

**Corrigendum for
“Crowdsourcing Annotations for Websites’ Privacy Policies: Can It Really Work?”
As Published in the Proceedings of WWW 2016**

Summary

The authors discovered errors in parts of the data analysis reported in this paper, as published at WWW 2016. The following are corrections to the reported results:

- The annotation accuracy by crowdworkers is slightly higher than previously reported (ranging from 92-98% instead of 87-98%).
- Due to an error in how agreement was calculated, more policy-question pairs should have been considered in the crowdworker/skilled annotator comparison. The correct accuracy numbers above are based on the corrected number of policy-question pairs, constituting higher accuracy over more data.
- The corrections result in slight changes regarding what questions were difficult to answer reliably by crowdworkers. Sharing of financial and contact information questions are most difficult to answer, not sharing of contact information.
- For the highlighting study, it was erroneously reported that reported legal training was lower in the NOHIGH condition. This is no longer the case. Instead, levels of legal training are approximately the same across all conditions.
- It was reported that highlighting of paragraphs “positively affected perceived task difficulty.” Instead the claim shall be revised to: “We found that highlighting relevant paragraphs can reduce task completion time without impacting annotation accuracy.”

Except as noted above, the conclusions of the paper remain the same.

Detailed corrections to the text:

The most significant changes to the paper’s claims are in bold below.

1 Introduction:

- In paragraph “First, we investigate...”: annotation accuracy was 95%; should be 96%

4.1 Overall Accuracy:

- Last sentence in the first paragraph was: “All crowdworker agreement thresholds demonstrate strong accuracy when evaluated against skilled annotators’ answers, with accuracies ranging from 87% (i.e., 132/151 at the 70% crowdworker agreement threshold) up to 98% (42/43 at the 100% crowdworker agreement threshold)”; the numbers should be “with accuracies ranging from 92% (i.e., 168/182 at the 60% crowdworker agreement threshold) up to 98% (47/48 at the 100% crowdworker agreement threshold)”

- Caption for Figure 2: number of policy-question pairs was 179; should be 194

4.2 Question-Specific Results

- Most difficult to answer question for crowdworkers was stated to be sharing of contact information (23 no converge); should be a tie between sharing financial information and sharing location information (both 20 no coverage)
- Most difficult to answer question for skilled annotators remains sharing of financial information, but frequency was stated to be 19; should be 15

5 Highlighting paragraphs

- **First sentence said: “crowdworkers can provide highly accurate privacy policy annotations for some questions”; however, the claim should be stronger, that they can do this “for most questions”**

4.2: Study: Effects of Highlighting

- **Sentence in first paragraph said: “We found that highlighting relevant paragraphs can reduce task completion time and positively affect perceived task difficulty without impacting annotation accuracy”; the phrase “and positively affect perceived task difficulty” should be removed from the claim**

4.2.2 Results

- Changes throughout the section to account for the correction to Figure 5: levels of legal training are approximately the same across the control and the conditions
- Change to caption for Figure 6: the sentence “Control group participants (NOHIGH) rated their ability substantially lower...” should be removed
- Changes to the description of Figure 7 to match accuracy percentages in the new figure

4.3 Discussion

- Final paragraph: “improve the user experience” should be “may improve the user experience”

WWW 2016 Corrections to Figures and Tables

The figures and tables below are corrections to those in the published manuscript.

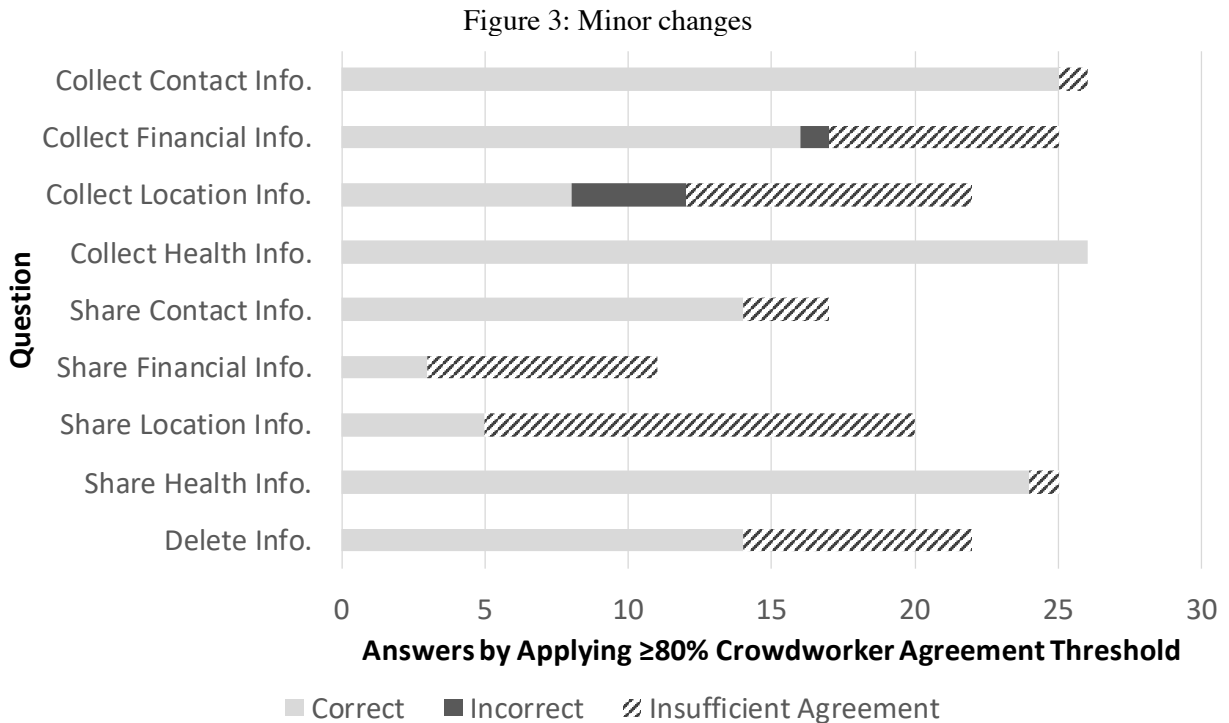
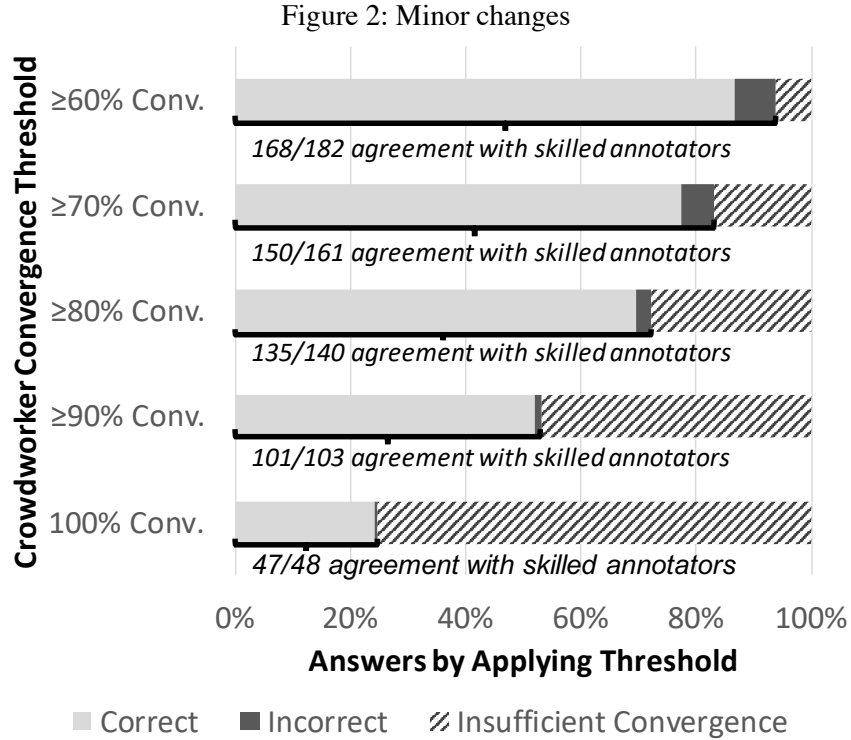


Table 2: Several changes to numbers

Question	Skilled Annotators			Crowdworkers		
	Yes	Unclear or N/A	No Conv.	Yes	Unclear or N/A	No Conv.
Collect Contact Info.	26			25		1
Collect Financial Info.	21	4	1	13	4	9
Collect Location Info.	10	12	4	14		12
Collect Health Info.	1	25		1	25	
Share Contact Info.	17		9	22		4
Share Financial Info.	7	4	15		6	20
Share Location Info.	1	19	6	2	4	20
Share Health Info.		25	1		24	2
Deletion of Info.	9	13	4	7	8	11
Total	92	102	40	84	71	79

Table 3: Several changes to numbers

	Gender			Age		Education	
	Male	Female	Undisclosed	Range	Median	No College Degree	College Degree
NOHIGH	50.2%	49.4%	0.4%	18-82	32.5	48.7%	51.3%
TOP10	58.0%	42.0%	0%	19-68	30.9	42.9%	57.1%
TOP05	58.3%	41.0%	0.8%	20-65	31.4	45.8%	54.2%

Figure 5: The distribution of responses from the NOHIGH condition is now approximately in line with the other conditions

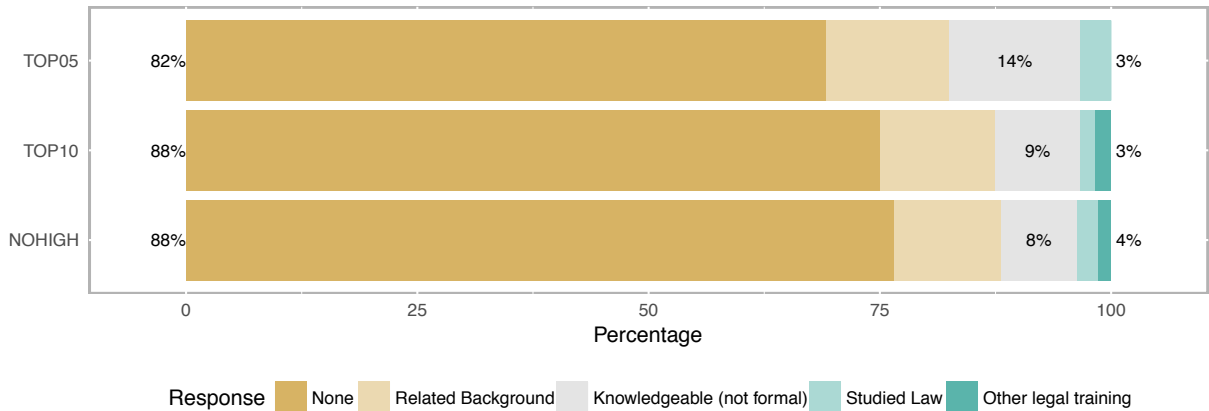


Figure 6: The distribution of responses from the NOHIGH condition is now approximately in line with the other conditions

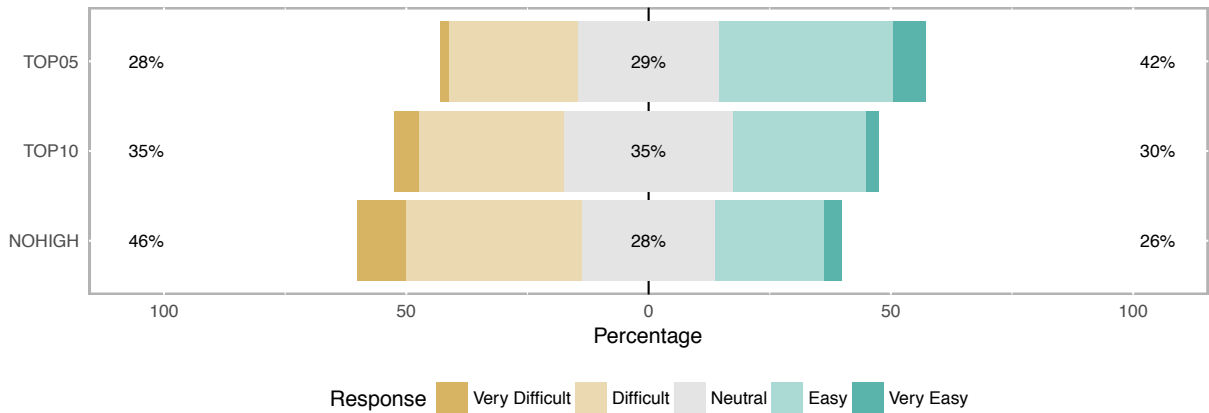


Figure 7: Minor changes

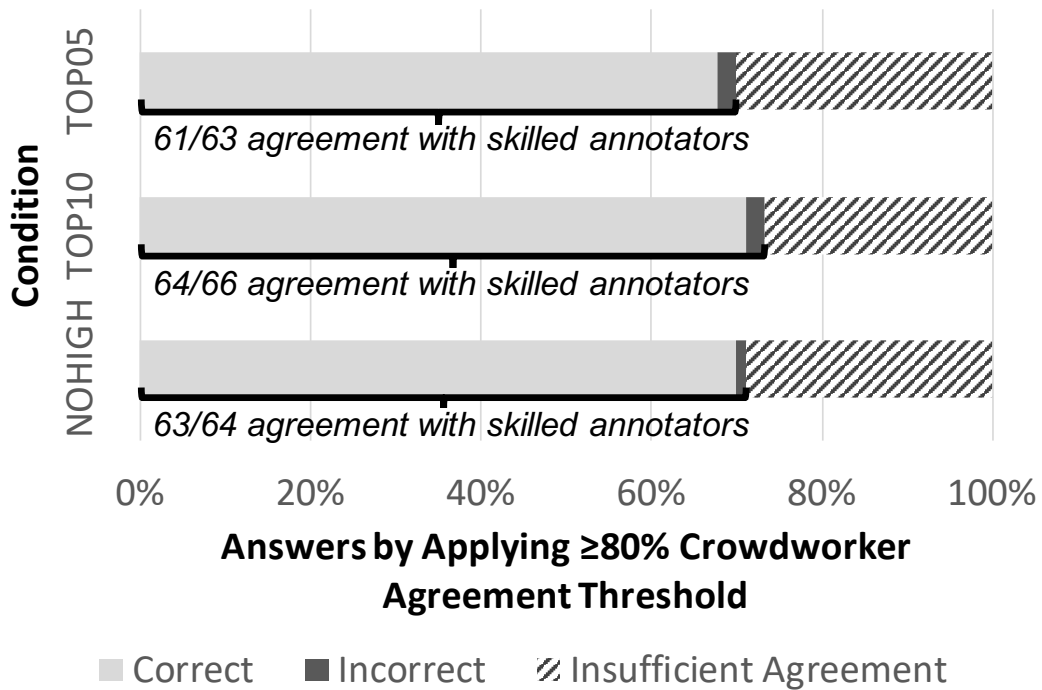


Figure 9: Minor changes, and a new layout for interpretability

