Task 1 Reflection

1. Empirical Analysis of Evaluation Processes and Technical Deficiencies

This evaluation verifies the robustness of the system through the Peer Pilot Protocol. Although the system follows the HTMX specification in technical implementation, the data of the balanced control experiment (N=4) reveals serious usability defects in the No-JS mode. The most significant data difference appears in task 4 (delete task). Server log analysis shows that the median operation time under No-JS conditions is 1.0 ms (MAD 0.0), which is significantly lower than 3.0 ms under HTMX conditions. Although the indicator is numerically "efficient", combined with the low confidence score and qualitative feedback of the participants **2/5**, the data actually reflects the lack of the system's security mechanism. The code review confirmed that the defect stems from the lack of confirmation steps in the form submission logic:

```
<form action="/tasks/1/delete" method="post">
    <button>Delete</button>
</form>
```

Although the above code meets the HTTP protocol standard, it violates the WCAG 3.3.4 (error prevention) principle at the interactive level. The empirical results show that mere technical effectiveness is not equivalent to the security of user interaction.

2. Criteria for Selecting Statistical Indicators

For the characteristics of the small sample (N=4) of this evaluation, the statistical analysis adopts Median and absolute median difference (MAD), instead of the mean and standard deviation. The mean index is extremely susceptible to the impact of a single abnormal interaction under such a sample size, and cannot accurately reflect the typical performance of the system. The data shows that in task 2 (add task), the MAD value of the HTMX group is 1.0 ms. This low-dispersion data provides strong evidence that the server-side processing performance is highly stable. This indicator effectively excludes the impact of back-end delay on user experience, thus confirming that the operation delay observed in No-JS mode is entirely attributed to the front-end Full-page Reloads mechanism.

3. Root Cause Analysis of Interaction Failures

Task 3 (editing task) has 100% functional failure under No-JS conditions, which is manifested by the user's inability to cancel the editing operation. An in-depth analysis of the implementation code shows that the function relies too much on client script attributes and lacks a semantic HTML backback mechanism. The following code snippet shows the implementation logic:

```
<button hx-get="/tasks/1/view" class="outline">Cancel</button>
```

In the running environment disabled by JavaScript, the above elements have neither the `type="submit"` attribute nor the `href` attribute, resulting in their appearance as an inoperable static element in DOM. This implementation violates the WCAG 2.1.1 (keyboard accessibility) standard, that is, all functions must still be operable without script assistance. This discovery provides a direct basis for the subsequent redesign of the progressive enhancement scheme using the `formaction` attribute.

4. Ethical Compliance and Research Limitations

This study strictly follows ethical guidelines and obtains informed consent before testing (see `consent-script.md`). However, the anxious reaction of No-JS users due to the accidental deletion of data during the observation process shows that the existing protocol is insufficient in psychological protection measures. Future research should carry out Debriefing immediately after the test to clearly inform users that errors are caused by system design defects. In addition, the conclusion of this evaluation about screen readers is mainly based on code inference. Given the difference in performance of ARIA Live Regions under different browsers and auxiliary technology combinations, future verification work must include live tests based on NVDA or VoiceOver to ensure the validity of accessibility conclusions.