A Supplementary material

A.1 Crowdsourcing details

Figure 2 shows the instructions that crowdworkers were given at the beginning of each question writing task. Figure 3 shows the interface they used to write unanswerable questions for each paragraph. In the interface, workers first write an unanswerable question, then highlight a plausible answer in the paragraph.

A.2 Plausible answers as distractors

As mentioned in Section 5.4, we measured how often systems were fooled into answering the plausible answers provided by crowdworkers for our unanswerable questions. For each system, we first isolated their false positive errors—cases where they predicted an answer to an unanswerable question—on the development set. Within this set of examples, we measured exact match and F1 scores between the system predictions and plausible answers. These numbers are shown in Table 5. Plausible answers account for roughly half of the false positive errors made by each of the computer systems, as well as by human answerers. We conclude that the plausible answers in our dataset do indeed serve their purpose of being distracting spans that could be mistaken for the correct answer.

<table>
<thead>
<tr>
<th>System</th>
<th>EM</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>BNA</td>
<td>48.6</td>
<td>63.0</td>
</tr>
<tr>
<td>DocQA</td>
<td>55.0</td>
<td>69.9</td>
</tr>
<tr>
<td>DocQA +ELMo</td>
<td>54.9</td>
<td>69.2</td>
</tr>
<tr>
<td>Human</td>
<td>46.4</td>
<td>60.6</td>
</tr>
</tbody>
</table>

Table 5: Exact match (EM) and F1 scores between system predictions and plausible answers, in cases where the system made a false positive error.
Ask Impossible Reading Comprehension Questions

Instructions
In this article about Geology, you will be asked to pose and answer reading comprehension questions based on the paragraph. There is a twist! The question you pose must be impossible to answer based on the paragraph alone, but should be about the same topic and same people/places/things! Additionally, the paragraph must contain a phrase/word that seems like a plausible answer to the question. Read each paragraph, pose an impossible question, and then highlight a phrase from the paragraph that looks like a plausible (but of course, incorrect) answer. We'll clarify the task with the help of an example below.

Estimated Time For Task Completion - 3.2 hours
This article consists of 25 paragraphs. We recommend a time of 7 minutes per paragraph. Submit each paragraph after you are done to save partial progress. Feel free to take breaks -- if you come back to the task, you do not need to resubmit paragraphs already submitted in an earlier session. After completing all paragraphs, click the submit task button at the end of the page.

Task Tutorial
1. On the left, you’ll see a reading passage and ‘prompt questions’ underneath it. First read the passage, and skim over the questions.
2. Now, based on what you’ve read in the passage, your task is to come up with questions that don’t have a correct answer in the passage, but have feasible answers.
3. Start by picking a question from the prompt questions. For the purposes of this example, let’s pick What is decentralization according to the United Nations Development Programme?
4. Note that this question does have an answer in the passage. We’re going to modify it so that it doesn’t have an answer. For instance, we can modify the question to What is decentralization according to the local government? Note that this question doesn’t have an answer in the passage, but still contains entities present in the passage such as local government.
5. You will also be asked to pick a plausible answer for our question. This is an answer that looks possibly correct if someone hadn’t read the passage. You select the plausible answer by highlighting a segment of the passage. For our example question, we would highlight It is “more than a process, it is a way of life and a state of mind.”
6. Let’s come up with another example. This time, we will use the inspiration question What is a way of life and state of mind, more than a

Figure 2: The instructions shown to crowdworkers at the beginning of each question writing task.
Figure 3: The interface crowdworkers used to write unanswerable questions and annotate plausible answers.