

I'm not a bot. SOUPS2022
Please answer to the following questions to verify you are not a bot.

Choose the correct answer from the choices below.
A: Connecting a USB flash drive to a computer in public is a security risk.
B: Charging a smartphone via USB port on a computer in public is a security risk.

You cannot re-select.

Only statement A is correct.
 Only statement B is correct.
 Both statements are correct.
 Both statements are wrong.
 I cannot tell.

DualCheck : Exploiting Human Verification Tasks for Opportunistic Online Safety Microlearning



Ryo Yoshikawa



Hideya Ochiai



Koji Yatani

IIS-Lab, The University of Tokyo

Background

Issues on online safety literacy

- **Learning online safety & ethics** is critical for everyone.
- However, not so many people learn such knowledge regularly.
 - **Only 17.9%** of smart device users had taken explicit training on online ethics^[1].
- It is important to learn about Internet security and ethics in **daily life**.
 - Conventional learning styles like e-learning may not fit well to their lifestyles.

[1] Japan Information-technology Promotion Agency. “Awareness survey on information security ethics in fy2019 report”
<https://www.ipa.go.jp/files/000080783.pdf> (2019)

Our Idea

We propose DualCheck

- Integrating online safety microlearning into human verification task.
- 

- **Microlearning** is a learning style that learners do small learning tasks in a short time.
- Often used for vocabulary learning etc. in HCI fields^[2,3] .

[2] Trusty et al. “Augmenting the Web for Second Language Vocabulary Learning” (CHI ’11)

[3] Cai et al. “Wait-Learning: Leveraging Wait Time for Second Language Education” (CHI ’15)

Our Idea

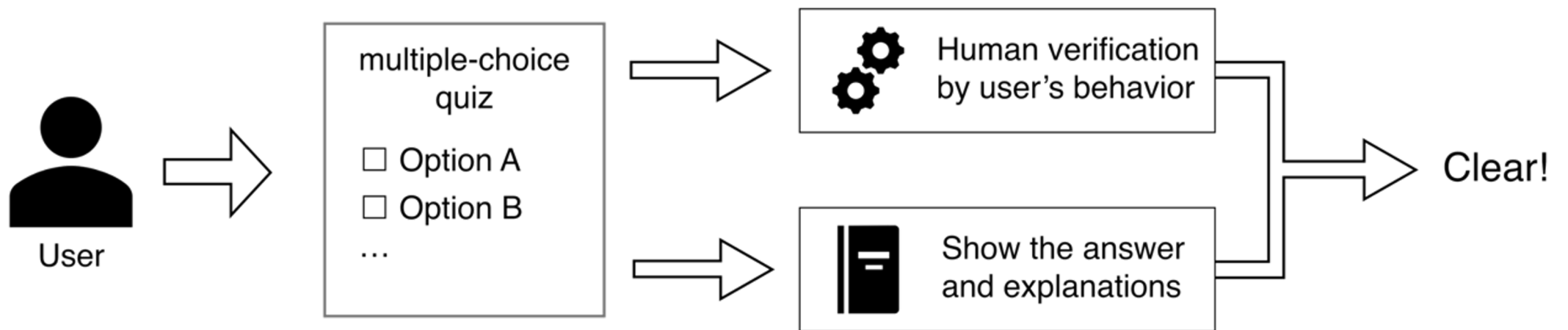
We propose DualCheck

- Integrating online safety microlearning into human verification task.
- Often appear when entering forms or logging in.
- The famous one is **CAPTCHA** (known as “I’m not a robot”).
- Our idea is to propose a question of online safety instead of conventional tasks.

Implementation

System design

- Users **respond to the question and read explanations.**
- DualCheck is expected to detect bot access by **cursor behavior** (like reCAPTCHA checkbox).
 - Human verification does **not depend on whether the answer is correct.**



Implementation

Prototype

- We implemented prototype in a Web environment.
- In our current implementation, a human verification mechanism is not integrated into DualCheck.
 - Our main purpose is to **investigate the learning effect and the usability** of DualCheck.
 - The code of reCAPTCHA is not publicly available.

1. Answering questionnaire

**This is the demonstration movie of DualCheck.
You can view it at the link below.**

<https://youtu.be/5W16fpjeg4M>

Please fill your name.

[Must]Q1: Have you read Internet news in a hour?

Yes
 No

[Optional]Q2: (For those who answered "Yes" to Q1,) which category of news have you read?

[Must]Q3: Please tell me the services what you will use today.[You can choose multiple choices]

SNS
 EC site
 Others

Q4: How do you think the use of Internet services will change in the future?

Will increase a lot
 Will increase a little
 Same as yesterday
 Will decrease a little

I'm not a bot.

Please answer to the following questions to verify you are not a bot.

Choose the correct answer from the choices below.

A: Connecting a USB flash drive to a computer in public is a security risk.
B: Charging a smartphone via USB port on a computer in public is a security risk.

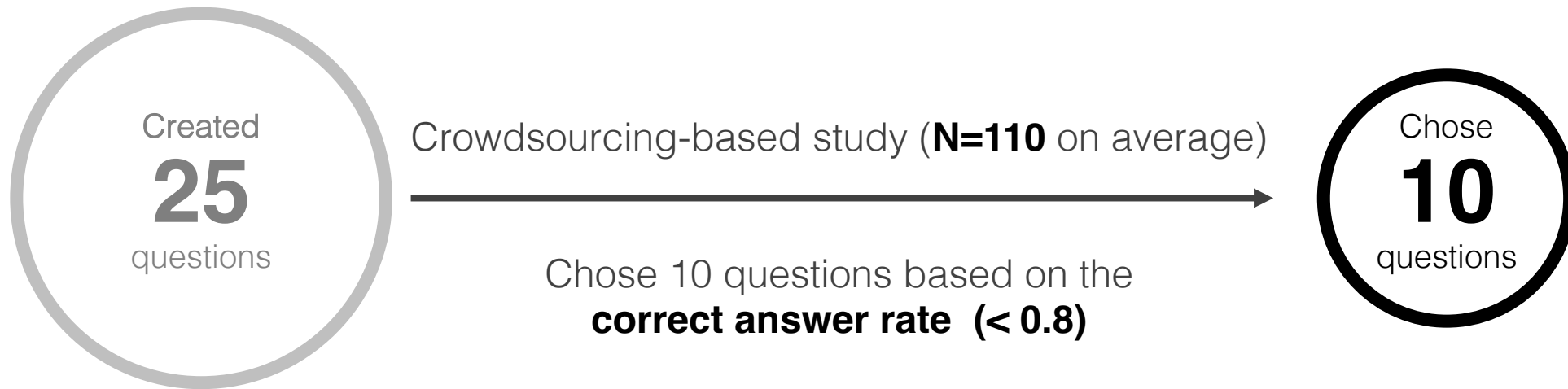
You cannot re-select.

Only statement A is correct.
 Only statement B is correct.
 Both statements are correct.
 Both statements are wrong.
 I cannot tell.

Question Curation

Prepared 10 questions

- We created a set of questions **related to online safety and ethics.**
 - The questions should be **moderately difficult.**
 - We collected common issues from **teaching materials for high school, etc.**



Deployment Study

Conducted 15-day study

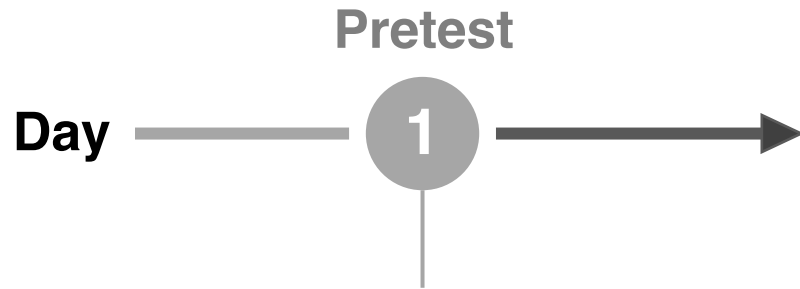
- We designed the 15-day deployment study to verify the **learning effect** and the **usability of DualCheck**.
- The task imitated the practical usage of DualCheck.
 - The participants answered to a questionnaire and then responded to DualCheck.

Deployment Study

2 modes of DualCheck

- We set two modes in DualCheck: OneTime / Repeat mode.
 - OneTime mode allows the users to go through **even if they answered wrong.**
 - Repeat mode demands them to answer **until they answer correctly.**
- We recruited participants through crowdsourcing service (N=34).
 - Split them randomly into two groups: OneTime **(N=16)** and Repeat **(N=18).**

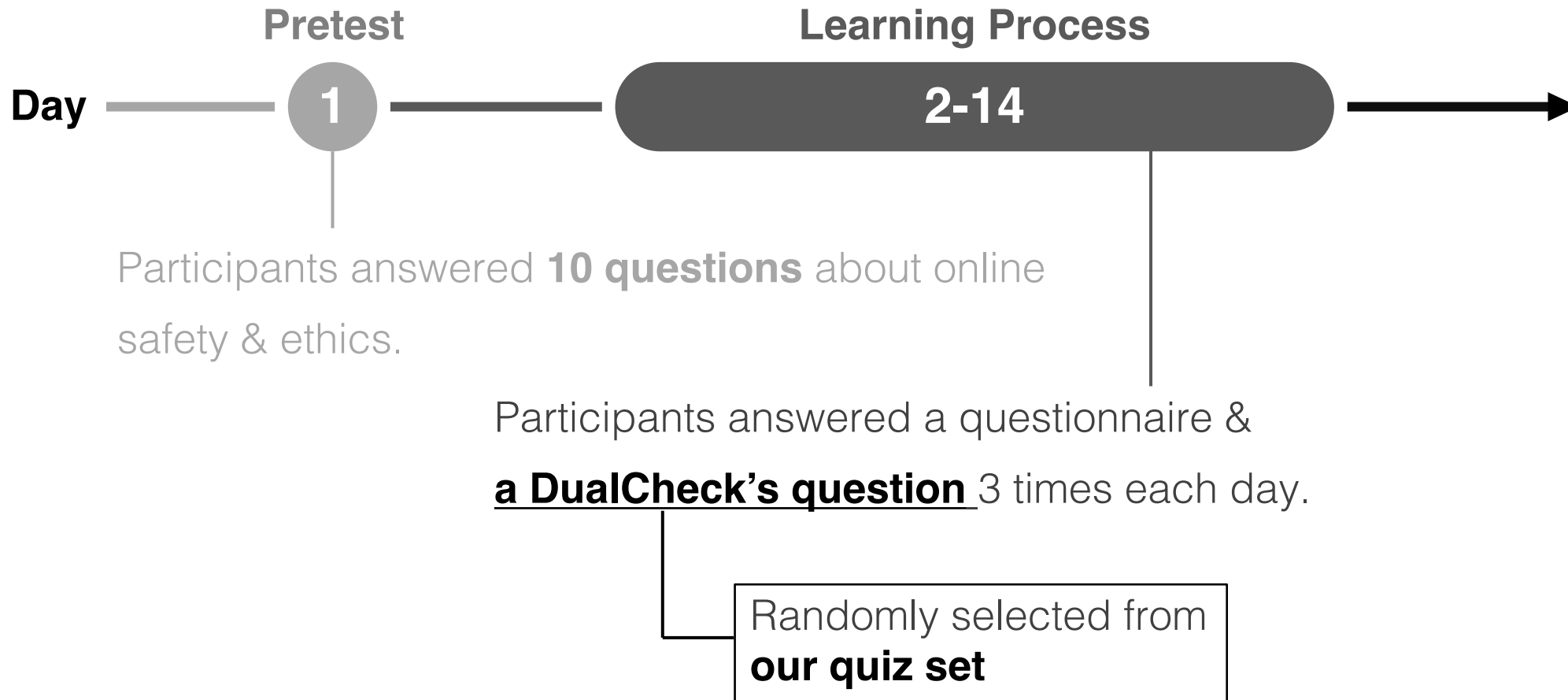
Deployment Study Procedure



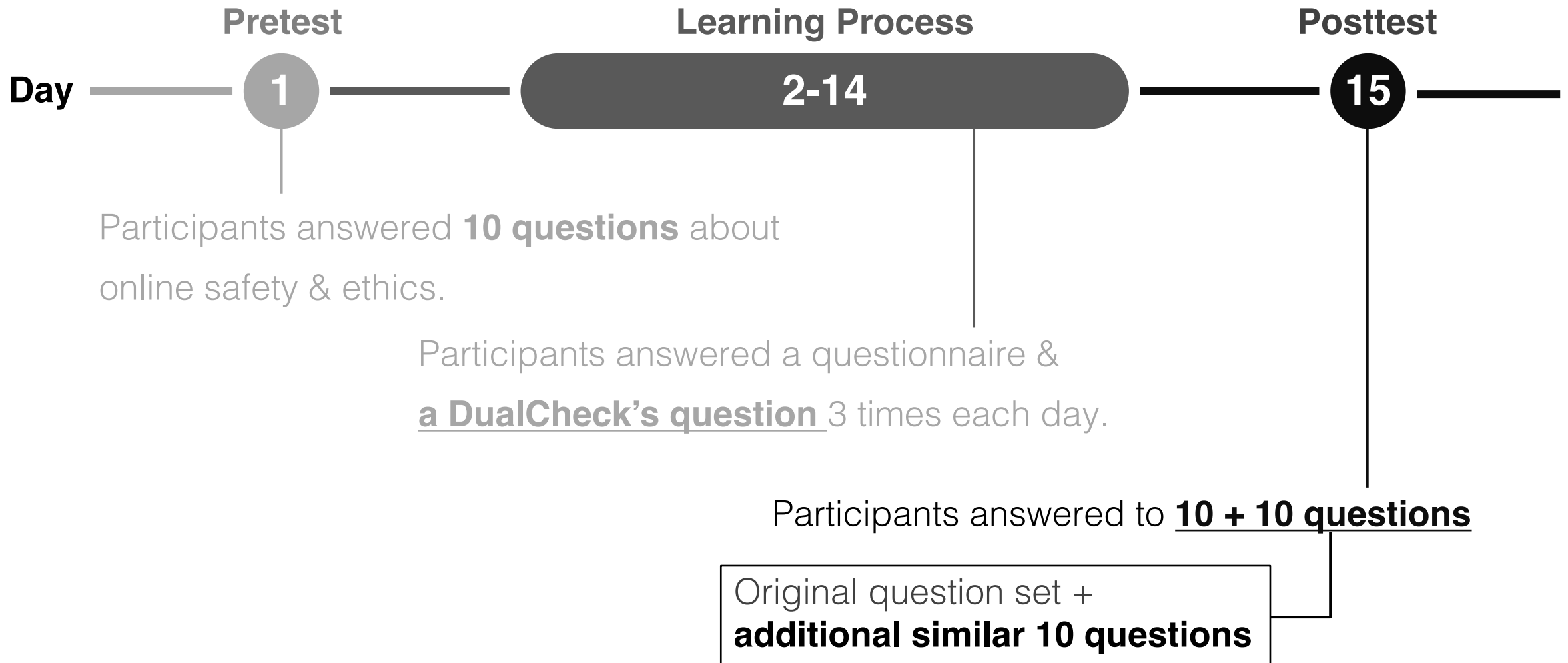
Participants answered **10 questions** about online safety & ethics.

The question set we made in our **quiz curation process**

Deployment Study Procedure



Deployment Study Procedure



Results

Mean accuracy rose after the study

- We conducted a two-way ANOVA for mean accuracy (=correct answer rate).
 - **Within** : Day1 vs Day15 / **Between** : OneTime vs Repeat
- The mean accuracy **rose significantly** after the deployment study.
 - 0.68 (SD=0.11) -> **0.94** (SD=0.04), $p < .001$
- We did **not find significant differences** in accuracy between OneTime / Repeat mode ($p = .30$).

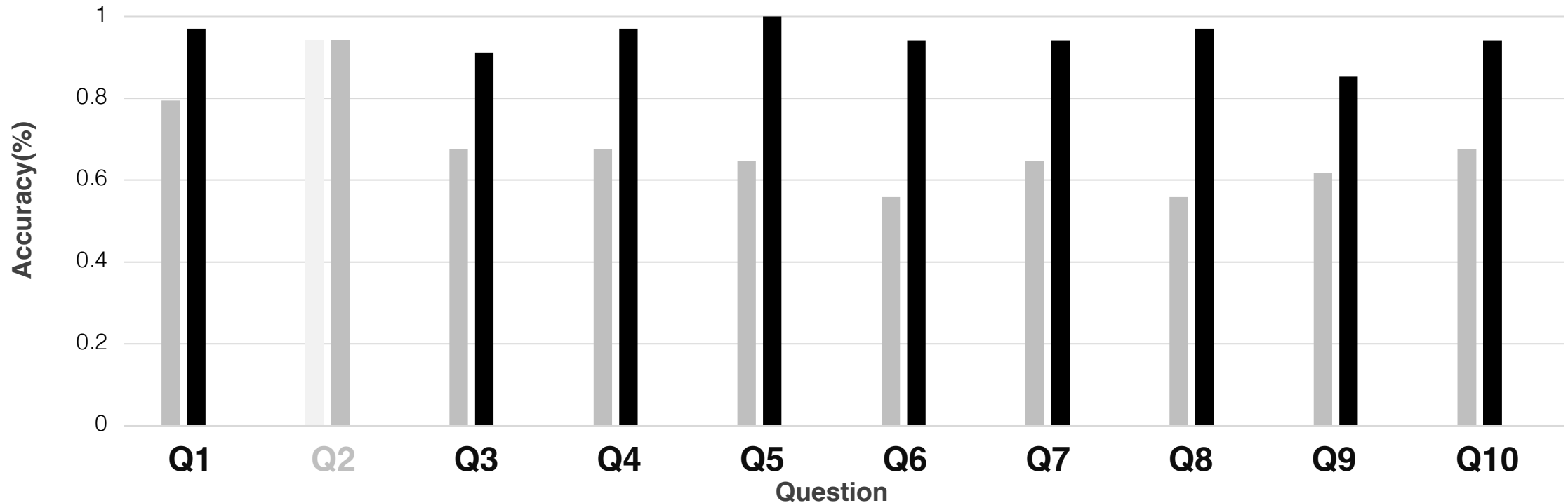
Results

Positive learning effect of original questions

Original Questions

9 questions showed **significantly higher accuracy** in day 15 than day 1.

■ Day 1
■ Day 15



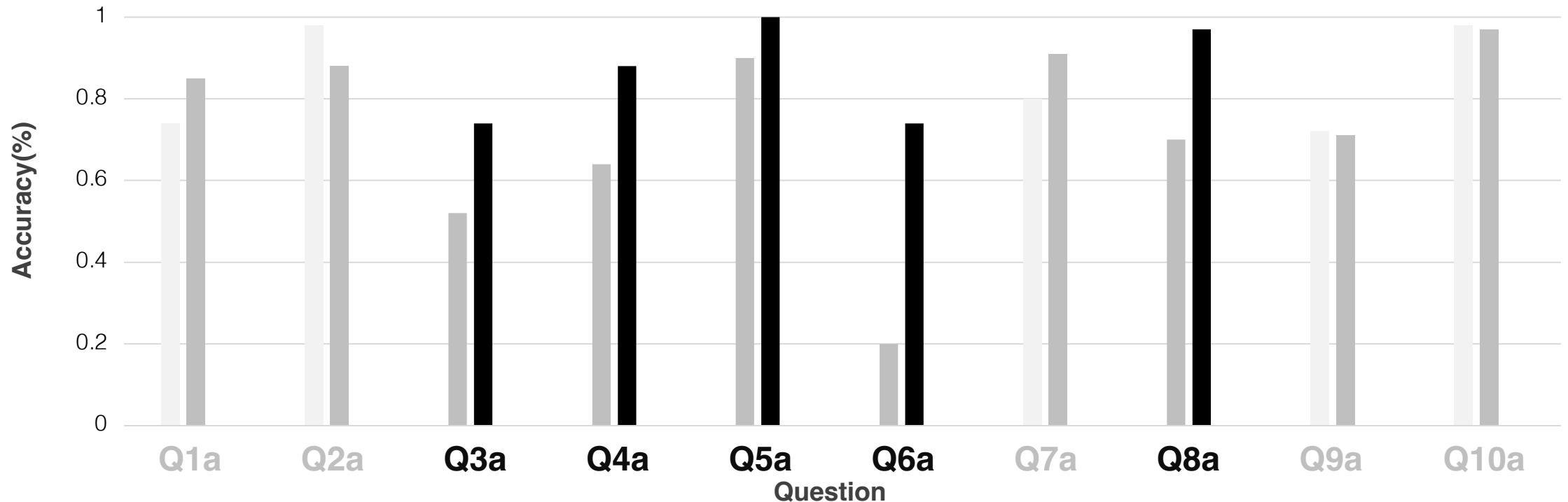
Results

Positive learning effect of similar questions

Similar Questions

5 questions showed **significantly higher accuracy** in day 15 than non-user's accuracy (N=100).

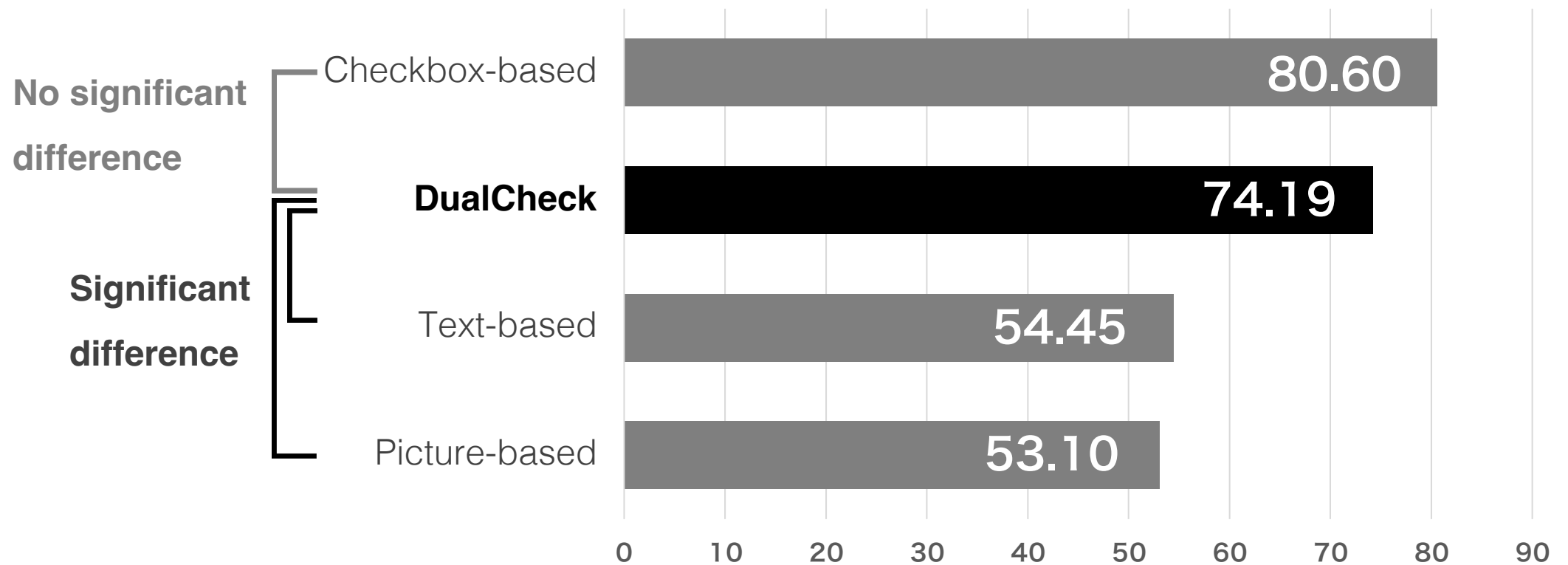
■ Reference accuracy
■ Day 15



Results

Higher SUS score

Compared SUS score with other CAPTCHAs.



Results

Comments for DualCheck

- Qualitative analysis

For more information, please read our paper.

The categorization of participants' comments

Questions	
Question difficulty	15
Issues on question presentation	9
Issues on answer explanations	5

Advantages of DualCheck	
Perceived advantages	15

Usability of DualCheck	
Positive opinions on usability	11
Issues on usability	10

Suggestions	
Possible improvements	8

Limitations and future work

Limitations and future work

- **Task design** might affect the evaluation of DualCheck.
 - Participants' attribute, sampling method, etc.
- Investigation of **other topics / formats of questions** is important.
- Thinking about more **practical aspects** of DualCheck is also important.
 - Integrate reCAPTCHA or other CAPTCHA.
 - Maintain and update question sets.

Conclusion

DualCheck's potential as a learning platform

- Our user study confirmed the **positive learning effect** of DualCheck.
- It also confirmed **higher perceived usability**.
- DualCheck would support Internet users **learning online safety and ethics**.

Summary

Thank you for listening!

Contact

Ryo Yoshikawa

ryo@iis-lab.org

Title	DualCheck : Exploiting Human Verification Tasks for Opportunistic Online Safety Microlearning
Design	<ul style="list-style-type: none">• Users answer to a multiple-choice question about online safety• DualCheck simultaneously process human verification
Study	<ul style="list-style-type: none">• Participants used DualCheck for 15 days• Collected the accuracy of target questions and analyzed
Results	<ul style="list-style-type: none">• Confirmed positive learning effect of DualCheck• Also confirmed higher perceived usability

Extra Questions

	Statement and answer
Q1	<p>A: Connecting a USB flash drive to a computer in public is a security risk.</p> <p>B: Charging a smartphone via USB on a computer in public is a security risk.</p> <p>Correct Answer: Both statements are correct.</p>
Q2	<p>A: On social networking sites, there is no privacy problem in sharing selfies and other information if you give limited access.</p> <p>B: On social networking sites, if you don't post any personal information, your identity will not be identified.</p> <p>Correct Answer: Both statements are wrong.</p>
Q3	<p>A: This is the first time I visited this Website, but I thought it was safe because it had a key symbol on my browser, so I entered my personal information.</p> <p>B: I entered my personal information on a Website beginning with http:// . It is risky to enter personal information on such a Website.</p> <p>Correct Answer: Only statement B is correct.</p>
Q4	<p>A: Passwords should be a combination of letters, numbers, and symbols that are difficult to remember.</p> <p>B: Passwords are safer if they are based on personal information, such as your hobbies, and avoid famous words that are easily guessed.</p> <p>Correct Answer: Only statement A is correct.</p>
Q5	<p>A: When the earthquake struck, local people posts the situation in the area. Even if you don't know whether it is true information, it is better to share the information quickly.</p> <p>B: When spreading information when an earthquake or other event occurs, it is better to only spread posts by the government or news organizations.</p> <p>Correct Answer: Only statement B is correct.</p>

Extra

Questions

Q6	<p>A: A cookie is a piece of information that sends a user's name and other personal information to a site administrator.</p> <p>B: Cookies are used for retargeting advertisements and other purposes.</p> <p>Correct Answer: Only statement B is correct.</p>
Q7	<p>A: Documents created with online storage services and document creation tools are not disclosed to the public.</p> <p>B: Documents created with online services can be seen by others through searches.</p> <p>Correct Answer: Only statement B is correct.</p>
Q8	<p>A: The procedure for requesting information about an offensive social networking account has been made easier due to a change in the law.</p> <p>B: Even if there is an offensive SNS account, it is difficult to identify their source address.</p> <p>Correct Answer: Only statement A is correct.</p>
Q9	<p>A: To verify that the email you received was sent from a real bank or other sources, you check the back of the @ in the source address.</p> <p>B: Checking the domain is one of the most important things to ensure that the URL sent to you is authentic.</p> <p>Correct Answer: Only statement B is correct.</p>
Q10	<p>A: Photos taken with a smartphone may contain location information.</p> <p>B: If you post a photo without the location information to a social networking site, your location will not be identified.</p> <p>Correct Answer: Only statement A is correct.</p>

Extra

Questions

Q1a	<p>A: If you use a computer's USB port only to charge your smartphone, no viruses or other devices will be transferred.</p> <p>B: If you connect a USB flash drive to a shared computer, viruses and other malicious programs may be copied.</p> <p>Correct Answer: Only statement B is correct.</p>
Q2a	<p>A: On social networking sites, if you limit the number of people you can follow, there is no problem if you tweet personal information.</p> <p>B: Your identity can be identified based on your following relationship on social networking sites.</p> <p>Correct Answer: Only statement B is correct.</p>
Q3a	<p>A: Websites that start with http://... do not support encrypted communication.</p> <p>B: If the Website is capable of encrypted communication, it is safe to send personal information.</p> <p>Correct Answer: Only statement A is correct.</p>
Q4a	<p>A: Passwords should be a meaningless string of characters with symbols.</p> <p>B: It is preferable to create a password based on a hobby or something that you keep secret from others.</p> <p>Correct Answer: Only statement A is correct.</p>
Q5a	<p>A: An earthquake occurred, but there was no information from the news media or government, so I spread a post made by a person claiming to be a local.</p> <p>B: When the earthquake occurred, a person claiming to be a scholar on Twitter explained the situation. It is considered as credible information.</p> <p>Correct Answer: Both statements are wrong.</p>

Extra

Questions

Q6a	<p>A: The use of cookies can customize ads.</p> <p>B: Allowing the use of cookies is likely to leak personal information.</p> <p>Only statement A is correct.</p>
Q7a	<p>A: Documents created with online document creation tools are not likely to show up in a Web search.</p> <p>B: It is important to check the publication settings of documents created with online tools.</p> <p>Correct Answer: Only statement B is correct.</p>
Q8a	<p>A: It is difficult to identify the source address of an anonymous social networking account.</p> <p>B: You can file a request for disclosure of sender information against an offensive social networking account.</p> <p>Correct Answer: Only statement B is correct.</p>
Q9a	<p>A: Checking the domain of the URL is important to confirm whether it is genuine or not.</p> <p>B: I received an email claiming to be from my bank. It was the same domain as the bank's email, so I figured it was the right email.</p> <p>Correct Answer: Only statement A is correct.</p>
Q10a	<p>A: The scenery and objects in the photo could lead to the identification of personal information.</p> <p>B: Location information may be stored in the photo.</p> <p>Correct Answer: Both statements are correct.</p>

Extra

Accuracy transition

