

APPENDIX D · QUESTIONNAIRE 01 / 03

Participant briefing and task sheet.

JARVIS & System Brain · exploratory evaluation · blank instrument v1.1

Purpose: understand whether people can interpret the interface, locate supporting information and identify the limits of delegated decision-making. Participation is voluntary. Skip any question or stop at any time. Do not enter names, contact details, passwords or confidential business information.

SESSION RECORD	COMPLETE BEFORE THE SESSION
Participant code (optional)	_____
Date / report or prototype version	_____
Viewing condition	☐ Report screenshots ☐ Isolated interactive demo
Main role	☐ Research ☐ Engineering ☐ Operations ☐ Other
Experience using AI tools	☐ None ☐ Occasional ☐ Weekly ☐ Daily

☐ I understand the activity and choose to participate. For an organised study, the facilitator must first provide the study contact, storage arrangements, retention period and withdrawal process. This document does not collect or submit responses.

Four observation tasks

TASK	PARTICIPANT INSTRUCTION	FACILITATOR RECORD
T1· Role	Find the Operations cell and describe its responsibility.	_____
T2· Evidence	Locate a stored record and identify its source or timestamp.	_____
T3· Limits	Identify one sign that this material is demonstration data.	_____
T4· Control	Explain when the example should return to a human operator.	_____

Record: C = completed without help; H = completed with help; U = not completed; N = not observed. For screenshots, assess interpretation only. Do not record navigation speed or execution success as if the participant used a live system.

APPENDIX D · QUESTIONNAIRE 02 / 03

Your assessment of the experience.

Circle one response per statement. Choose N/A when the material did not let you judge.

Response scale · 1 Strongly disagree · 2 Disagree · 3 Neither agree nor disagree · 4 Agree · 5 Strongly agree · N/A
Unable to assess

STATEMENT	RESPONSE
01· I understood the purpose of the System Brain view.	1 2 3 4 5 N/A
02· I could distinguish the responsibilities of different cells.	1 2 3 4 5 N/A
03· I could identify where a cell's instructions are shown.	1 2 3 4 5 N/A
04· I could identify where stored information is shown.	1 2 3 4 5 N/A
05· I could distinguish sample data from measured results.	1 2 3 4 5 N/A
06· The labels made the system's current state clear.	1 2 3 4 5 N/A
07· I knew what evidence I would need before accepting a recommendation.	1 2 3 4 5 N/A
08· I understood when a human should review an action.	1 2 3 4 5 N/A
09· I could identify limits to the authority assigned to a cell.	1 2 3 4 5 N/A
10· The information was readable at my viewing size.	1 2 3 4 5 N/A
11· I would feel able to question a recommendation.	1 2 3 4 5 N/A
12· The material helped me identify a suitable task for a controlled trial.	1 2 3 4 5 N/A

INTERPRETATION

This is an original exploratory questionnaire, not a validated psychological scale. Responses describe perceptions of the material shown. They do not measure intelligence, decision accuracy, safety or equivalence to human specialists.

APPENDIX D · QUESTIONNAIRE 03 / 03

Decision boundaries and feedback.

Participant code (optional): _____

For each example, select the highest level you would currently permit based on the evidence available to you. A = human only; B = AI draft with human approval; C = bounded automatic action; D = insufficient evidence.

EXAMPLE TASK	YOUR CHOICE
Summarise a non-sensitive internal document.	A B C D
Route a support request to a department.	A B C D
Recommend a stock transfer from an old record.	A B C D
Approve a financial commitment or change access rights.	A B C D

What was clear, useful or difficult?

What evidence or control would you need before relying on it?

Which routine task would you test first, and how would you judge success?

What should change before a wider trial?

Researcher analysis plan: report sample size, recruitment, viewing condition and missing answers. Summarise item distributions and task outcomes separately; compare screenshots and interactive sessions separately. Code free-text themes with a documented rubric. Do not invent a combined “trust” or “intelligence” score. No participant responses have been collected for this edition.