A PILOT STUDY COMPARING CHATGPT AND GOOGLE SEARCH
IN SUPPORTING VISUALIZATION INSIGHT DISCOVERY

Chen He[1],
Robin Welsch[2],
Giulio Jacucci[1]

[1]University of Helsinki,
[2]Aalto University

MOTIVATION

► Discovering insight from visualization takes time and effort.
► Researchers proposed techniques to **automate** data insights (e.g., [1]).
► However, automated data insights, such as data clusters, often lose the context of the **domain**, which hinders actionability [2].

Research Question. ChatGPT's immense reservoir of information could provide **domain-related insight** when prompted with the data under exploration as contexts.

What are the similarities and differences between contextualized ChatGPT and Google search in supporting visualization insight generation?

RELATED WORK

	ChatGPT	Google search
Medical information retrieval (IR) [3, 4, 5]	Information is more difficult to read and comprehend [4, 3]. More relevant information, but without sources [4]. Better at general medical information [5].	More reliable information with sources [4]. Better at medical recommendation [5].
Learning to program [6]	Better success rate with less time spent.	Students were better at understanding the topic.

CHATGPT-SUPPORTED CO_2 EXPLORER

We integrated ChatGPT 4 (vision) with an existing CO_2 Explorer (Fig. 1). The initial system prompt at the backend informs ChatGPT of the CO_2 emission data in CSV format, describes the visualization and the user task, and asks the chatbot to assist with the user task. User clicks on the year and country are prompted as system messages (Fig. 1E), with the resulting visualization as an image attached to the prompt.

STUDY DESIGN

As a between-subjects study, we compared contextualized ChatGPT with Google search in supporting insight discovery of the CO2 Explorer.

Participants. We recruited 25 internationals from a large university, 12 of which used Search (age range: 21-53, median: 25.5; female: 10) and 13 used Chat (age range: 21-45, median: 25; female: 6). The search group is familiar with Google search, while the chat group is not so familiar with ChatGPT (effect size = 0.76, $p < 0.001$).
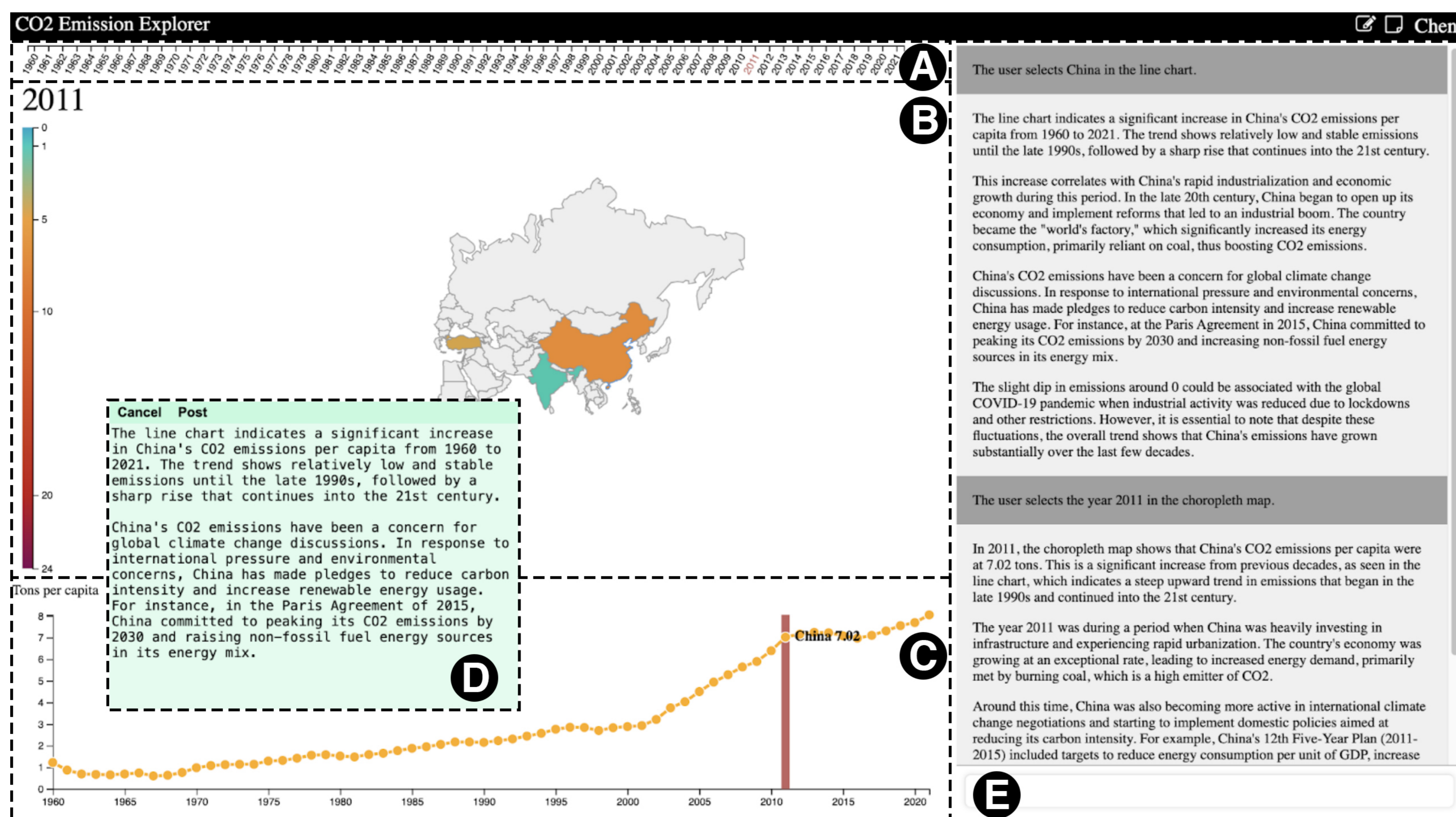
Procedure and tasks.

Table 1: Two insight discovery tasks.

Quantity	Quality
The more insights the merrier; An insight must contain external evidence.	A hypothesis or generalization with rationale and external evidence.

RESULTS

Results showed no significant differences between the two conditions in the time taken (Fig. 2) or the grades of notes for the tasks (Fig. 3); neither did the number of notes generated for the quantity task. In both conditions, the quantity task took more time than the quality task, while the quality task produced notes with higher grades.

Search has a better SUS score (85) than Chat (77.5). 38% mentioned that they did not like waiting for the chat's answer after a click. Two users in the chat condition used search.

Three failed to find the answer with search; one got a wrong answer with chats. Besides asking for facts and reasons, queries in chat also include when- and how-type of questions. 42% in Search put URLs in notes.

CONCLUSION

Both platforms have their merits and demerits. A future study will integrate both platforms for reliable and efficient IR to 1) avoid failure in IR, 2) improve answer correctness, and 3) obtain information sources.

REFERENCES

[1] Ding et al.: QuickInsights: Quick and Automatic Discovery of Insights from Multi-Dimensional Data. SIGMOD Conference 2019: 317-332

[2] Karer et al.: Insight Beyond Numbers: The Impact of Qualitative Factors on Visual Data Analysis. IEEE Trans. Vis. Comput. Graph. 27(2): 1011-1021 (2021)

[3] Bellinger et al.: BPPV Information on Google Versus AI (ChatGPT), Otolaryngology-Head and Neck Surgery (2023).

[4] Hristidis et al.: ChatGPT vs Google for Queries Related to Dementia and Other Cognitive Decline: Comparison of Results, J Med Internet Res 25 (2023) e48966.

[5] Ayoub et al: Head-to-Head Comparison of ChatGPT Versus Google Search for Medical Knowledge Acquisition, Otolaryngology–Head and Neck Surgery (2023).

[6] Arias Sosa and Godow: Comparing Google and ChatGPT as Assistive Tools for Students in Solving Programming (Bachelor thesis), 2023.

Figure 1: Screenshot of the interface of ChatGPT-empowered CO_2 Explorer for insight discovery. Users can select a year from the top list (A) to view that year's CO2 emission of various countries on the map (B), select countries from the map to view their historic CO2 emission in the line chart (C), chat with the chatbot (E) to gain more information about the data, such as news and events, and compose a note recording their discoveries (D).
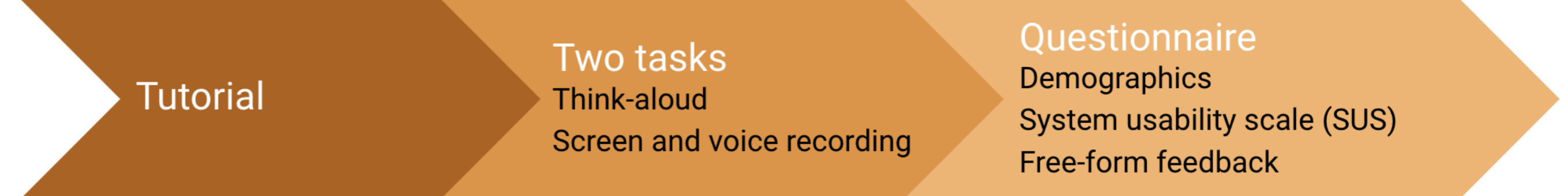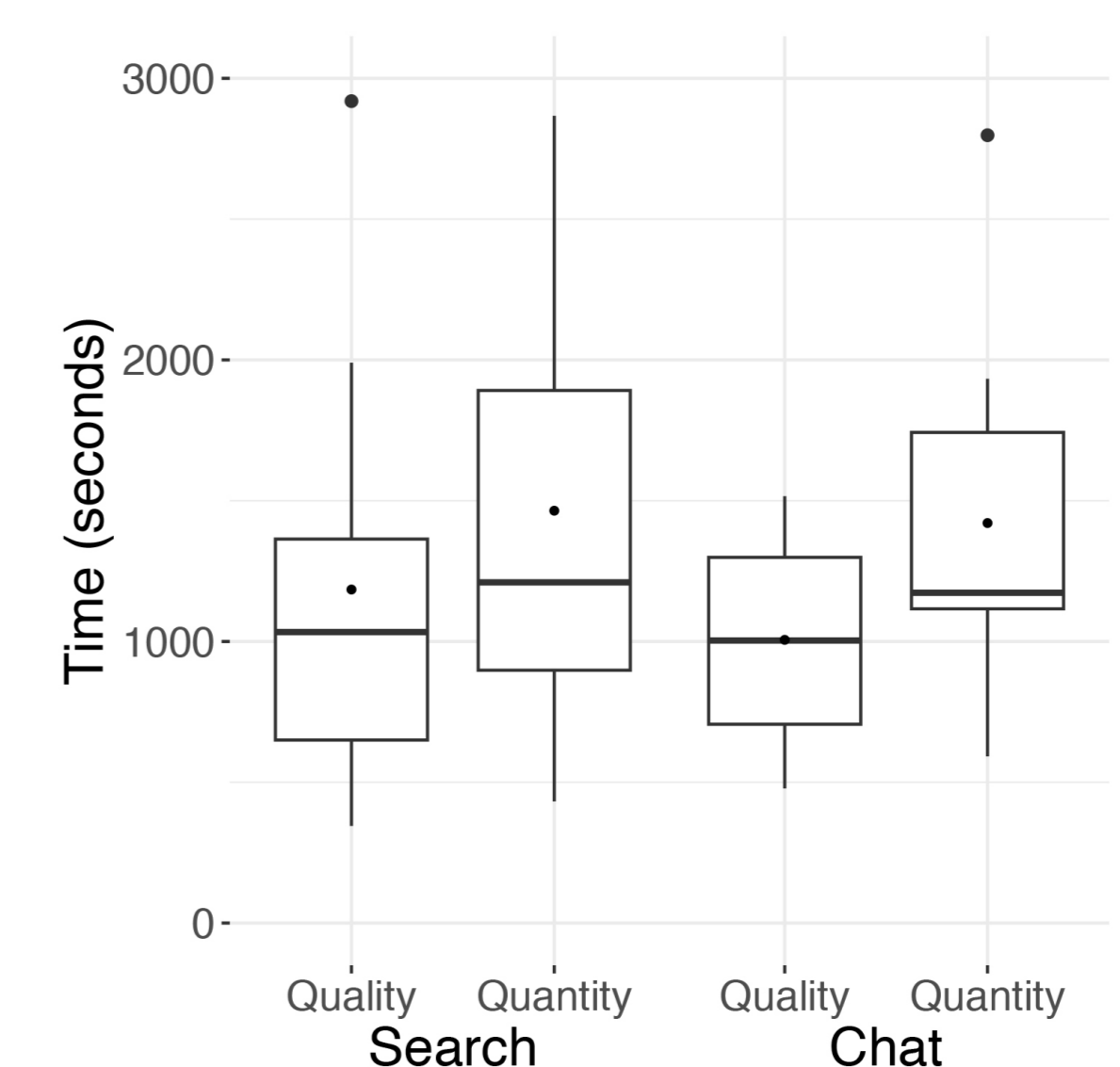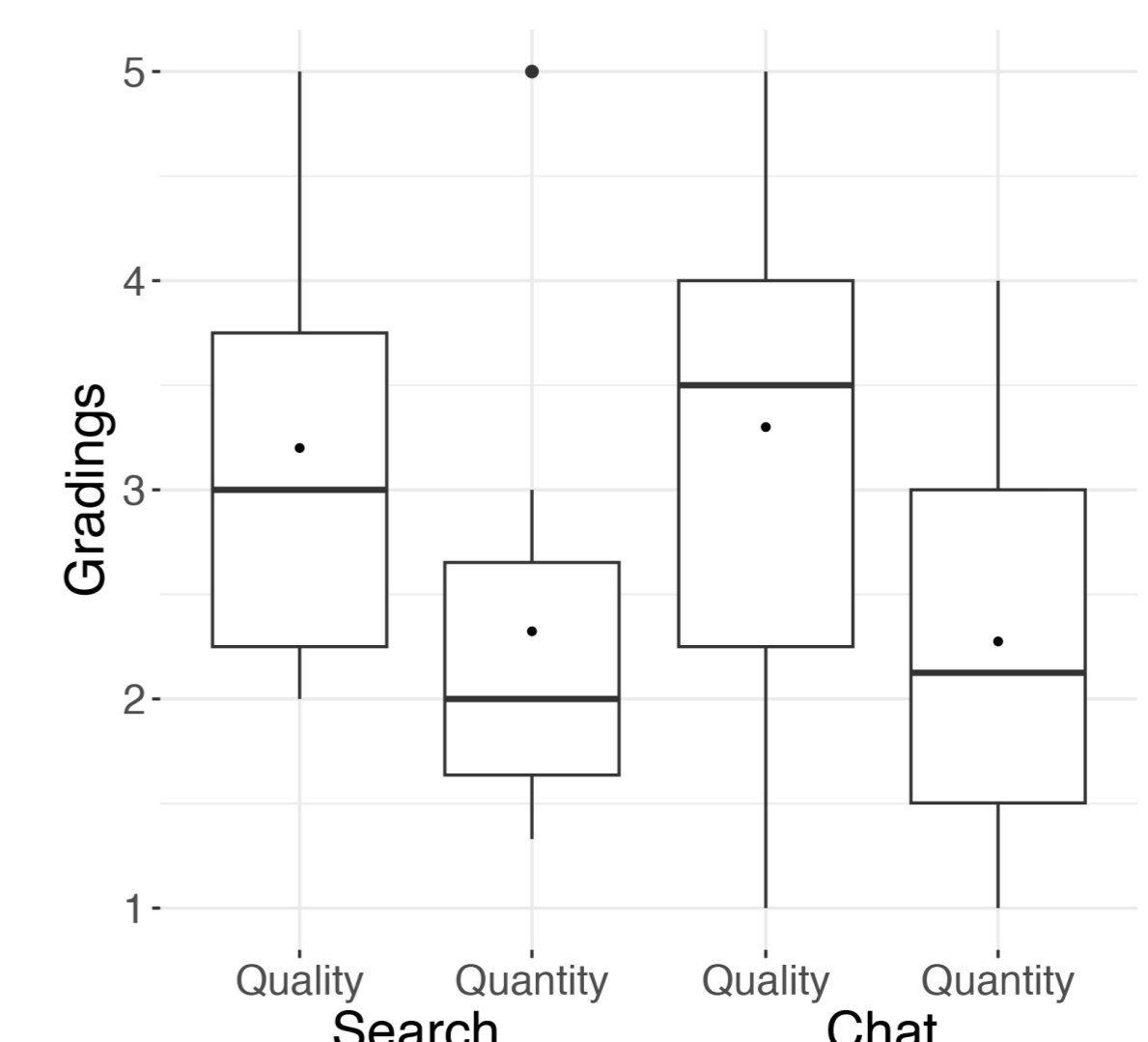

Figure 2: Time taken for the tasks.

Figure 3: Note gradings in a 5-point Likert scale with 5 being the best.