey, it is unlikely that the participant would know the true purpose of the summary statistics, or answer the questions differently.

Another potential flaw lies in the fact that frustrated participants might have been encouraged toward a higher rate of “I don’t know” responses. Suppose the participant made a genuine effort to answer the first 4 questions, only to find they did not know any of the animals. The resulting frustration might have eroded their willingness to engage fully with the remaining questions, choosing “I don’t know” in an effort to simply finish the task. Given the short length of the survey, we judge this to be a small risk. However, that risk also could have been mitigated by including a few images of well-known animals that were easy to identify throughout the survey.

The most profound challenge to validity in this experiment is the question of what is truly being measured. Ideally, we wanted to operationalize each person’s willingness to admit that they didn’t know the answer. However, the summary statistics themselves may have unintentionally introduced bias to that measurement. Each list of previous responses was constructed to be plausible, but to never contain a truly positive identification for the animal. Unfortunately, for some animals, the responses included values that were “closer” to the real identity than others. For example, the frilled shark does in fact closely resemble an eel. The previous responses for that question included eel as the top choice. By suggesting a highly plausible answer, participants are presumably less likely to approach the question of whether or not they should guess. In those cases, we’ve measured the ability of that particular table of statistics to sway the participant, not their actual willingness to answer “I don’t know.”

A very similar argument could also be made that the statistics might be capable of increasing participant willingness to answer “I don’t know.” If the top previous response was obviously incorrect, it might drive the participant away from even considering the summary statistics at all. Injecting additional uncertainty into their process for answering the question might then lead to a higher rate of “I don’t know” responses.

Solving these problems would require more careful selection of the values in the summary statistics tables. There may never be a perfect balance between answers that are plausible enough to not be ridiculous and answers that are distant enough from providing a direct mental association to a known animal. It might be possible to explore the effect of this phenomenon by restructuring this experiment to replace the moderate/strong treatment levels with varying levels of plausibility among the previous responses. If we measured the treatment effects under nonsense responses, compared to marginally believable responses, compared to highly plausible responses, it might be possible to measure more precisely how much weight these concerns actually carry.

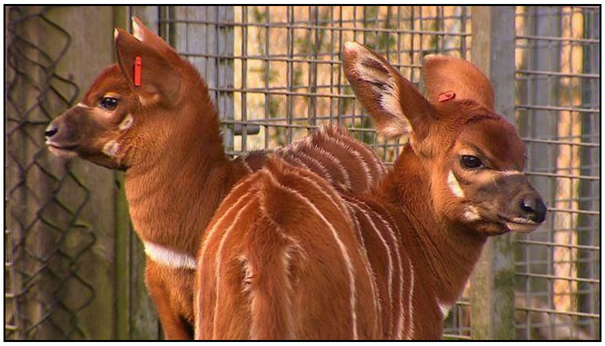
CONCLUSIONS

This field experiment tested whether the presence of summary statistics makes people less likely to admit when they don’t know the answer to a question. We found that they do indeed have a general dampening effect on people’s willingness to answer “I don’t know.” There was no difference in the magnitude of this effect when the summary statistics included a very confident top answer as opposed to a moderately confident top answer. And no differential effects under treatment were found based on gender, possession of a Bachelor's (or higher) degree, or age above or below 30.

Flaws in the survey design do call into question whether the magnitude of the measured effect is reliable. The use of obscure but real animal images as opposed to completely false images, as well as the substance of the summary statistics may have led to unintentional bias in participants' responses, potentially biasing outcomes in both directions. However, we do not believe that such bias would completely negate the measured effect. Many of the difficulties encountered during this experiment would be relatively easy to resolve in another iteration of the experiment, and we hope that these results might provide motivation for further study.

Appendix A - Animal Images and Previous Response Tables

These animal images were used in the final study. True names of the animals are shown here, but were never revealed to participants. Fake animals are noted. Previous response tables are also shown here as they appeared to participants.



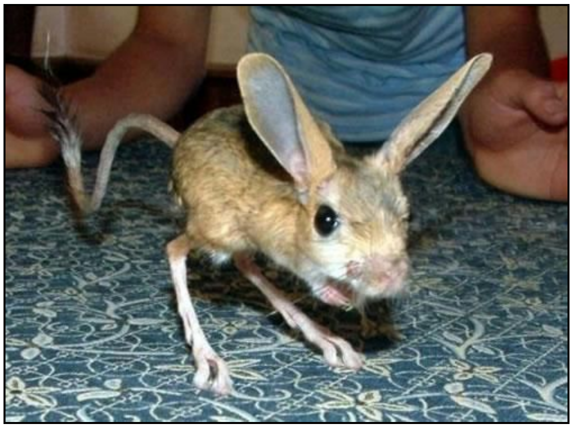
Bongo Antelope

Previous Responses (Moderate)

- Deer (61%)
- Alpaca (30%)
- Zebra (5%)
- Other (4%)

Previous Responses (Strong)

- Deer (71%)
- Alpaca (20%)
- Zebra (5%)
- Other (4%)



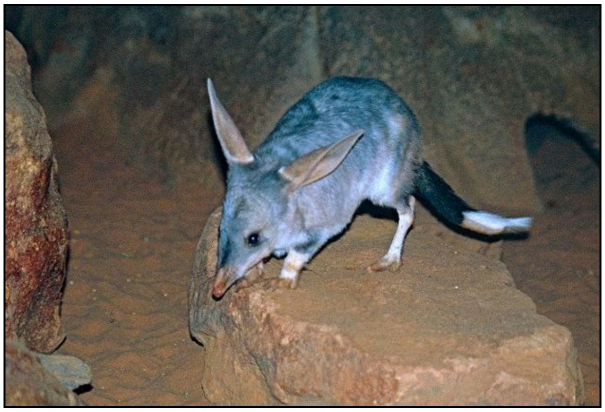
Jerboa

Previous Responses (Moderate)

- Mouse (55%)
- Rabbit (29%)
- Rat (13%)
- Other (3%)

Previous Responses (Strong)

- Mouse (80%)
- Rabbit (9%)
- Rat (8%)
- Other (3%)



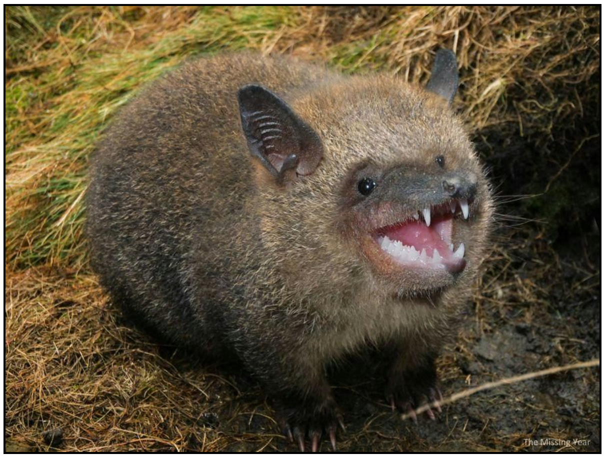
Bilby

Previous Responses (Moderate)

- Possum (58%)
- Rat (37%)
- Rabbit (3%)
- Other (2%)

Previous Responses (Strong)

- Possum (74%)
- Rat (21%)
- Rabbit (3%)
- Other (2%)



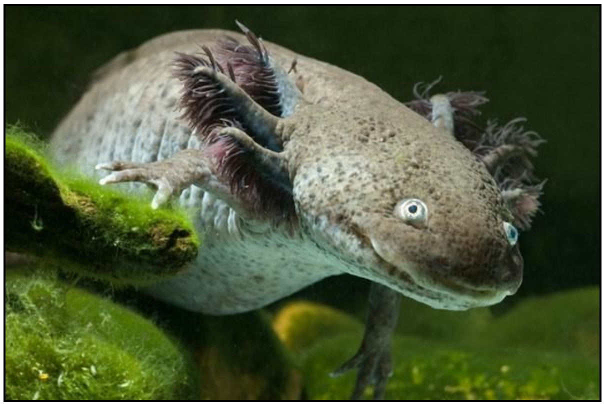
"Wombatbat" (fake)

Previous Responses (Moderate)

- Wolverine (61%)
- Porcupine (26%)
- Badger (10%)
- Other (3%)

Previous Responses (Strong)

- Wolverine (75%)
- Porcupine (12%)
- Badger (10%)
- Other (3%)



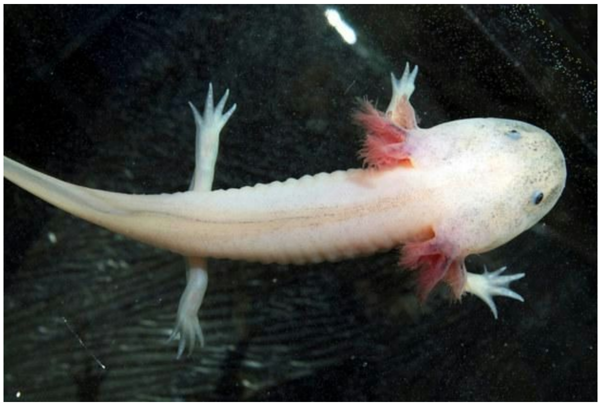
Axolotl (adult)

Previous Responses (Moderate)

- Otter (56%)
- Manatee (28%)
- Fish (12%)
- Other (4%)

Previous Responses (Strong)

- Otter (74%)
- Manatee (18%)
- Fish (6%)
- Other (4%)



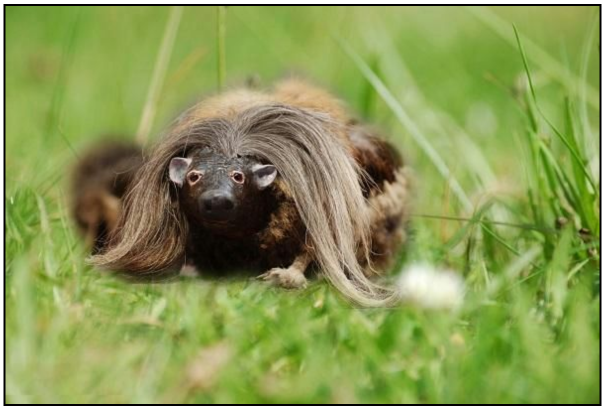
Axolotl (young)

Previous Responses (Moderate)

- Lizard (62%)
- Tadpole (24%)
- Fish (12%)
- Other (2%)

Previous Responses (Strong)

- Lizard (72%)
- Tadpole (15%)
- Fish (12%)
- Other (1%)



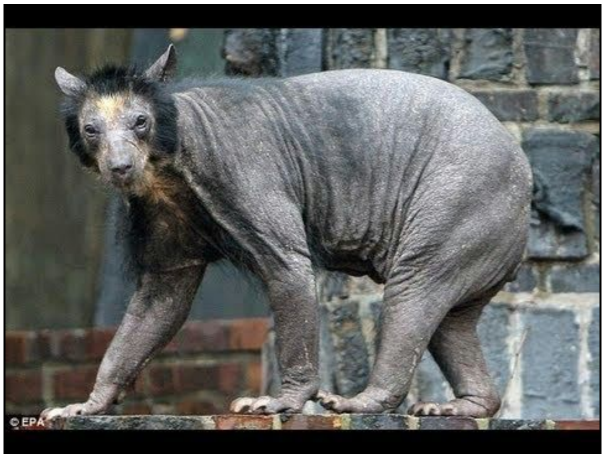
“Haggis” (fake)

Previous Responses (Moderate)

- Badger (58%)
- Beaver (20%)
- Aardvark (19%)
- Other (3%)

Previous Responses (Strong)

- Badger (78%)
- Beaver (10%)
- Aardvark (9%)
- Other (3%)



© EPA

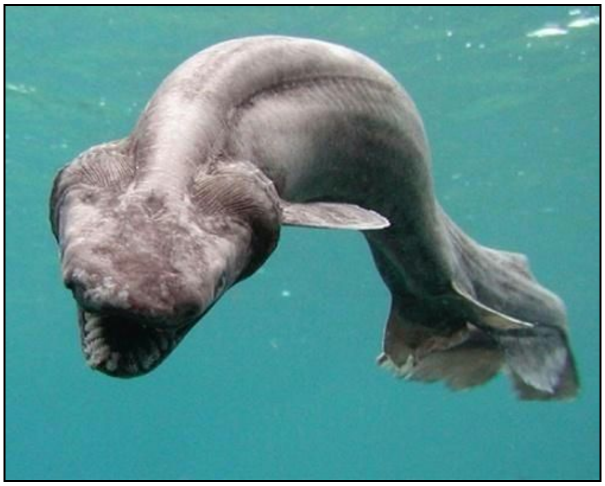
Bear (hairless)

Previous Responses (Moderate)

- Sloth (57%)
- Elephant (16%)
- Baboon (14%)
- Other (13%)

Previous Responses (Strong)

- Sloth (73%)
- Elephant (21%)
- Baboon (3%)
- Other (3%)



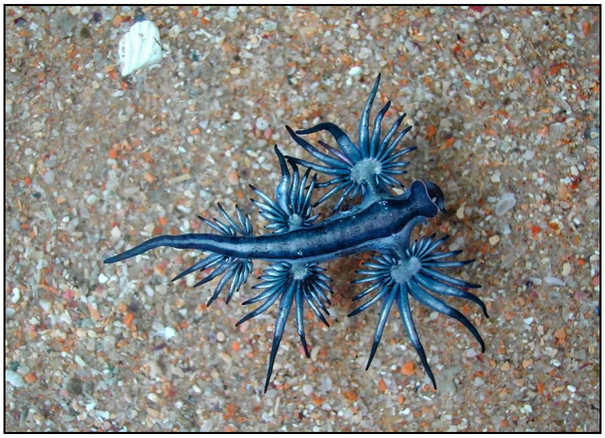
Frilled Shark

Previous Responses (Moderate)

- Eel (59%)
- Whale (22%)
- Fish (17%)
- Other (2%)

Previous Responses (Strong)

- Eel (79%)
- Whale (12%)
- Fish (7%)
- Other (2%)



Nudibranch (a.k.a. Blue Dragon)

Previous Responses (Moderate)

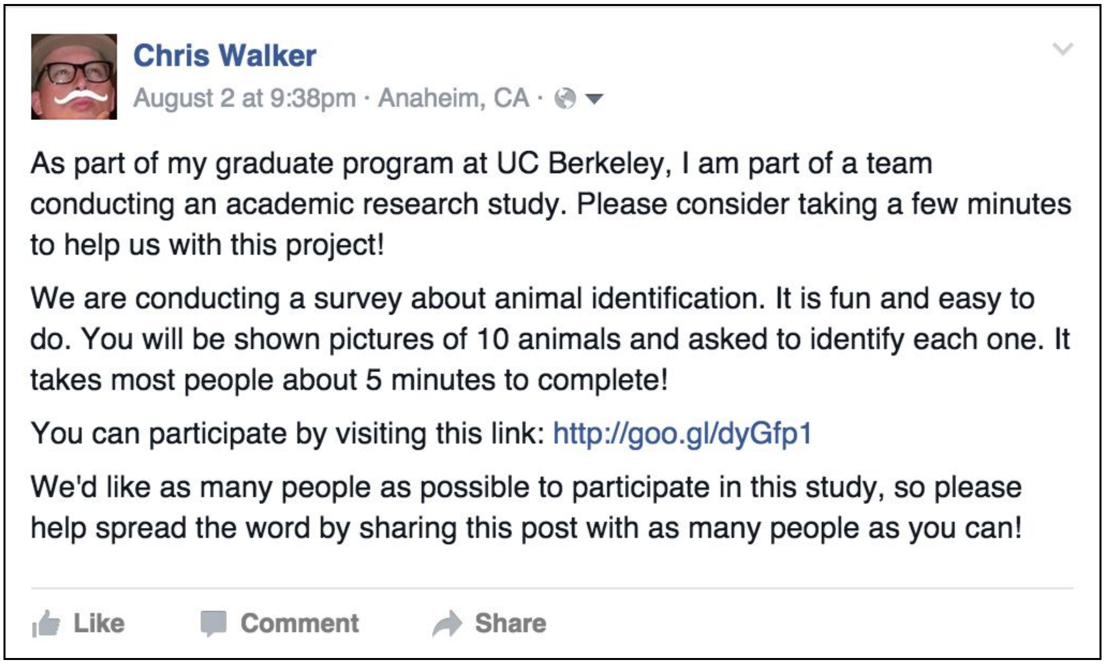
- Fish (57%)
- Coral (22%)
- Crab (18%)
- Other (3%)

Previous Responses (Strong)

- Fish (73%)
- Coral (15%)
- Crab (8%)
- Other (4%)

Appendix B - Recruitment Messages

Facebook:



Chris Walker

August 2 at 9:38pm · Anaheim, CA ·

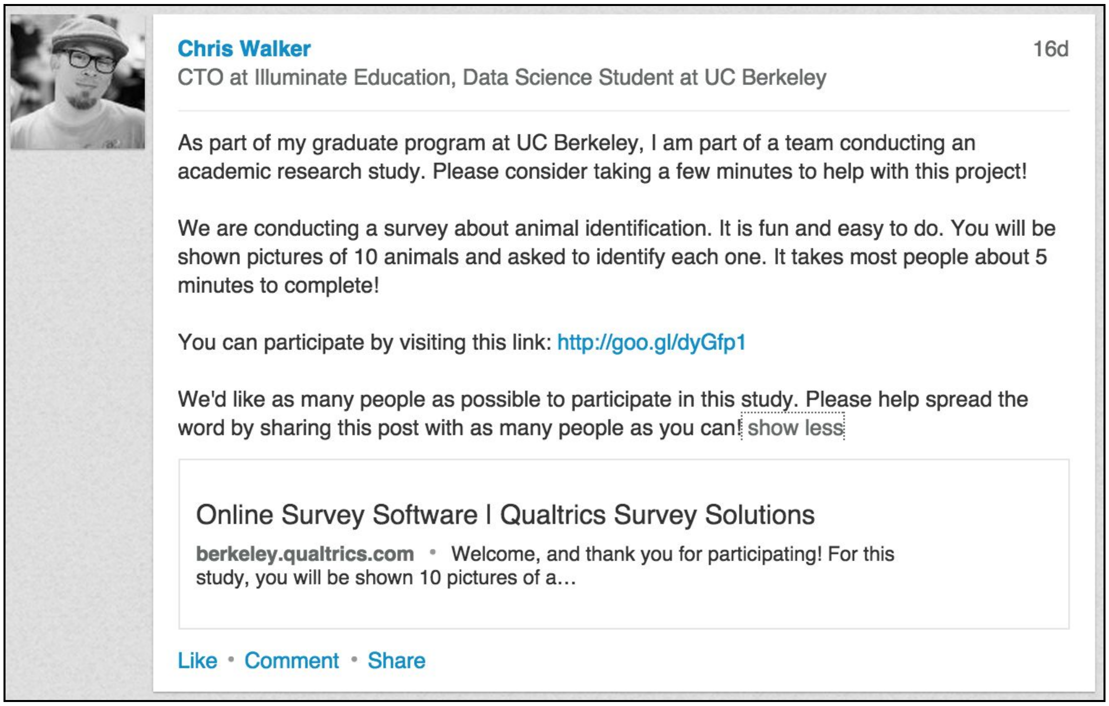
As part of my graduate program at UC Berkeley, I am part of a team conducting an academic research study. Please consider taking a few minutes to help us with this project!

We are conducting a survey about animal identification. It is fun and easy to do. You will be shown pictures of 10 animals and asked to identify each one. It takes most people about 5 minutes to complete!

You can participate by visiting this link: http://goo.gl/dyGfp1

We'd like as many people as possible to participate in this study, so please help spread the word by sharing this post with as many people as you can!

LinkedIn:



Chris Walker

16d

CTO at Illuminate Education, Data Science Student at UC Berkeley

As part of my graduate program at UC Berkeley, I am part of a team conducting an academic research study. Please consider taking a few minutes to help with this project!

We are conducting a survey about animal identification. It is fun and easy to do. You will be shown pictures of 10 animals and asked to identify each one. It takes most people about 5 minutes to complete!

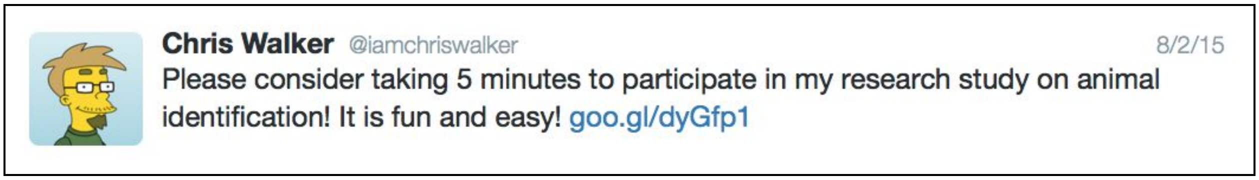
You can participate by visiting this link: http://goo.gl/dyGfp1

We'd like as many people as possible to participate in this study. Please help spread the word by sharing this post with as many people as you can! show less

Online Survey Software | Qualtrics Survey Solutions

berkeley.qualtrics.com • Welcome, and thank you for participating! For this study, you will be shown 10 pictures of a…

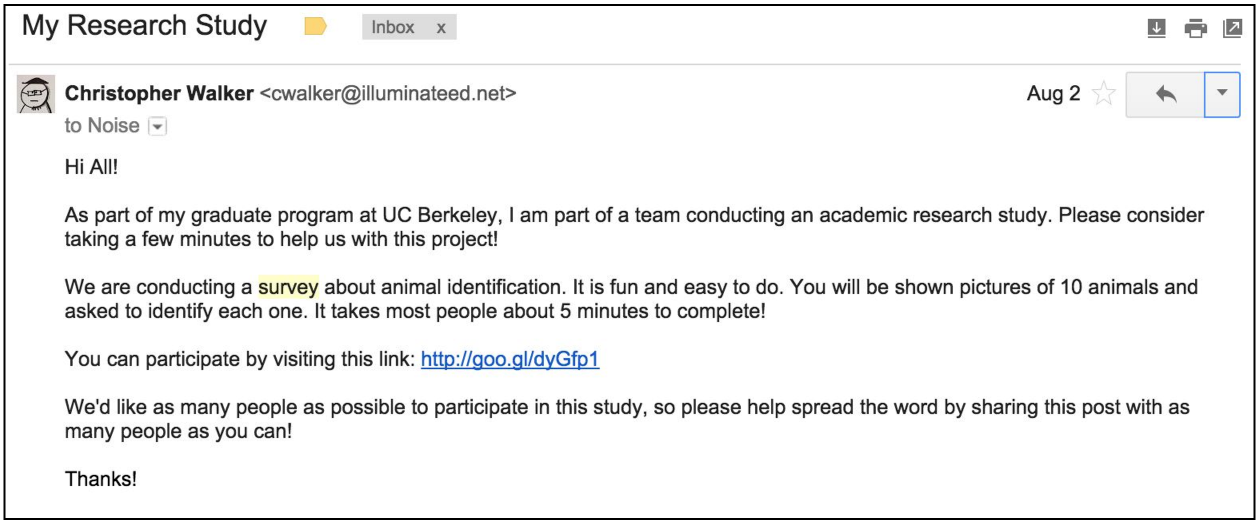
Twitter:



Chris Walker @iamchriswalker 8/2/15

Please consider taking 5 minutes to participate in my research study on animal identification! It is fun and easy! goo.gl/dyGfp1

Email 1:



My Research Study

Christopher Walker <cwalker@illuminateed.net> Aug 2

to Noise

Hi All!

As part of my graduate program at UC Berkeley, I am part of a team conducting an academic research study. Please consider taking a few minutes to help us with this project!

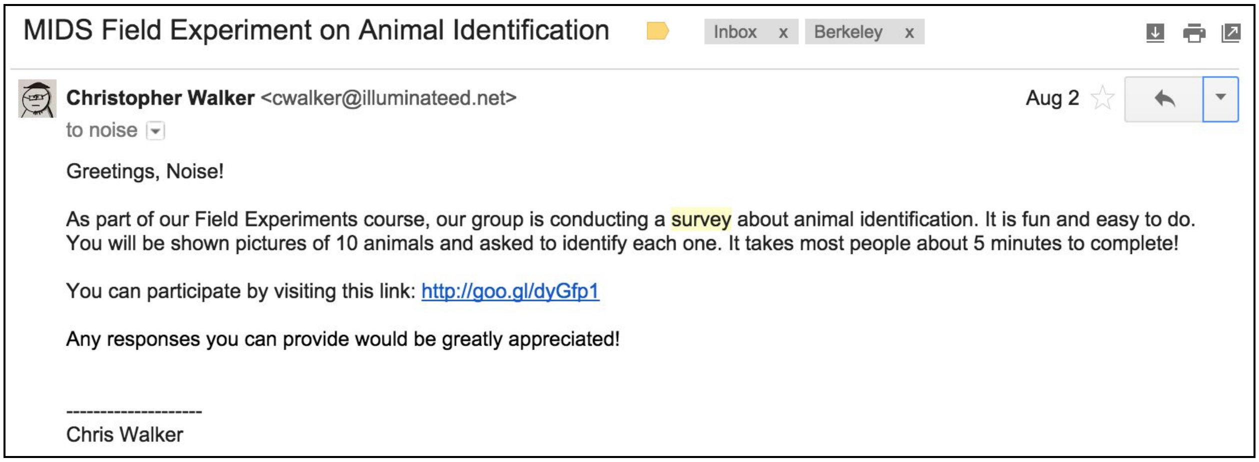
We are conducting a survey about animal identification. It is fun and easy to do. You will be shown pictures of 10 animals and asked to identify each one. It takes most people about 5 minutes to complete!

You can participate by visiting this link: http://goo.gl/dyGfp1

We'd like as many people as possible to participate in this study, so please help spread the word by sharing this post with as many people as you can!

Thanks!

Email 2:



MIDS Field Experiment on Animal Identification

Christopher Walker <cwalker@illuminateed.net> Aug 2

to noise

Greetings, Noise!

As part of our Field Experiments course, our group is conducting a survey about animal identification. It is fun and easy to do. You will be shown pictures of 10 animals and asked to identify each one. It takes most people about 5 minutes to complete!

You can participate by visiting this link: http://goo.gl/dyGfp1

Any responses you can provide would be greatly appreciated!

Chris Walker

Appendix C - Responses Treated as Equivalent to “I Don’t Know”

I don't know.	I don't know!	I really don't know
I DON'T KNOW	I don't know.	I. Don't know
I don't know	I don't know.	i.dont know
80's hair metal Andre Braugher (I don't know)	I don't know. Looks like a wingless bat!	I've seen this one before... but don't know its name
I don't know	I don't know. None of there are real.	albino medusa newt (I don't know)
I do not know	i don't know...oh come on!	Albinodontknow
I don;t know	i don't know...this shit is hard	don't know
I don;t know.	I don't konw	Don't know
I don't k ow	I don't kow	Don't know
I don't kmow	I don't ky	dont know
I don't knkw	I don't lnow	Dont know
I don't knoe	I don't[know	elephagoat (I don't know)
i don't know	i don'tknow	glamour pig (I don't know)
I don't know	I don'tknow	idk
I don't Know	I don'tknow.	Idk
I don't KNOW	i dont know	IDK
I Don't know	I dont know	Idon't know
I DON'T KNOW	I dont Know	kangaroo mouse? (I don't know)
i don't know	i dont know	medusa salamander (I don't know)
I don't know	I dont know	opossum with super sonic hearing (I don't know)
I don't Know	i dont' know	Same as the other one I don't know
I don't know - looks like 'nessie' / cgi	I dont' know.	teddy bear with fangs (I don't know)
i don't know :(	I dont't know	What! I don't know
I don't know :(	I dontknow	I don't know, but it's cute
I odn't know		