# Evaluating Explanations of a Computer Program

## Instructions

Thank you for agreeing to participate in this study.

Please take part in this study only if you are willing to consider the tasks seriously and take some time to work out your answer. Participation in the study will take up to 12 minutes and you will be paid a sum of £7.50/hr in recognition of your participation.

The study can ONLY be completed on a desktop or laptop computer - the information will NOT display correctly on a smartphone or similar device. Please do NOT attempt to take part in the study on a smartphone, as you will not be able to complete the study.

**The main task in this study is to evaluate a computer program that can label objects in a picture.**

Please do not take part in the study if you have previously taken part in any study that has asked you to do this same task.

Your participation is voluntary and you are free to withdraw during any stage of the study without penalty, up to the point when the data is pooled for analysis. Should you wish to withdraw, any data that has already been collected will be destroyed.
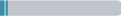
Annonymized

Next

0%

# Evaluating Explanations of a Computer Program
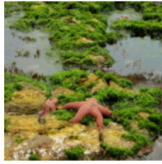
Please enter your Prolific ID:

Back    Next

2%

## Evaluating Explanations of a Computer Program

**The main task in this study is to evaluate a computer program that can label objects in a picture.**

**You will also be asked to evaluate explanations of the program's labelling.**

**You will be shown two different ways to explain the program's labelling.**

Original



For example, when the program is shown the above picture, it labels it as a: **Starfish**.

The program has learned to do this, by being shown many pictures with their correct labels, provided by human labellers. Then, it uses what it has learned to label new pictures that it has not been shown before.
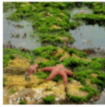
Back Next

4%

Instructions: Task 1

*In this study, on each page, you are asked to do **four** simple tasks to help us evaluate the computer program.*

*In the first task, you will be shown a picture and the label given by the program for this picture:*



*Here, your task is to carefully note the label that the program has given for the object in the picture.*

Back    Next

6%

Instructions: Task 2

*Then, on the same page, you are asked to do a **second** task.*

*In this task, you will be shown **one of two possible explanations** for why the program labelled the picture this way (in Task 1).*

**First Possible Explanation**



Task 2:
Explanation: The program labelled the object in the picture (Original) this way because of what it learned from this other picture (Explanation) that people labelled as: **Starfish**.

Original          Explanation

*Here your task is to carefully note the original and explanation pictures.*

*You will be show an explanation using the above pictures with no highlighting **or** an explanation using the below pictures with highlighting.*

*The **second possible explanation** highlights the **main part** of the picture that was used by the program when it labelled the object with a **yellow box**.*

**Second Possible Explanation**



Task 2:
Explanation: The program labelled the object in the picture (Original) this way because of what it learned from this other picture (Explanation) that people labelled as: **Starfish**.

Original          Explanation

See Yellow Box Highlighting

*Again, with these **yellow box highlighted pictures** your task is to carefully note the original and explanation pictures.*

Back    Next

8%

# Evaluating Explanations of a Computer Program

Instructions: Task 3

*On the same page, you are asked to do a third task.*

*In this task, you are asked to **rate the correctness** of the label provided by the program for the picture (in Task 1).*

**Task 3:**
The program's labelling of the image in Task 1 is **correct**:

| 1<br>I disagree strongly | 2<br>I disagree somewhat | 3<br>I'm neutral about it | 4<br>I agree somewhat | 5<br>I agree strongly |
|:---:|:---:|:---:|:---:|:---:|
| ○ | ○ | ○ | ○ | ○ |

*Here, you should **assess the correctness** of the **program's labelling** by picking one of the numbers on the rating scale.*

Back     Next

10%

## Evaluating Explanations of a Computer Program

Instructions: Task 4

*Finally, on the same page you are asked to do a fourth task.*

*In this task you will be asked to **rate the explanation** (in Task 2).*

*That is, to evaluate whether the explanation helps you to **understand** the program's labelling*

**Task 4:**
The explanation in Task 2 helps me **understand** the program's labelling of the object in the picture: •

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|
| I disagree strongly | I disagree somewhat | I'm neutral about it | I agree somewhat | I agree strongly |
| ○ | ○ | ○ | ○ | ○ |

*Here, you should assess whether the explanation helps you **understand** the program's labelling, by picking one of the numbers on the rating scale.*

Back    Next

12%

Instructions

# Now, practice the tasks using the examples shown on the next two pages…

14%

# Evaluating Explanations of a Computer Program

Practice: Your tasks will look like this

**Task 1:**
The program was presented with this picture:

Original



The program labelled this picture as: **Starfish**

**Task 2:**
Explanation: The program labelled the object in the picture (Original) this way because of what it learned from this other picture (Explanation) that people labelled as: **Starfish**.

Original          Explanation

    

**Task 3:**
The program's labelling of the image in Task 1 is _**correct**_:

| 1 I disagree strongly | 2 I disagree somewhat | 3 I'm neutral about it | 4 I agree somewhat | 5 I agree strongly |
|:---:|:---:|:---:|:---:|:---:|
| ○ | ○ | ○ | ○ | ○ |

**Task 4:**
The explanation in Task 2 helps me **understand** the program's labelling of the object in the picture:

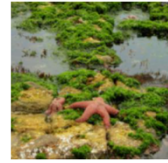| 1 I disagree strongly | 2 I disagree somewhat | 3 I'm neutral about it | 4 I agree somewhat | 5 I agree strongly |
|:---:|:---:|:---:|:---:|:---:|
| ○ | ○ | ○ | ○ | ○ |

Back    Next

16%

# Evaluating Explanations of a Computer Program

Practice: Your tasks will look like this

### *Task 1:*
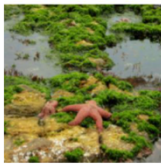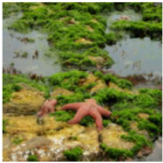The program was presented with this picture:

Original



The program labelled this picture as: **Starfish**

### *Task 2:*
Explanation: The program labelled the object in the picture (Original) this way because of what it learned from this other picture (Explanation) that people labelled as: **Starfish**.

Original          Explanation



### *Task 3:*
The program's labelling of the image in Task 1 is ***correct***:

| 1<br>I disagree strongly | 2<br>I disagree somewhat | 3<br>I'm neutral about it | 4<br>I agree somewhat | 5<br>I agree strongly |
|:---:|:---:|:---:|:---:|:---:|
| ○ | ○ | ○ | ○ | ○ |

### *Task 4:*
The explanation in Task 2 helps me **understand** the program's labelling of the object in the picture:

| 1<br>I disagree strongly | 2<br>I disagree somewhat | 3<br>I'm neutral about it | 4<br>I agree somewhat | 5<br>I agree strongly |
|:---:|:---:|:---:|:---:|:---:|
| ○ | ○ | ○ | ○ | ○ |

Back   Next

18%

Instructions

So, if you feel you now understand what you have to do, please continue with the study and progress through the pages, in order, at your own pace.

Back  Next

20%

Evaluating Explanations of a Computer Program

Instructions

So, if you feel you now understand what you have to do, please continue with the study and progress through the pages, in order, at your own pace.

Instructions

# Please only take part in this study if you are willing to consider the tasks and your answers to them carefully.

Back  Next

22%

Instructions

# Press Next Page to Begin the Study.

Back   Next

24%

Question 19 - AB

***Task 1:***

The program was presented with this picture:

Original



The program labelled this picture as: **Flute**

***Task 2:***

Explanation: The program labelled the object in the picture (Original) this way because of what it learned from this other picture (Explanation) that people labelled as: **Flute**.

Original          Explanation



***Task 3:***

The program's labelling of the image in Task 1 is ***correct***:

| 1<br>I disagree strongly | 2<br>I disagree somewhat | 3<br>I'm neutral about it | 4<br>I agree somewhat | 5<br>I agree strongly |
|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ |

***Task 4:***

The explanation in Task 2 helps me **understand** the program's labelling of the object in the picture:

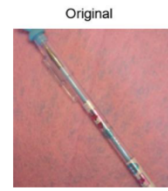| 1<br>I disagree strongly | 2<br>I disagree somewhat | 3<br>I'm neutral about it | 4<br>I agree somewhat | 5<br>I agree strongly |
|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ |

Back    Next

27%

# Evaluating Explanations of a Computer Program

***Task 1:***
The program was presented with this picture:

Original



The program labelled this picture as: **Electric Guitar**

***Task 2:***
Explanation: The program labelled the object in the picture (Original) this way because of what it learned from this other picture (Explanation) that people labelled as: **Electric Guitar**.

Original          Explanation



***Task 3:***
The program's labelling of the image in Task 1 is ***correct***:

| 1<br>I disagree strongly | 2<br>I disagree somewhat | 3<br>I'm neutral about it | 4<br>I agree somewhat | 5<br>I agree strongly |
|---|---|---|---|---|
| ○ | ○ | ◉ | ○ | ○ |

***Task 4:***
The explanation in Task 2 helps me **understand** the program's labelling of the object in the picture:

| 1<br>I disagree strongly | 2<br>I disagree somewhat | 3<br>I'm neutral about it | 4<br>I agree somewhat | 5<br>I agree strongly |
|---|---|---|---|---|
| ○ | ○ | ◉ | ○ | ○ |

Back    Next

33%

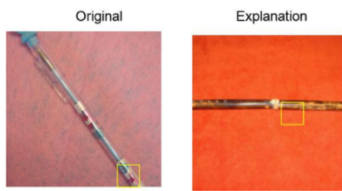## Evaluating Explanations of a Computer Program

**Task 1:**
The program was presented with this picture:

Original



The program labelled this picture as: **Shovel**

**Task 2:**
Explanation: The program labelled the object in the picture (Original) this way because of what it learned from this other picture (Explanation) that people labelled as: **Shovel**.

Original          Explanation



**Task 3:**
The program's labelling of the image in Task 1 is ***correct***:

| 1 I disagree strongly | 2 I disagree somewhat | 3 I'm neutral about it | 4 I agree somewhat | 5 I agree strongly |
|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ |

**Task 4:**
The explanation in Task 2 helps me **understand** the program's labelling of the object in the picture: ·

| 1 I disagree strongly | 2 I disagree somewhat | 3 I'm neutral about it | 4 I agree somewhat | 5 I agree strongly |
|---|---|---|---|---|
| ○ | ○ | ○ | ○ | ○ |

Back    Next

31%

# Evaluating Explanations of a Computer Program

## *Task 1:*
The program was presented with this picture:

Original



The program labelled this picture as: **Snail**

## *Task 2:*
Explanation: The program labelled the object in the picture (Original) this way because of what it learned from this other picture (Explanation) that people labelled as: **Snail**.

Original                Explanation



## *Task 3:*
The program's labelling of the image in Task 1 is **_correct_**:
·

| 1<br>I disagree strongly | 2<br>I disagree somewhat | 3<br>I'm neutral about it | 4<br>I agree somewhat | 5<br>I agree strongly |
|:---:|:---:|:---:|:---:|:---:|
| ○ | ○ | ○ | ○ | ○ |

## *Task 4:*
The explanation in Task 2 helps me **understand** the program's labelling of the object in the picture: ·

| 1<br>I disagree strongly | 2<br>I disagree somewhat | 3<br>I'm neutral about it | 4<br>I agree somewhat | 5<br>I agree strongly |
|:---:|:---:|:---:|:---:|:---:|
| ○ | ○ | ○ | ○ | ○ |

Back    Next

33%

Continue for 32 Questions…. See main paper for study layout.

All original materials may be sampled from ImageNet in the supplement.

Recall there are two versions of the study, one were material Set-A has a CCR explanation, and Set-B does not, and then the other study which does the opposite. All incorrect classifications are randomized.

Correctly classified (e.g., the snail above) materials were spaced out every 4 questions and were not randomized (i.e., question 4, 8, 12… 32)

End

# Congratulations!

## You have completed the study.

Back  Submit

98%

# Debrief Page

This study is being conducted to determine the effect of the use of explanations for the outputs of computer programs. Your responses will be used to compare the perceived usefulness of computer program explanations such as this used here.

<div style="background-color:red; color:white; text-align:center;">Annonymized</div>

Thank you for taking our survey. Your response is very important to us.

If you have any further questions about the study, or would like to access a summary of the findings, please contact:

<div style="background-color:red; color:white; text-align:center;">Annonymized</div>

**You must enter this code on the Prolific site.**

**If after 30 seconds you are not automatically redirected to the Prolific completion page, please click the link below. You must be redirected or click the link to prove that you have completed this survey.**

<div style="background-color:red; color:white; text-align:center;">Annonymized</div>