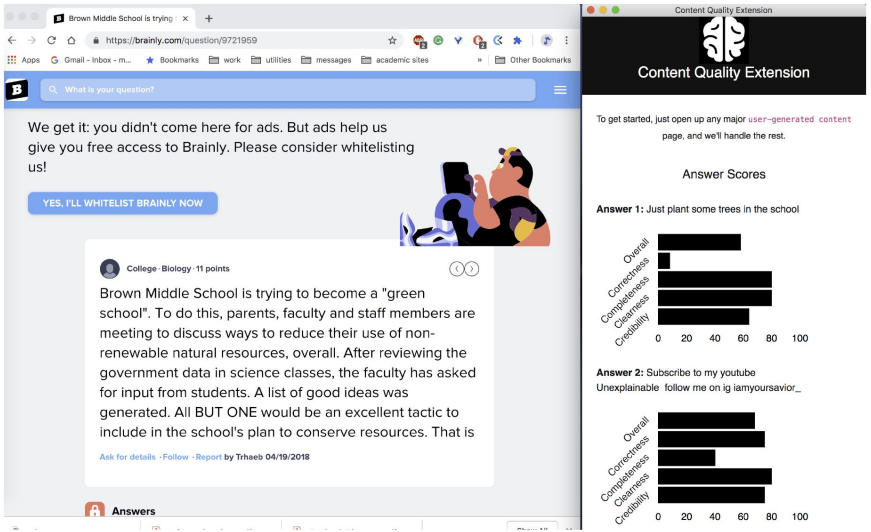
UGCQ-Checker [Doctoral Research]

UGCQ-Checker, a browser-based extension which assists users in evaluating the quality content (i.e. displaying scores) in question-answering sites for young learners

Watch the video:
https://www.youtube.com/watch?v=r1s5clqtyMo&ab_channel=ManasaRath



RUTGERS

Role

As a doctoral research, I built end to end processes of architecting software building, conducted usability testing studies to gain insights into how users make use of **the browser-based extension** (MVP) as a part of a **federal grant**

Team

Project lead, **collaborated with software engineers, supervisor**

Duration

June, 2018 - June, 2020

Skills and Tools

Tools: R, Python, HTML, CSS, JavaScript, PHP, xampp, MS Office

Skills: Usability testing, Design thinking, Data Analysis, Front-end coding Contextual Inquiry, Interviews

Problem:

- Users (K-12 students) make use of content on educational question-answering (Q&A) for course requirements
- No full-proof way to conceptualize the quality of content independent of consumers task at hand

Research Questions

RQ1: What were users' perceptions regarding the presence of automated quality scores while completing the task?

RQ2: How does users' engagement vary with the change in task type (such as ranking. synthesis) in the presence or absence of automated quality scores?

Process

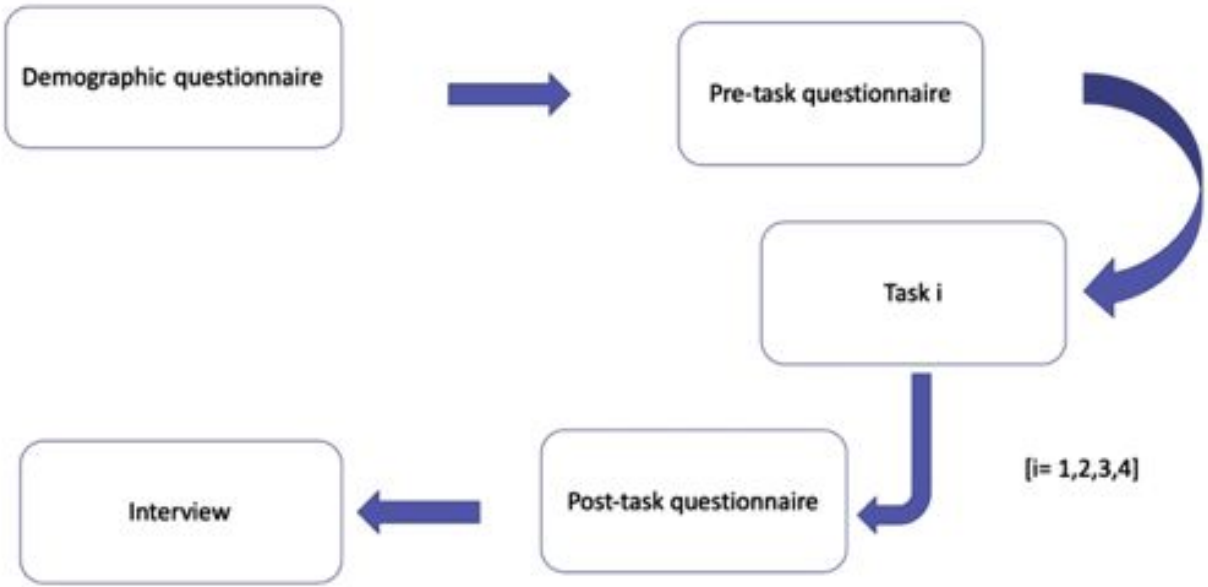
- To build the plugin, I collaborated with software engineers and my supervisor to ideate the concept and operationalize the idea
- Several prototypes were built to ideate and later engineered to build the extension
- After the extension was built, a control user-study was built to study users' perception of it while using it
- **A/B testing was performed** to understand to what extent users made use of the browser based extension
- A web-registration system was created consisting of pre-task/tasks/ post-task questionnaires to under the process
- To understand users' engagement with the quality scores, an eye-tracker was employed to record users' gazes

Process

- While carrying out the lab-studies, three group of participants were considered: Control, Treatment 1 & Treatment 2
    - Control condition: Participants were not shown the quality scores while performing the tasks
    - Treatment condition 1: Participants were shown scores while performing tasks
    - Treatment condition 2: Participants were shown manipulated scores while performing the tasks

- 15 Undergraduate students from Rutgers University were considered for the lab study
- After the users completed the tests, a semi-structured interview was conducted

Study Design

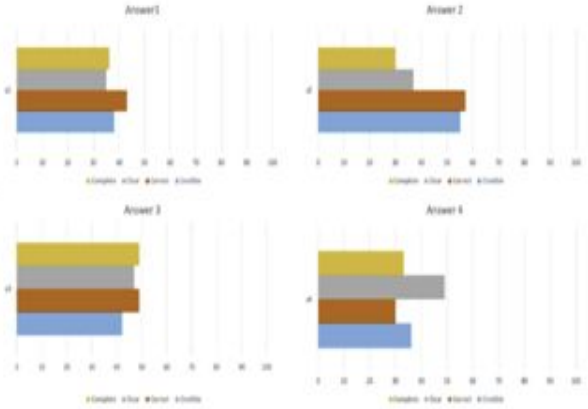


Study materials included in the experiments

Study interface : ranking & synthesis tasks

Question 1: When did recycling become popular?

Submit Order



Question 1: Why did dinosaurs become extinct? A consensus of scientific opinion says it was because of a meteorite impact 65 million years ago. Seems to me SOME would have survived.

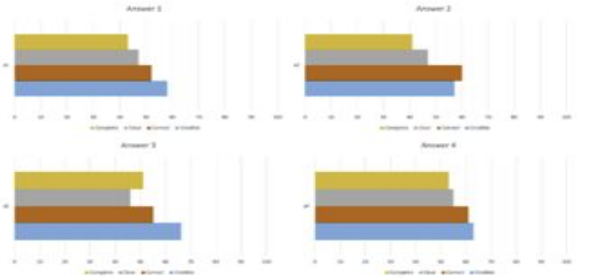
Answer 1: Some did. 65 million years is a long time to evolve into the lizards and birds we see today.

Answer 2 : Because behemoths (dinosaurs as a word was coined in 1849) are mentioned in the Bible it could also be they died in the Noahaxic Flood.

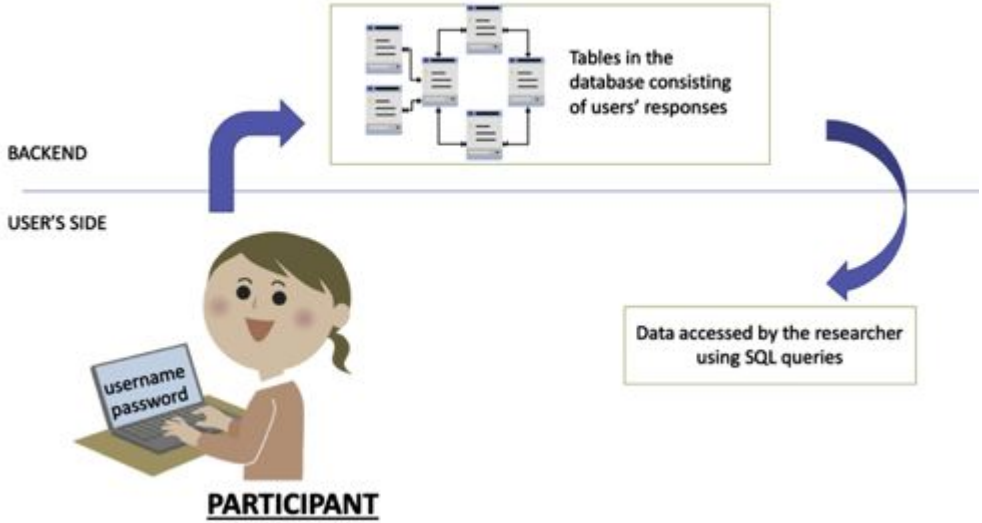
Answer 3: Surely they did? There are lots of animals around today which are descendents from dinosaurs.. birds, lizards etc.

Answer 4: The tuatara is considered to be from the oldest living line of dinosaurs.

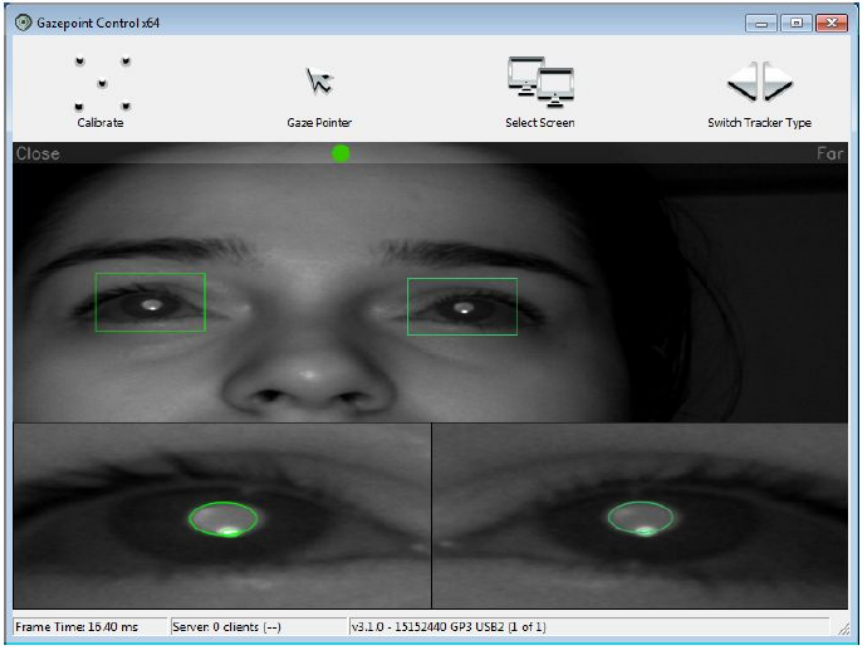
Submit



Specifics of a web-registration



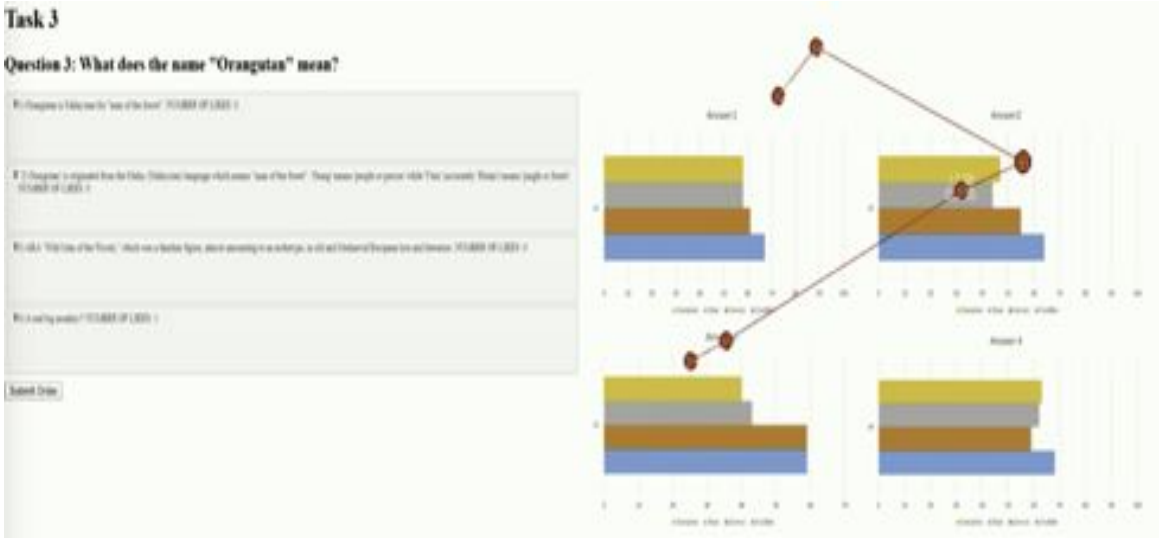
Eye-tracker: Calibration



The eye-tracker was calibrated and pilot tests were conducted before performing the actual study

Variables: Measuring users' engagement

-Areas of interests, Total number of fixations, Gaze duration, Average time per gaze were calculated



Analysis

- Statistical analysis was conducted to observe differences in users' behavior while completing the tasks

- Interviews were **transcribed and coded under relevant themes** to further understand users' perceptions on the use of quality scores in presence of quality scores

Findings

- Users only used quality scores when unfamiliar with the topic
  - *"I took help of the charts when I wasn't knowledgeable about the topic"*
- No differences found between difference in task type and users use of quality scores
- Genuine/Manipulated treated quality scores as real: Placebo effect
- Users primarily used: correct, credible
  - *"I used credibility and correctness as the criterions, also went through all the answers along with the questions to give in more context."*

Limitations

- Eye-tracking did not provide insights on users' intention on looking a particular point on the screen
- Higher sample size would have provided more insights between users' use of quality score
- Conducting a lab session along with a naturalistic study could have provided more context on users' use of quality scores