University of Nottingham
UK | CHINA | MALAYSIA

COMP2044 Human Computer Interaction

Coursework 2: Evaluation

Group 7

Haotao Wang, 20517820

Zelin Xia, 20513999

Qinghua Zhang, 20514001

Yuyang Zhang, 20514470

Yingcong Wei, 20413599

Sihan Zhang, 20417195

Yixuan Zhou, 20511652

May 2, 2025

School of Computer Science

University of Nottingham Ningbo China

Chapter 1

Introduction

1.1 Purpose of the Prototypes

To meet the growing demand for campus event management, the University of Nottingham Ningbo (UNNC) needs to develop a mobile application to help teachers and students easily discover and RSVP for various campus activities, such as social, academic and professional activities. The development of this low-fidelity prototype aims to provide a reference model for subsequent application development, ensure a smooth development process, and achieve good human-computer interaction design at an early stage, thereby reducing the additional modifications that may occur in subsequent system development, saving time, manpower and resources.

1.2 Research Objectives

The main goal of this evaluation is to record the user experience and determine the best design solution to improve the usability and user satisfaction of the application. The evaluation will focus on several specific goals, such as whether the prototype has sufficient feedback mechanisms (such as pop-up prompts), whether there are clear guidelines for different tasks to help users complete specified behaviors, and whether the overall page information is sufficient and clear. By achieving these goals, the evaluation will provide more sufficient reference for subsequent prototype improvements, helping to design prototypes that are more in line with human-computer interaction principles and improve user experience.

Chapter 2

Methodology

This evaluation uses three evaluation methodologies to evaluate the usability of the RSVP function and the reminder settings in Coursework 1, such as cognitive walkthrough(CW), collaborative evaluation, and the system usability scale (SUS). These methodologies analyze prototype task efficiency, user feedback, and usability.

2.1 Cognitive Walkthrough

A task-oriented technique called cognitive walkthrough mimics the user's problem-solving process to assess the usability of a prototype (Polson et al., 1992). Each team member conducted a CW on another member's CW1 prototype, focusing on key tasks such as the RSVP function and the reminder settings. Evaluators answered questions about the prototype. By having individuals evaluate others' prototypes rather than their own, it helps simulate a user's mindset more effectively and avoids the emotional attachment that might occur when developers spend too much time on their own prototypes. Emotional attachment can make it difficult for designers and stakeholders to objectively assess their designs' strengths, potentially hindering collaborative evaluation and reducing openness to changes (Dam & Teo, 2020). The question about CW in FigureA.4. This method identifies potential usability issues based on cognitive principles.

2.2 Collaborative Evaluation

Collaborative evaluation is a user-centered approach that collects qualitative usability data through direct user participation (Monk et al., 1993). After the cognitive walk-

through (CW), our group recruited participants from other COMP2044 groups and selected one of our prototypes for group evaluation. Participants completed tasks such as RSVP function and the reminder setting while expressing their thoughts using the Think Aloud Protocol (TAP) (Nielsen, 1994). This process allowed us to capture users' cognitive processes and help identify potential issues in the interface. Before participating, users signed an informed consent form. A moderator guided interactions while another team member recorded key data including task completion time and accuracy. We used a structured data collection method based on the HTA in Figure 2.1 and Figure 2.2 and analyzed quantitative and qualitative feedback to guide prototype improvements. Finally, we collected quantitative data based on the recorded results of the collaborative assessment and plotted the time it took participants to complete the task in a bar graph, and also used a line graph to show the maximum, minimum, median, and average time. In addition, we conducted a qualitative analysis and recorded the confusion questions that participants encountered during the task.

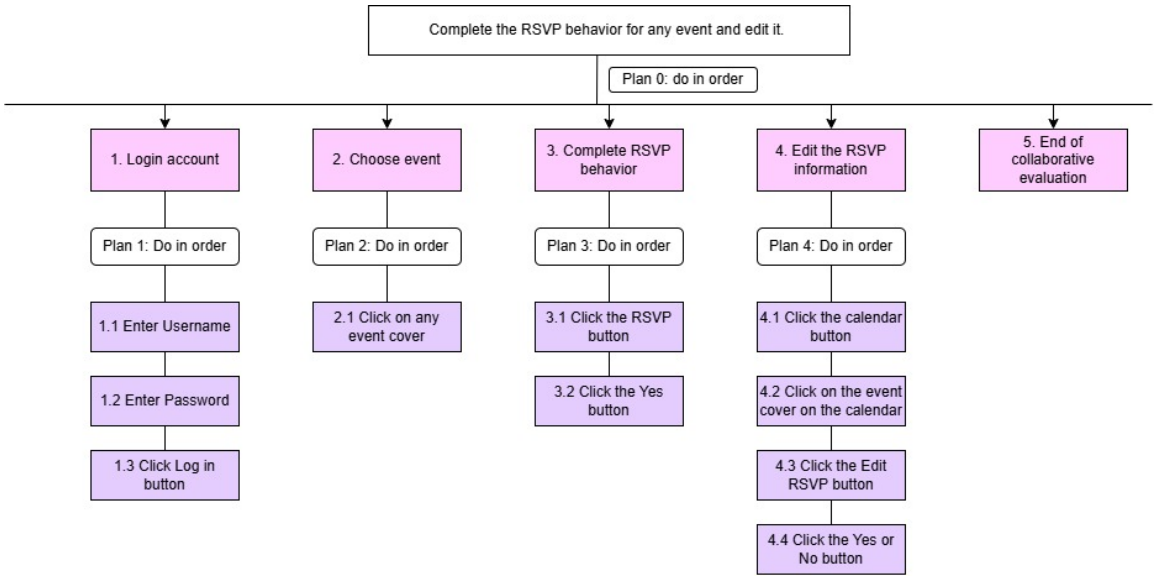


Figure 2.1: *HTA for Collaborative Evaluation Task 1: Complete the RSVP behavior for any event and edit it.*

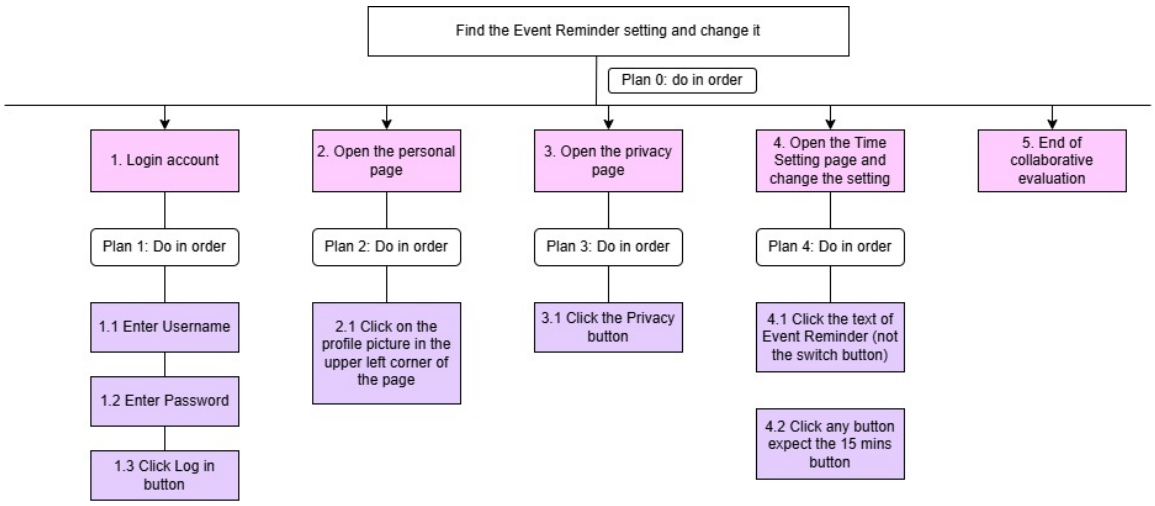


Figure 2.2: *HTA for Collaborative Evaluation Task 2: Find the Event Reminder setting and change it.*

2.3 System Usability Scale

The group used a standardized ten-item questionnaire developed by (Jordan et al., 1996) to quantify user satisfaction. The System Usability Scale (SUS) is a widely used evaluation tool designed to measure the overall user satisfaction with a system. It consists of ten questions covering aspects such as the ease of use, functionality, and user experience of the system. Each question is rated on a five-point scale from "strongly disagree" to "strongly agree." After the collaborative evaluation, we asked participants to complete a paper version of the SUS questionnaire to rate the overall usability of the prototype. All feedback was recorded and organized into a systematic dataset as an important part of the usability evaluation. This data will be used for quantitative analysis and the results will be made into a bar chart to help us further understand the user's perception and feedback on the prototype, thereby providing guidance for subsequent prototype improvements.

2.4 Evaluation Planning and Implementation

Before the data collection phase, the team first conducted a cognitive walkthrough (CW) and a system usability scale (SUS) to evaluate other members' prototypes, and then voted to select the prototypes to be used in the future. In order to ensure that the evaluation covered the core functions and user experience of the prototype, we designed two specific evaluation tasks: the use of the RSVP function and the setting of the reminder time and modified it to 20 minutes in advance.

During the process of recruiting participants to use the prototype, the evaluation was guided by a moderator to ensure that the participants could proceed smoothly according to the scheduled tasks, while another team member was responsible for recording the time and various reactions of the participants during the task execution. After all participants completed the task, the participants filled out the paper version of the SUS questionnaire for system satisfaction evaluation. The paper questionnaire was selected for offline testing in order to standardize the test conditions and avoid the potential interference of the unstable network environment on the test results. At the same time, the researchers could directly observe the operation process of the participants, thereby effectively improving the reliability and validity of the data (Albert & Tullis, 2013).

Subsequently, we entered the feedback data of all questionnaires into the SUS website to generate the final score. Finally, the collected quantitative and qualitative data were systematically summarized and analyzed. In terms of data processing, literature search, and drawing tasks, two team members were assigned to each task to ensure that each member could participate in the specific tasks. After the data was collated, the team members wrote the final report in accordance with the division of labor to ensure that the tasks were distributed fairly and all members could participate.

Chapter 3

Results

3.1 Prototype Selection

In terms of prototype selection, the team chose the prototype of Member 2. The following is a proof of the rationality of the prototype selection.

3.1.1 Selection Justification

To select the prototypes suitable for collaborative evaluation, the team systematically evaluated each prototype by completing the Cognitive Walkthrough (CW). The following is an overview of the evaluation results for each prototype.

- **Member 1**: The prototype passed all CW questions, none suggestion.
- **Member 2**: The prototype meet all CW requirements. However, many lack of guidance and feedback pages, overall completion is good.
- **Member 3**: The basic functions have all been implemented. When users want to delete the activities they have selected, the interactive keys are missing.
- **Member 4**: It feels not detailed enough.Then the Settings section is missing.Although the picture is in color, it seems that the overall color is not suitable.Then there are too many characters, which are not intuitive enough.
- **Member 5**: It meets basic functional requirements, but the page design does not meet the UNNC brand requirements and lacks guidance.
- **Member 6**:

- **Member 7**: Lack of guidance. Page layouts are lack of organization.

In order to have a more intuitive understanding of the peer evaluation results of the prototype, after completing the CW questionnaire, the team completed the SUS questionnaire based on their understanding of the prototype of the evaluation object, and statistically calculated the SUS scores of each prototype as follows:

An overly perfect prototype can hinder collaborative evaluation by appearing too pol-

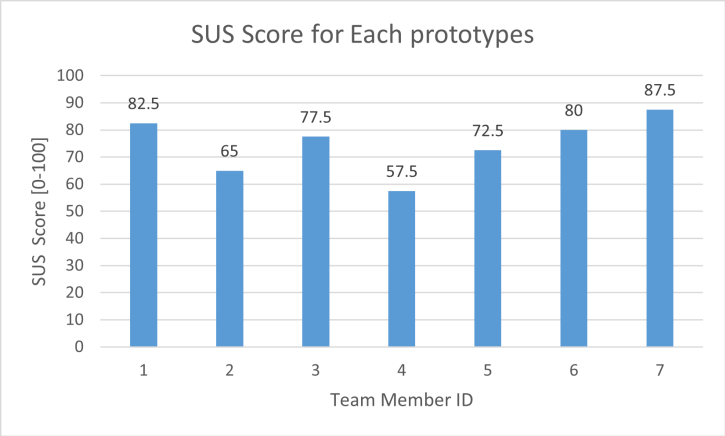


Figure 3.1: *SUS Scores for Cross-Member Prototypes*

ished or stereotyped, which could discourage stakeholders from providing feedback or suggesting improvements (Yeomans et al., 2006). After considering both the quality of the prototypes and the feedback of the participants, the team ultimately selected the prototype of member 2.

3.2 Collaborative Evaluation Results

3.3 System Usability Scale (SUS) Results

3.4 Result Discussion

This discussion successfully examined the usability of the RSVP system by analyzing the results of collaborative evaluation (cognitive walkthrough and task-based user testing) and the System Usability Scale (SUS). Although the average SUS score of Task 1 is ?? (Task 1 score) and that of Task 2 is ?? (Task 2 score) , this indicated that the prototype used as a survey template demonstrated higher usability than expected. A more in-depth

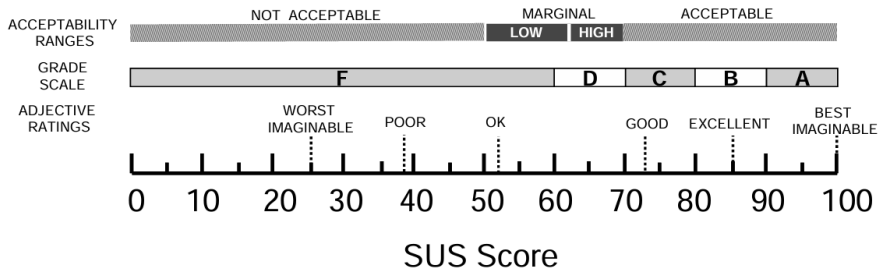


Figure 3.2: *A comparison of the adjective ratings, acceptability scores, and school grading scales, in relation to the average SUS score*

analysis uncovered subtle discrepancies and consistencies between subjective user perceptions and objective task performance metrics. By placing these insights in the context of Human-Computer Interaction (HCI) principles, we further illustrate our deficiencies in system design and thus gain deeper insights into user behavior patterns.

3.4.1 Consistency Between Methods

- **High SUS score reflects overall efficiency**

  The score of SUS reflects the overall efficiency. The high average score of SUS is consistent with the high task completion rate in the collaborative evaluation where users must complete simple operations such as logging in and selecting time options. It also demonstrates that, when completing the collaborative evaluation task, users find the system's design to be supportive in two key aspects. First, they appreciate the use of intuitive icons for tasks such as logging in and navigating to the calendar (e.g., the calendar icon for scheduling). Second, they think of a predictable linear workflow, which is structured (e.g., progressing from viewing events to confirming participation), allowing them to follow a logical sequence of steps to complete the evaluation. These advantages also correspond to the options selected by users in SUS.

- **The tasks verified by the two methods have common deficiencies**

  During the collaborative evaluation process, key issues were identified when editing RSVP information and accessing privacy settings, which were responded to in the SUS comments on complexity (e.g., "The system is unnecessarily complex"). Controls like the "Edit RSVP" button lack clear indications, which violates users'

expectations of immediate feedback (Norman, 2013).

3.4.2 Contradictions Between Methods

- **Paradox of high SUS score amid task failures**

  Users with high SUS scores still face non-negligible challenges when completing collaborative evaluations. In Task 3 (accessing the privacy page), without assistance from collaborators, most users failed to associate the time reminder with the "privacy" label. However, the SUS did not explicitly capture this inconsistency. This indicates that users may have made this mistake due to the illogical grouping of the system. In Task 5 (editing RSVP), a small number of users struggled to identify ambiguous labels, yet the SUS score remained high. This discrepancy could arise because users evaluated the system holistically rather than focusing on granular frustrations (Sauro & Lewis, 2016). This represents an identified conflict.

- **Time efficiency and cognitive load**

  The average task completion time (approximately 4 minutes) masks extreme outliers (for example, User 4 of Task 1: 6m 63s). SUS focuses on users' subjective perception of efficiency (for example, "I think the system is easy to use"), but fails to capture the tension experienced by users when they need to browse features such as the "personal page" for a relatively long time in the collaborative evaluation. This contradiction highlights the inability to represent the survey results with behavioral indicators in SUS.

3.4.3 Deeper Insights

- **Mismatched mental models**

  Problems such as terms (e.g., classifying time reminders under "Privacy") and ambiguous labels (e.g., "Edit RSVP") stem from the disconnection between user expectations and the classification of system modules. These issues conflict with (Johnson-Laird, 1983) mental model theory, which posits that users interpret interfaces based on pre-existing cognitive frameworks. When the system logic deviates from its internal framework, users will experience cognitive dissonance. Similarly, (Kieras, 1984) demonstrated that mismatched terms (e.g., "Edit RSVP" vs. RSVP "Modify participation status") increase cognitive load, and arbitrary labels of the system must be understandable to users.

- **Low visibility of key buttons**

  The low visibility of critical control buttons, such as the "Event Reminder" toggle and the "RSVP" button, forces users to rely on trial and error. Users often overlook these elements not due to a lack of knowledge, but because a weak visual hierarchy—such as insufficient contrast and diminutive size—induces uncertainty in selection and leads to wasted effort. Therefore, control buttons must be perceptually prominent to show their functions. This design defect disproportionately affects novice users, who may lack the experience to infer hidden functions and thus fail to achieve the desired effect (Carroll, 1990).

- **Accumulated frustration**

  The extended task completion time (e.g., 4 minutes and 63 seconds for the user) indicates system inefficiency may induce user frustration. This phenomenon is consistent with (Sweller, 1988) cognitive load theory, which holds that poor design brings additional mental demands and overwhelms the user's working memory. For instance, confusion when accessing event details during a task can result in delays in subsequent steps, as the user expends cognitive resources across different stages without progressing efficiently toward the goal. (Polson & Lewis, 1990) emphasized that although an individual task might eventually be completed, cumulative usability problems will reduce user satisfaction.

- **Role of evaluator prompts**

  The performance of users varies depending on the different prompt contents of the assisting evaluators (e.g., User 1 succeeded in following the prompt, whereas User 4 required additional time despite the prompt), highlighting the influence of the intervention of the assisting evaluators on the task outcome. This observation resonates with the meta-analysis of usability studies by (Hornbæk & Law, 2007), which demonstrated that subtle guidance, such as prompts, can artificially inflate success rates while potentially obscuring underlying design issues.

Chapter 4

Discussion

4.1 Finding During Evaluation

4.1.1 Contrasting User Feedback in CW and SUS Evaluations

During the process of selecting and evaluating prototypes by the team, the team found that there were some differences between the user feedback of CW and SUS. The following team will elaborate on these differences and conduct in-depth discussions.

- **Difference between CW and SUS**
  - **Description**: When the group was doing Stage 1 of the task, the group conducted an interactive cognitive walkthrough evaluation of the prototypes of the group members. During the evaluation process, in order to more intuitively display the evaluation results, the group asked each member to fill out the SUS questionnaire based on the evaluated prototype and calculate the SUS questionnaire score. After comparing the cognitive walkthrough results with the corresponding SUS questionnaire scores, the group found that some prototypes with good cognitive walkthrough results obtained lower SUS scores than some prototypes with poor cognitive walkthrough results.
  - **Discussion**: After observing this phenomenon, in order to avoid the same problem in the collaborative evaluation stage, the team conducted a corresponding cause analysis and literature research on this problem, and finally came up with the following reasons:
    - **Instability in the quality of online questionnaire feedback**:

* **Two-dimensional impact of prototype appearance and function**:
* **Insufficient sample size**:

4.1.2 Key Usability Issues

Our Stage 2 collaborative walkthrough uncovered several "speed bumps" in the interface that led to task failures and user frustration. Below we describe each issue, note which HCI principle it breaks, and report how many participants were affected.

1. **Confusing Menu Labels.** Two of seven participants looked under "Calendar" or "Notifications" but never found the Personal Page (Task 2). The label "Personal" felt too generic and didn't match users' real-world expectations (Nielsen, 1994). Those two users ended up with a 0% success rate on Task 2.

2. **Misleading Iconography.** All seven participants failed to open the Privacy page (Task 3) because the padlock icon didn't intuitively signal time-related settings. Breaking consistency and standards makes people second-guess what an icon means (Monk et al., 1993). Task 3 suffered a 100% error rate.

3. **Hidden Primary Controls.** Two participants completely overlooked the "Event Reminder" switch (Task 4). The toggle was small, low-contrast, and buried in a submenu—important controls should jump out at you (Norman, 2013). Those users had a 0% completion rate on Task 4.

4. **Ambiguous Action Buttons.** During RSVP editing (Task 1), three participants hesitated because the "Edit RSVP" button offered no clear cue and provided no immediate success feedback. Every action should deliver instant feedback so people know it worked (Polson et al., 1992). Their hesitation added an average of 30 seconds to Task 1.

5. **Lack of Confirmation Feedback.** Although everyone could select "20 Minutes" for reminders (Task 5), one person suggested adding a "Confirm" button to feel certain their choice stuck. Without that, they felt "in the dark" about whether it actually changed (Albert & Tullis, 2013). This uncertainty slightly lowered satisfaction scores in our post-task survey.

4.2 Design Recommendations

Based on the comprehensive analysis of cognitive walkthroughs, collaborative evaluations, and System Usability Scale (SUS) scores, we propose the following design recommendations to enhance the usability and user satisfaction of the prototype. These recommendations are divided into two main categories: immediate fixes and long-term improvements.

4.2.1 Immediate Fixes

- **Enhance Feedback Mechanisms:** A dedicated feedback page should be added to provide users with clear, actionable responses after key interactions, such as submitting an RSVP or changing a reminder setting. In addition, revise ambiguous feedback messages to make them more intuitive and informative. For instance, after editing RSVP details, a confirmation popup should state "Your RSVP has been successfully updated" instead of a vague "Saved."

- **Improve Guidance and Instructions:** Essential guidance elements should be integrated, especially for first-time users. Tooltips and short, task-specific tutorials can assist users in navigating unfamiliar features. For example, a step-by-step guide can explain how to locate and modify event reminders, addressing confusion noted during the evaluation.

- **Refine Button Labels and Icons:** Labels such as "Edit RSVP" and icons like the padlock for reminder settings should be replaced with more meaningful alternatives (e.g., "Modify Attendance" or a bell icon for reminders). This aligns with users' mental models and improves clarity.

4.2.2 Long-Term Improvements

- **Redesign Layout of Event Reminders:** The event reminder feature, currently buried within the privacy section, should be relocated to a more prominent and logically connected area, such as directly under the event details page. This redesign requires collaboration between design and development teams to ensure feasibility and effectiveness, and should be reassessed through further usability testing.

- **Increase Visual Hierarchy and Button Visibility:** Critical control elements like the RSVP and reminder toggles need improved visual prominence. This can be

achieved through adjustments in size, contrast, and positioning. Enhanced visibility will reduce errors, particularly for novice users.

- **Introduce Confirmation for Critical Actions:** Adding an explicit "Confirm" button for changes in settings (e.g., reminder time) will reinforce user confidence and reduce uncertainty, as suggested by participant feedback.

4.2.3 Prioritization Based on SUS Results

According to the SUS questionnaire data, the prototype's weakest points relate to complexity and insufficient feedback—these items received lower average scores compared to ease-of-use metrics. Thus, the initial priority should be improving feedback clarity and simplifying task flows, especially around RSVP editing and reminder adjustments.

4.2.4 Considerations for Future Evaluations

- **Broaden the Scope of Testing:** The current evaluation primarily focused on RSVP-related functions. Future evaluations should cover a wider range of prototype features to ensure comprehensive usability assessment.

- **Increase Sample Size:** While existing literature (e.g., Sauro & Lewis, 2016) supports the validity of SUS with small sample sizes, our current sample of 8 participants remains limited. Expanding the sample in future rounds will enhance the robustness of the findings and increase confidence in generalizability.

By implementing these recommendations, the prototype will be more aligned with established human-computer interaction principles, reduce user frustration, and ultimately improve the overall user experience.

References

Albert, B., & Tullis, T. (2013). *Measuring the user experience: Collecting, analyzing, and presenting usability metrics*. Newnes.

Carroll, J. M. (1990). *The nurnberg funnel: Designing minimalist instruction for practical computer skill*. MIT Press.

Dam, R. F., & Teo, Y. S. (2020). Stage 4 in the design thinking process: Prototype. *Interaction Design Foundation - IxDF*. https://www.interaction-design.org/literature/article/stage-4-in-the-design-thinking-process-prototype

Hornbæk, K., & Law, E. L.-C. (2007). Meta-analysis of correlations among usability measures. *CHI '07*, 617–626. https://dl.acm.org/doi/10.1145/1240624.1240722

Johnson-Laird, P. N. (1983). *Mental models: Towards a cognitive science of language, inference, and consciousness*. Harvard University Press.

Jordan, P. W., Thomas, B., McClelland, I. L., & Weerdmeester, B. (1996). *Usability evaluation in industry*. CRC press.

Kieras, S., David E.and Bovair. (1984). The role of a mental model in learning to operate a device. *Cognitive Science*, *8*(3), 255–273.

Monk, A., Davenport, L., Haber, J., & Wright, P. (1993). *Improving your human-computer interface: A practical technique*. Prentice-Hall New York.

Nielsen, J. (1994). *Usability engineering*. Morgan Kaufmann.

Norman, D. (2013). *The design of everyday things*. Basic Books.

Polson, P. G., Lewis, C., Rieman, J., & Wharton, C. (1992). Cognitive walkthroughs: A method for theory-based evaluation of user interfaces. *International Journal of Man-Machine Studies*, *36*(5), 741–773.

Polson, P. G., & Lewis, C. H. (1990). Theory-based design for easily learned interfaces. *Human–Computer Interaction*, *5*(2), 191–220. https://doi.org/10.1080/07370024.1990.9667154

Sauro, J., & Lewis, J. R. (2016). *Quantifying the user experience: Practical statistics for user research*. Elsevier.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science*, *12*(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4

Yeomans, S. G., Bouchlaghem, N. M., & El-Hamalawi, A. (2006). An evaluation of current collaborative prototyping practices within the aec industry. *Automation in construction*, *15*(2), 139–149.

Appendix A

Appendix

A.1 Evidence of Group Work

- **Offline Meetings**

  In Figure A.1, we held two offline meetings to discuss the project progress and plan and divide the next steps. In these meetings, team members shared their work progress and gave each other feedback. These face-to-face discussions ensured that team members reached a consensus.

- **Wechat Group**

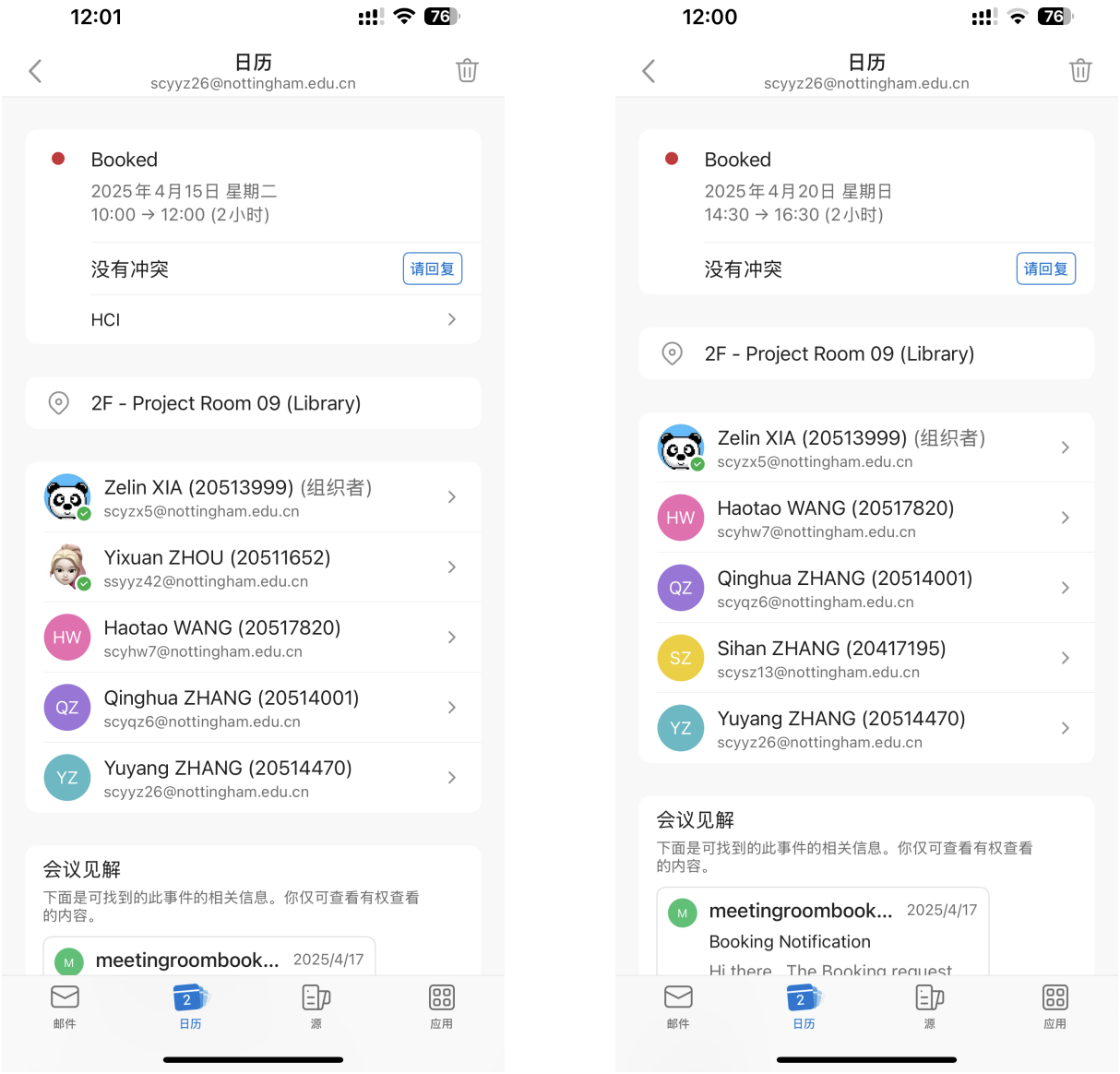
  To maintain continuous communication during the project, we set up a WeChat group in Figure A.2 where members exchange ideas, discuss issues, and coordinate tasks. The WeChat group is particularly helpful for solving problems outside of meetings, ensuring that everyone can get timely information and participate in discussions.

- **Overleaf Group**

  We created an Overleaf shared document platform in Figure A.3 for team collaboration in writing and editing reports. All team members can make changes, comments, and suggestions in real time. This collaborative approach improves work efficiency and ensures consistency and quality of report content.

- **Workshop**

  During the workshop, each team member actively participated and completed their tasks, ensuring the efficient progress of the tasks.



(a) *First meeting* **(b)** *Second meeting*

Figure A.1: *Meetings*

A.2 Cognitive Walkthrough Evaluation Table

The CW table in Figure A.4 we made when we selected a prototype within the group.

A.3 SUS responses

The SUS response forms in Figure A.5, Figure A.6, Figure A.7, Figure A.8, Figure A.9, Figure A.10, Figure A.11 and Figure A.12 are manually filled out by the participants. We will later enter the SUS responses into a website to calculate the scores, which will then

Figure A.2: *Wechat Group*

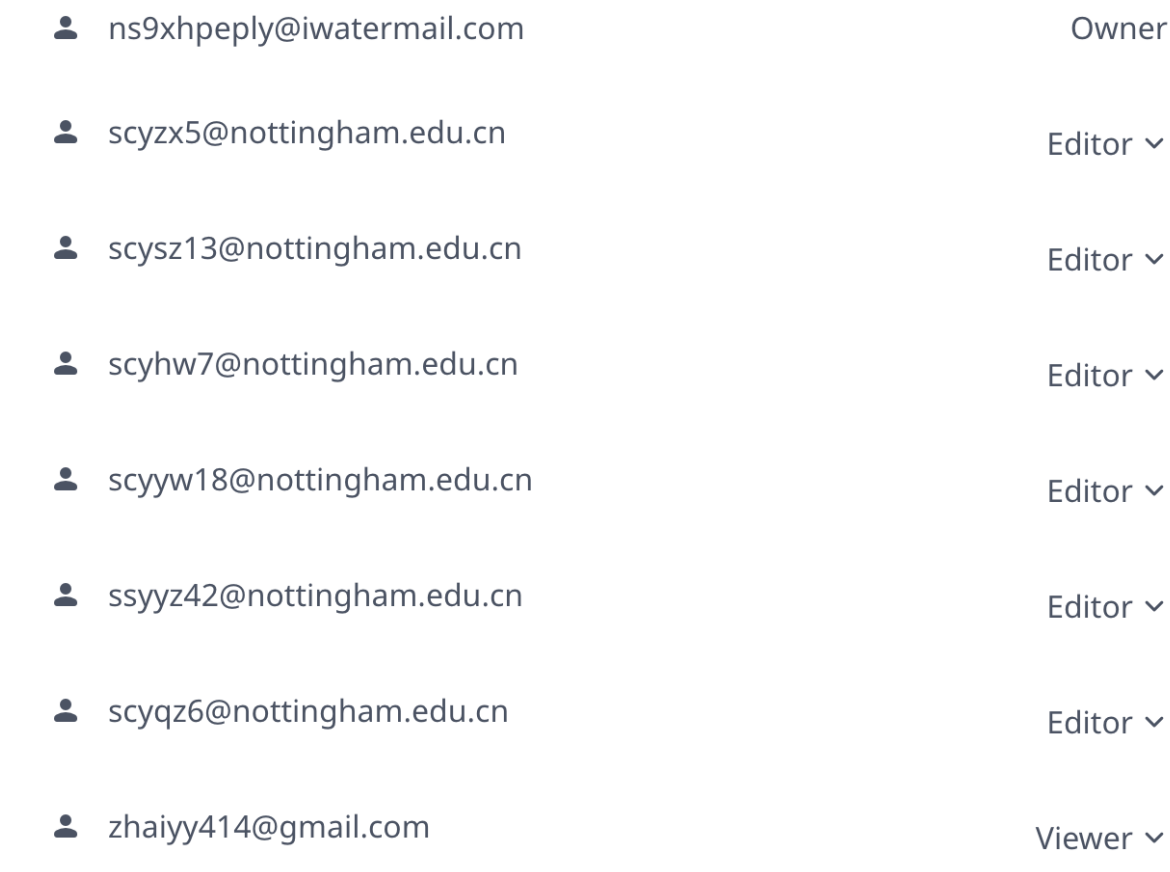


Figure A.3: *Overleaf Group*

Cognitive Walkthrough (CW) Evaluation Table

Evaluation Content and Questions

1. Task Description
   - Task Goal: Event Display and RSVP

2. Evaluation Questions
   - Can the user locate and perform the necessary actions to complete the task?
   - Are the interface elements clear and easy to understand?
   - Did the user encounter any difficulties while performing the task?
   - Was the task completed successfully without obstacles?

3. Interface and Feedback
   - Was the interface feedback timely and useful?
   - Can the user obtain sufficient information from the interface to know if their actions are correct?
   - Does the system provide clear guidance where necessary?
   - Does the user clearly know when the task is completed?

4. Expectations vs Reality
   - Are the user's expectations of the task and the actual operations aligned?
   - Does the interface design match the user's thinking patterns?

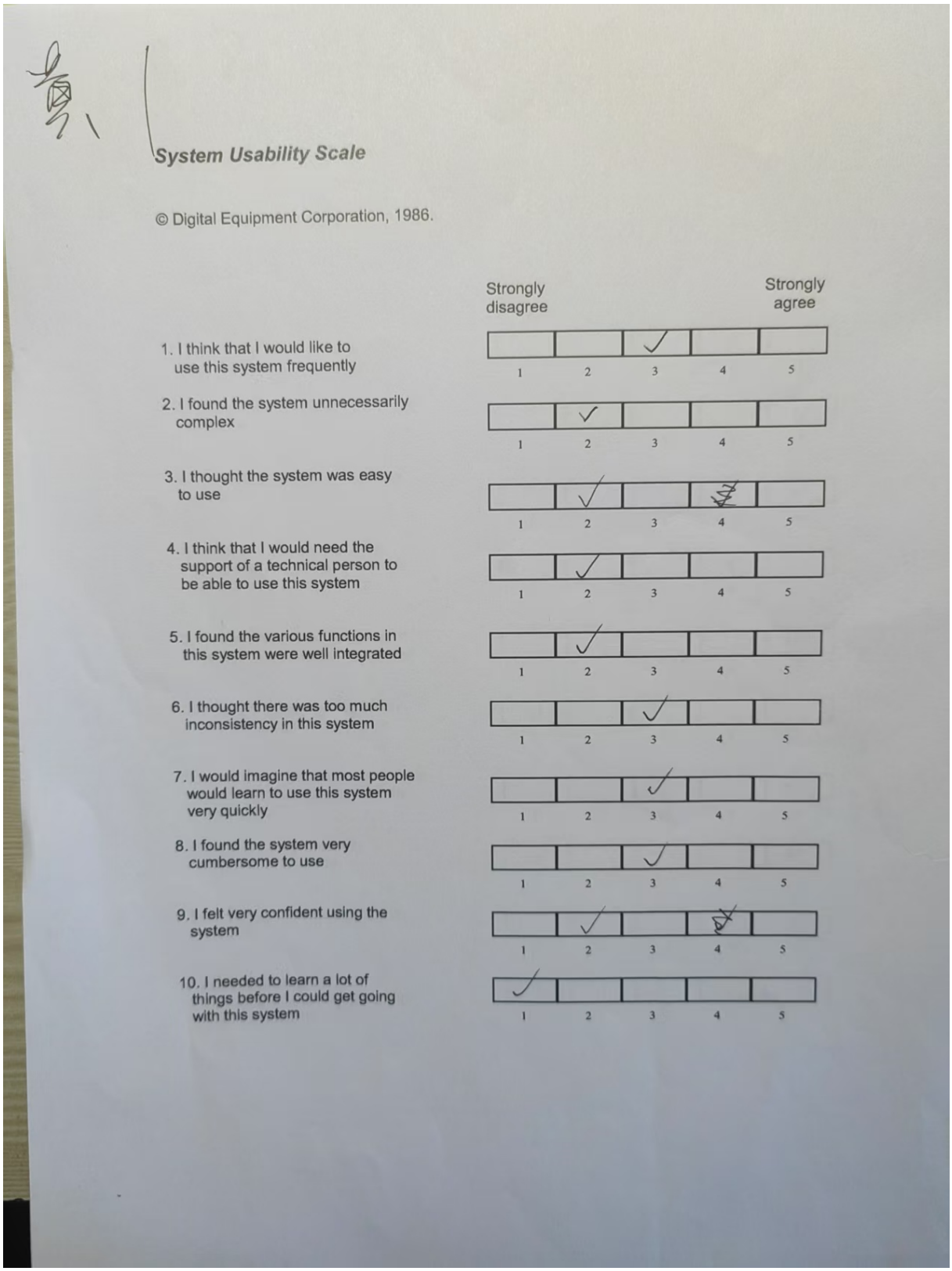
5. Factors for Task Success
   - What design elements helped the user complete the task?
   - What design elements may have hindered the task completion?

6. Summary and Recommendations
   - Overall assessment: Was the task easy and smooth to perform?
   - Improvement suggestions based on the evaluation results.

Figure A.4: *CW Question*

be presented in Figure 3.1.

4.4 Evidence of Informed Consent from Participants

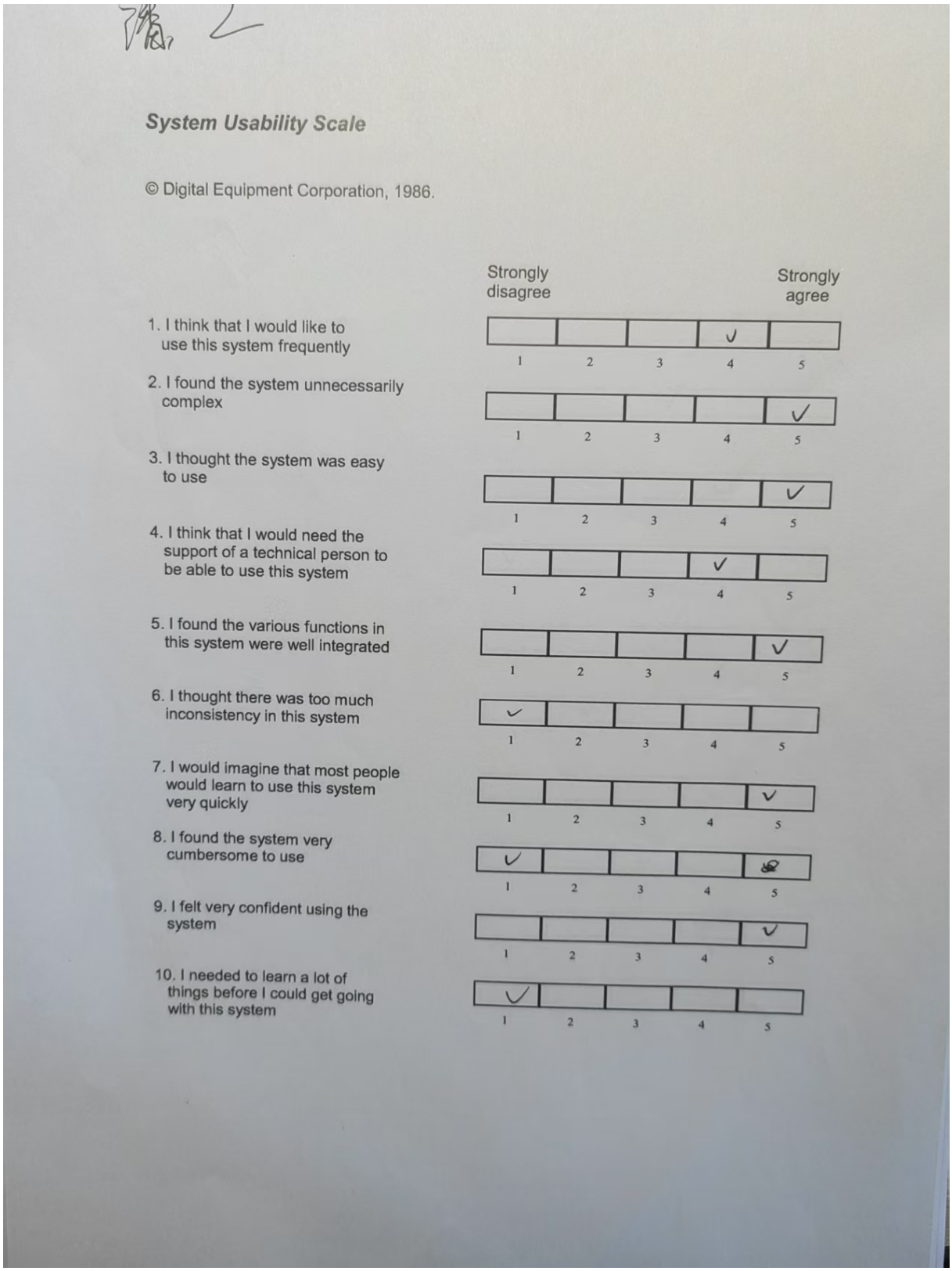


System Usability Scale

© Digital Equipment Corporation, 1986.

	Strongly disagree				Strongly agree
1. I think that I would like to use this system frequently			✓		
2. I found the system unnecessarily complex		✓			
3. I thought the system was easy to use		✓		✓	
4. I think that I would need the support of a technical person to be able to use this system		✓			
5. I found the various functions in this system were well integrated		✓			
6. I thought there was too much inconsistency in this system			✓		
7. I would imagine that most people would learn to use this system very quickly			✓		
8. I found the system very cumbersome to use			✓		
9. I felt very confident using the system		✓		✓	
10. I needed to learn a lot of things before I could get going with this system	✓				

Figure A.5: *SUS Reponse Form*



System Usability Scale

© Digital Equipment Corporation, 1986.

	Strongly disagree				Strongly agree
1. I think that I would like to use this system frequently				✓	
2. I found the system unnecessarily complex					✓
3. I thought the system was easy to use					✓
4. I think that I would need the support of a technical person to be able to use this system				✓	
5. I found the various functions in this system were well integrated					✓
6. I thought there was too much inconsistency in this system	✓				
7. I would imagine that most people would learn to use this system very quickly					✓
8. I found the system very cumbersome to use	✓				
9. I felt very confident using the system					✓
10. I needed to learn a lot of things before I could get going with this system	✓				

Figure A.6: *SUS Response Form 2*

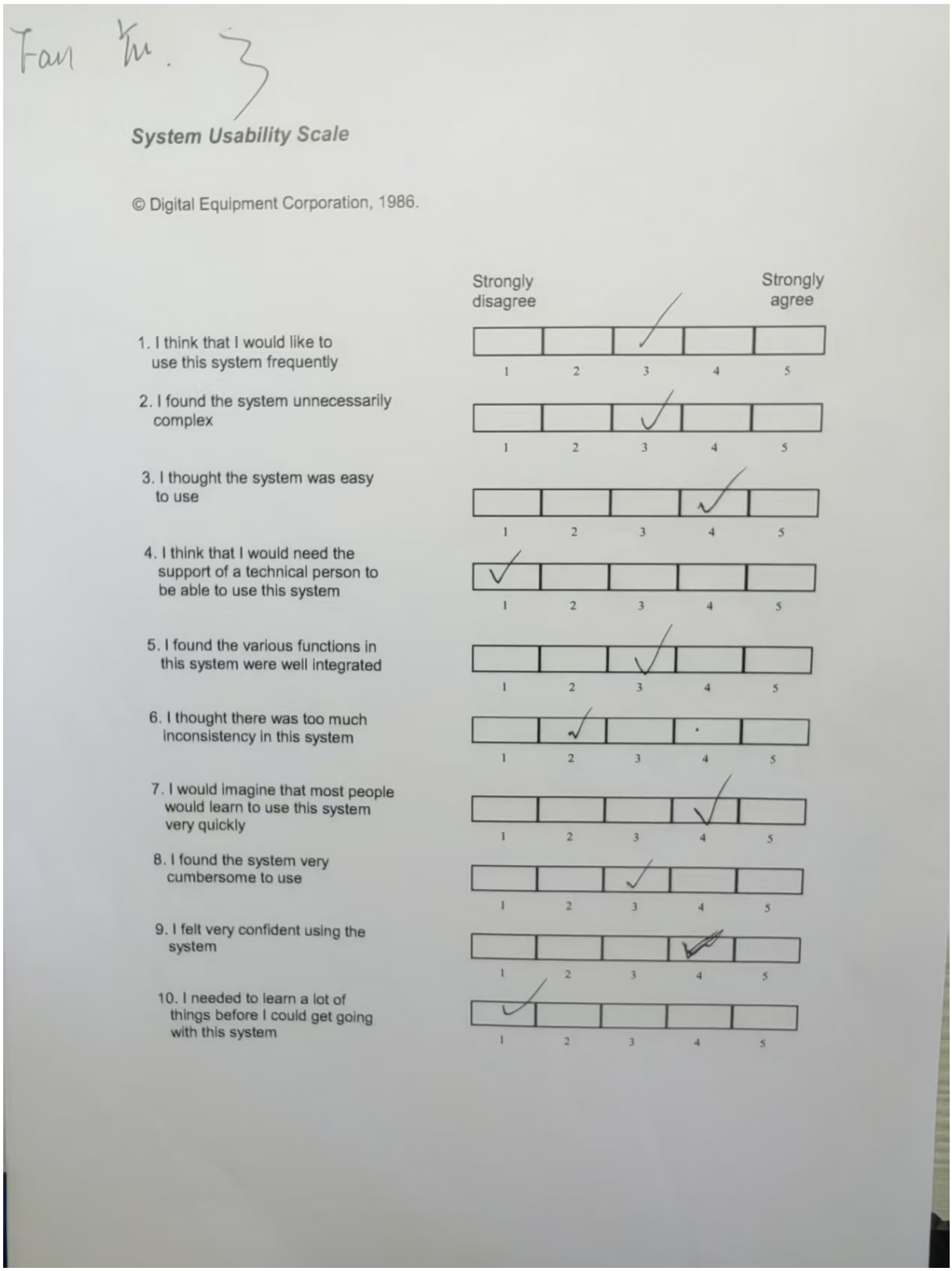
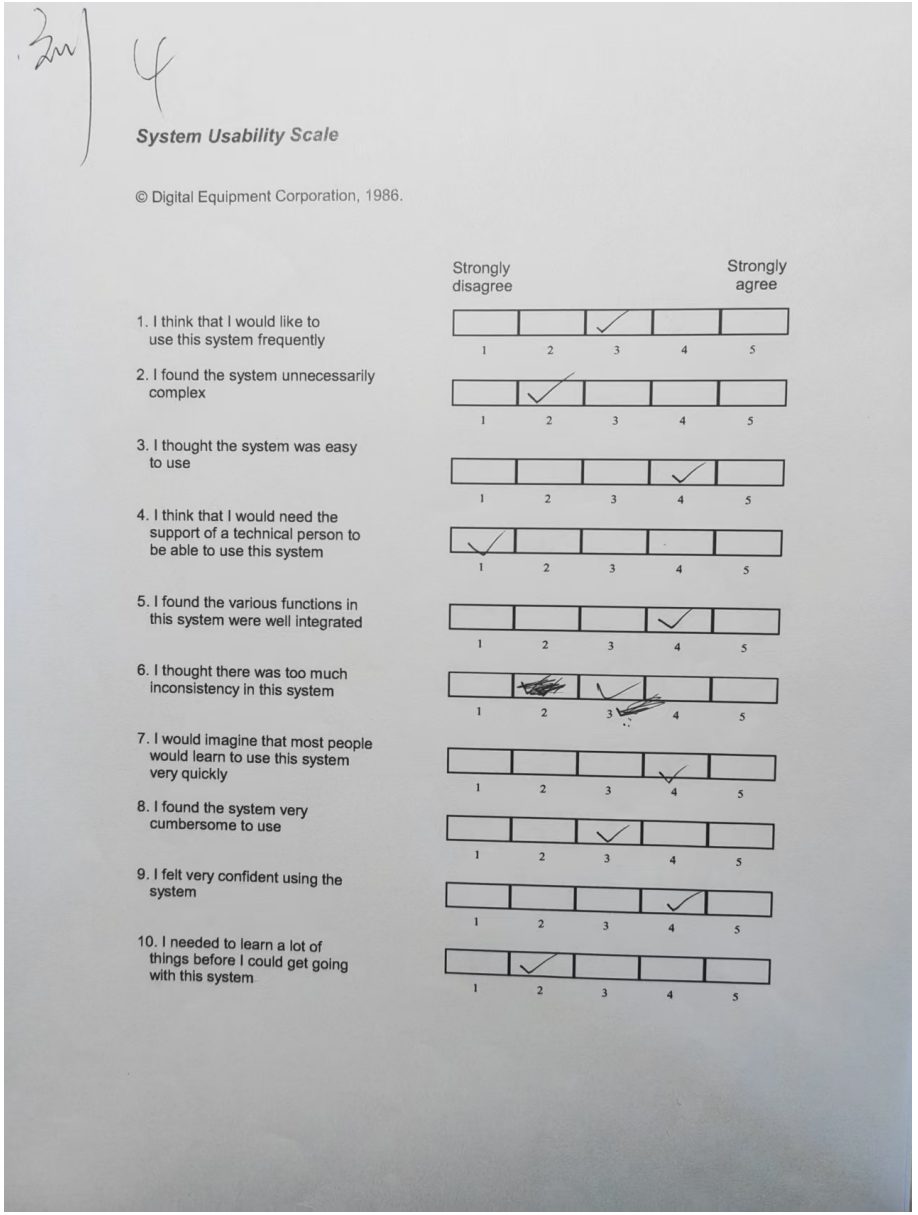


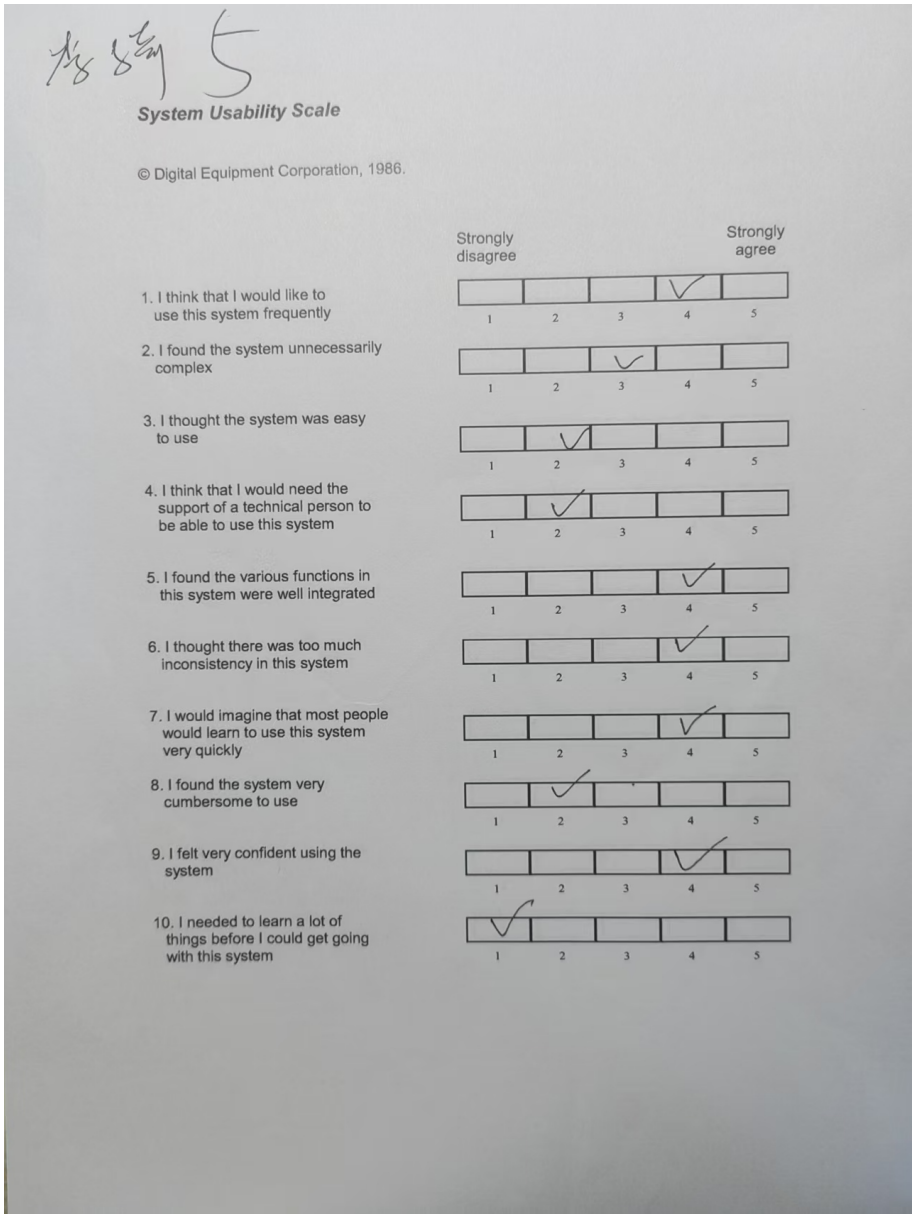
Figure A.7: *SUS Response Form 3*



System Usability Scale

	Strongly disagree				Strongly agree
1. I think that I would like to use this system frequently	1	2	✓ 3	4	5
2. I found the system unnecessarily complex	1	✓ 2	3	4	5
3. I thought the system was easy to use	1	2	3	✓ 4	5
4. I think that I would need the support of a technical person to be able to use this system	✓ 1	2	3	4	5
5. I found the various functions in this system were well integrated	1	2	3	✓ 4	5
6. I thought there was too much inconsistency in this system	1	✓ 2	✓ 3	4	5
7. I would imagine that most people would learn to use this system very quickly	1	2	3	✓ 4	5
8. I found the system very cumbersome to use	1	2	✓ 3	4	5
9. I felt very confident using the system	1	2	3	✓ 4	5
10. I needed to learn a lot of things before I could get going with this system	1	✓ 2	3	4	5

Figure A.8: *SUS Response Form 4*

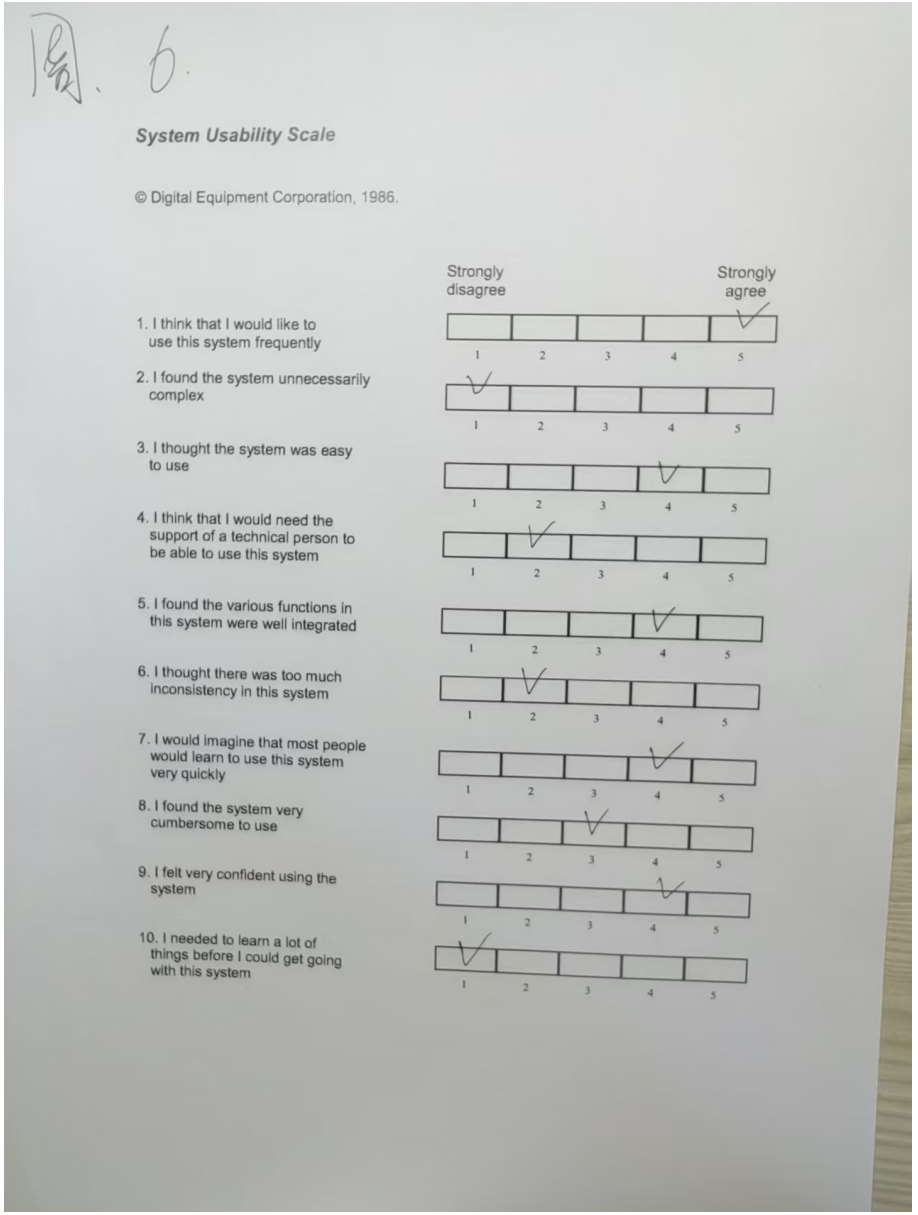


System Usability Scale

© Digital Equipment Corporation, 1986.

	Strongly disagree				Strongly agree
1. I think that I would like to use this system frequently	1	2	3	✓ 4	5
2. I found the system unnecessarily complex	1	2	✓ 3	4	5
3. I thought the system was easy to use	1	✓ 2	3	4	5
4. I think that I would need the support of a technical person to be able to use this system	1	✓ 2	3	4	5
5. I found the various functions in this system were well integrated	1	2	3	✓ 4	5
6. I thought there was too much inconsistency in this system	1	2	3	✓ 4	5
7. I would imagine that most people would learn to use this system very quickly	1	2	3	✓ 4	5
8. I found the system very cumbersome to use	1	✓ 2	3	4	5
9. I felt very confident using the system	1	2	3	✓ 4	5
10. I needed to learn a lot of things before I could get going with this system	✓ 1	2	3	4	5

Figure A.9: *SUS Response Form 5*

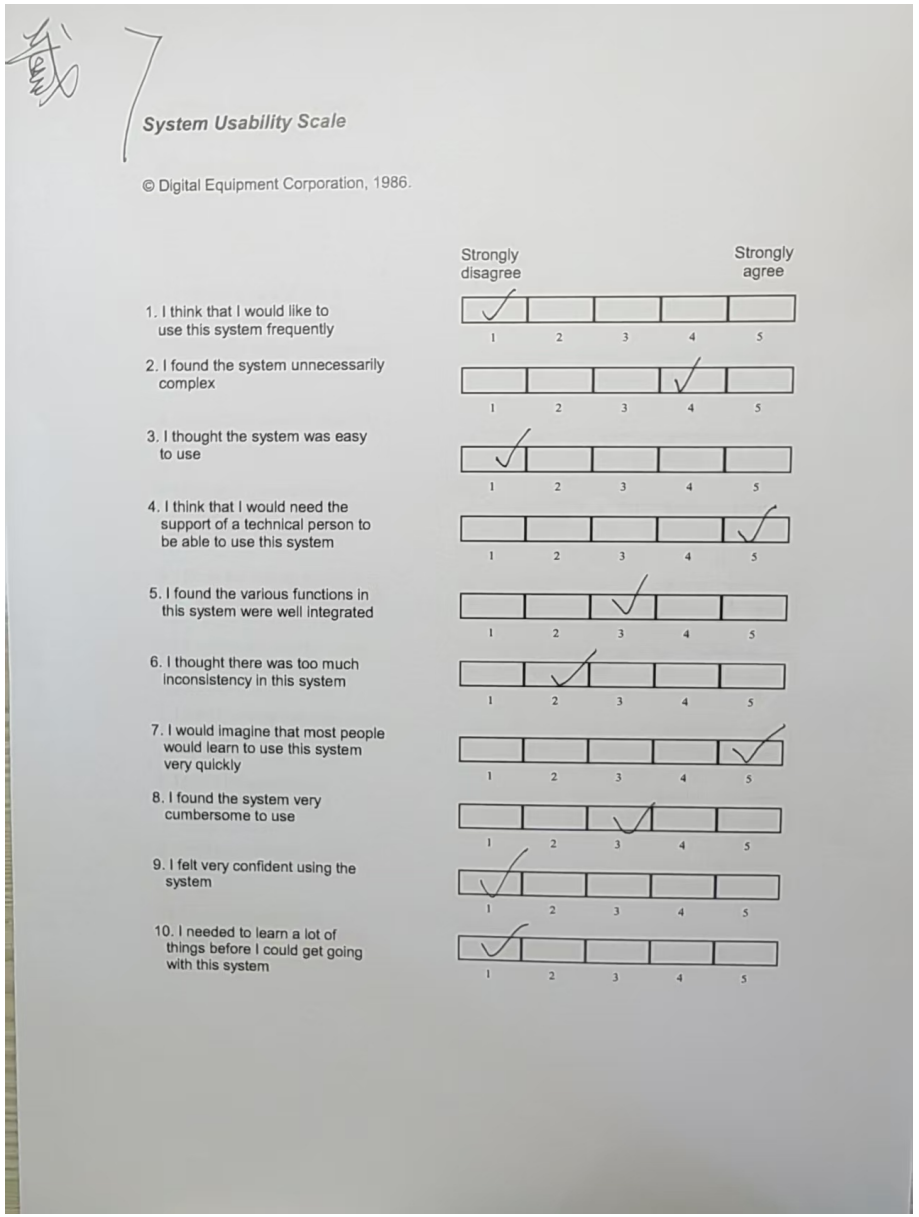


System Usability Scale

© Digital Equipment Corporation, 1986.

	Strongly disagree				Strongly agree
1. I think that I would like to use this system frequently	1	2	3	4	5 ✓
2. I found the system unnecessarily complex	1 ✓	2	3	4	5
3. I thought the system was easy to use	1	2	3	4 ✓	5
4. I think that I would need the support of a technical person to be able to use this system	1	2 ✓	3	4	5
5. I found the various functions in this system were well integrated	1	2	3	4 ✓	5
6. I thought there was too much inconsistency in this system	1	2 ✓	3	4	5
7. I would imagine that most people would learn to use this system very quickly	1	2	3	4 ✓	5
8. I found the system very cumbersome to use	1	2	3 ✓	4	5
9. I felt very confident using the system	1	2	3	4 ✓	5
10. I needed to learn a lot of things before I could get going with this system	1 ✓	2	3	4	5

Figure A.10: *SUS Response Form 6*



7

System Usability Scale

	Strongly disagree				Strongly agree
1. I think that I would like to use this system frequently	1	2	3	4	5
2. I found the system unnecessarily complex	1	2	3	4	5
3. I thought the system was easy to use	1	2	3	4	5
4. I think that I would need the support of a technical person to be able to use this system	1	2	3	4	5
5. I found the various functions in this system were well integrated	1	2	3	4	5
6. I thought there was too much inconsistency in this system	1	2	3	4	5
7. I would imagine that most people would learn to use this system very quickly	1	2	3	4	5
8. I found the system very cumbersome to use	1	2	3	4	5
9. I felt very confident using the system	1	2	3	4	5
10. I needed to learn a lot of things before I could get going with this system	1	2	3	4	5

Figure A.11: *SUS Response Form 7*

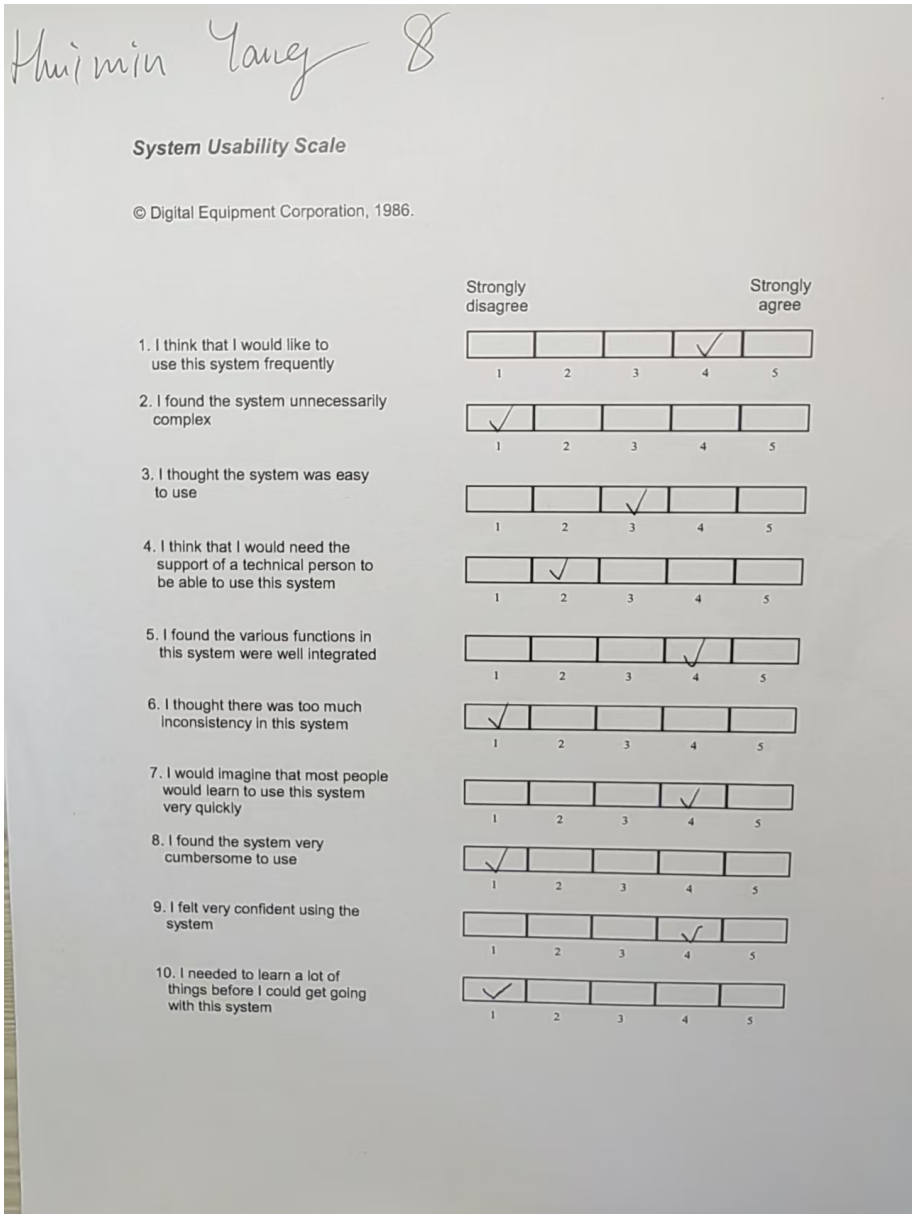


Figure A.12: *SUS Response Form 8*