Think-Aloud Evaluation – Observer Notes

Purpose of this document:
This document is used by the research team to record observations during the think-aloud evaluation session. Observers should not intervene during the tasks, except to encourage participants to continue thinking aloud if they fall silent.

Session Information

Date:
Participant ID: 2
Evaluator(s):
Persona role played by participant: Environmental Scientist

Observer Instructions

- Do not explain the interface or tasks.
- Let the participant speak as much as possible.
- Only prompt with neutral reminders such as: 'Please keep thinking aloud.'
- Record notable quotes verbatim where possible.

Task 1 — Explore global NbS distribution

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas
Looked at heatmap	I'm a bit confused Colours mean the concentration of projects	The three maps being present are confusing - choose 1		Base on the current atlas - number of projects Legend of what the colours mean (e.g. red = >100 projects)

Task 2 — Narrow the scope of analysis

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas
		What do the values mean - scale (SP matrix makes more sense)		Instead of scatter plot, use a line graph
		Scatter plot could be overwhelming		
		Matrix with diagrams is very confusing		

Task 3 — Compare NbS across multiple criteria

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Task 4 — Identify promising solutions

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Task 5 — Drill down into a specific project

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Task 6 — Explain outcomes using governance

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Task 7 — Reflect on evidence and data quality

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Task 8 — Draw a research-oriented conclusion

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Right now confused, but can see the
Start with abstract; start with more simple ones, then become more detailed

Think-Aloud Evaluation – Observer Notes

Purpose of this document:
This document is used by the research team to record observations during the think-aloud evaluation session. Observers should not intervene during the tasks, except to encourage participants to continue thinking aloud if they fall silent.

Session Information

Date:
Participant ID: Michael
Evaluator(s):
Persona role played by participant: Environmental Scientist

Observer Instructions

- Do not explain the interface or tasks.
- Let the participant speak as much as possible.
- Only prompt with neutral reminders such as: 'Please keep thinking aloud.'
- Record notable quotes verbatim where possible.

Task 1 — Explore global NbS distribution

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas
Thinks slide 1 is about comparisons	Implicit colour encoding - don't like red	Interpret size as no. of success factors		Have different colours based on projects
	Dots ones shows what you want more			How much they spend vs how successful they are - not on a world map
				Switch metrics (colour coding based on the different goals)
				Relative change (amount spent

vs how rich the
country is)

Task 2 — Narrow the scope of analysis

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas
			Likes the zoom in feature	

Task 3 — Compare NbS across multiple criteria

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas
			Likes the scatterplot matrix	Hover over one category and can see all the
				Only half the matrix is needed
				Split the matrix into negative and positive -> red ones and green ones

Task 4 — Identify promising solutions

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Task 5 — Drill down into a specific project

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Task 6 — Explain outcomes using governance

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Task 7 — Reflect on evidence and data quality

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Task 8 — Draw a research-oriented conclusion

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Have the matrix as a comparison - so you click on an aspect with the matrix and can see the different results

Think-Aloud Evaluation – Observer Notes

Purpose of this document:
This document is used by the research team to record observations during the think-aloud evaluation session. Observers should not intervene during the tasks, except to encourage participants to continue thinking aloud if they fall silent.

Session Information

Date: 12/12/2025
Participant ID: 8
Evaluator(s):
Persona role played by participant: Environmental Scientist

Observer Instructions

- Do not explain the interface or tasks.
- Let the participant speak as much as possible.
- Only prompt with neutral reminders such as: 'Please keep thinking aloud.'
- Record notable quotes verbatim where possible.

Task 1 — Explore global NbS distribution

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas
Assumes red is more prevalent	Density maps, bigger dots = more projects			

| areas (bottom left)

Larger circle is more projects | More projects in northern hemisphere | | | |

Task 2 — Narrow the scope of analysis

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas
Chose Australia (currently researching in - similar climate)		Was confused by the task itself		

Task 3 — Compare NbS across multiple criteria

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas
Assuming cost is a good thing	I assume that this is more of a correlation rather than impact on cooling, I am getting a lot of biodiversity Table is mirrored I am assuming that cost is a good thing… a	Didn't seem as confident with the matrices Don't see how cost comes into the matrix - can't tell if it is high cost/low cost etc Table is mirrored		Get rid of mirrored effect of matrix

		lot of keys in the table		
		I don't really see where I need to go into a specific project on this slide		

Task 4 — Identify promising solutions

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas
Essentially skipped		Didn't know the initiating actor		
		No legend as to the outcome category		

Task 5 — Drill down into a specific project

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Task 6 — Explain outcomes using governance

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Task 7 — Reflect on evidence and data quality

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Task 8 — Draw a research-oriented conclusion

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Think-Aloud Evaluation – Observer Notes

Purpose of this document:
This document is used by the research team to record observations during the think-aloud evaluation session. Observers should not intervene during the tasks, except to encourage participants to continue thinking aloud if they fall silent.

Observer Instructions

- Do not explain the interface or tasks.
- Let the participant speak as much as possible.
- Only prompt with neutral reminders such as: 'Please keep thinking aloud.'
- Record notable quotes verbatim where possible.

Task 1 — Explore global NbS distribution

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Task 2 — Narrow the scope of analysis

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Task 3 — Compare NbS across multiple criteria

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Task 4 — Identify promising solutions

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Task 5 — Drill down into a specific project

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Task 6 — Explain outcomes using governance

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Task 7 — Reflect on evidence and data quality

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Task 8 — Draw a research-oriented conclusion

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

He felt that the different diagrams could be combined - liked the left matrix; did not like the right matrix

Think-Aloud Evaluation – Observer Notes

Purpose of this document:
This document is used by the research team to record observations during the think-aloud evaluation session. Observers should not intervene during the tasks, except to encourage participants to continue thinking aloud if they fall silent.

Session Information

Date:
Participant ID: 4
Evaluator(s):
Persona role played by participant: Environmental Scientist

Observer Instructions

- Do not explain the interface or tasks.
- Let the participant speak as much as possible.
- Only prompt with neutral reminders such as: 'Please keep thinking aloud.'
- Record notable quotes verbatim where possible.

Task 1 — Explore global NbS distribution

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas
	White		Colour coded clearly	Heatmap is most effective but might not have enough projects to make it worthwhile

Task 2 — Narrow the scope of analysis

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas
			All pretty obvious	
			Communicates it all very well	

Task 3 — Compare NbS across multiple criteria

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas
	We talked about	Don't instantly know what I am looking at		Write out the abbreviations fully (e.g. c = cost)
		Nothing about the abbreviations is clear		
		Takes a few thinking steps		

in order to
understand it

Task 4 — Identify promising solutions

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas
Took a long time to read	Would look at [top left] as they show different types	Getting back to me not knowing what I am looking at		
	Start counting per column in the cost region			

Task 5 — Drill down into a specific project

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas
		Graph isn't clicking with me		

Task 6 — Explain outcomes using governance

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas
Took a while to look and understand the graph	Different governance models = different colours

At some point the governance type and NbS type are combined but I am not exactly sure how | I would trace the lines to answer the question | | |

Task 7 — Reflect on evidence and data quality

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

Task 8 — Draw a research-oriented conclusion

Observed behaviour / actions	Verbalised thoughts (quotes)	Confusion / difficulties	Positive feedback	Improvement ideas

1. Overall Impression Across Groups

- Many visualizations were perceived as overwhelming due to:

  - Too many visual elements

  - Too many categories shown at once

  - Lack of a clear starting point for interpretation

- Users often felt lost when:

  - The interaction flow was unclear

  - The relationship between categories or variables was not well explained

- Several groups struggled with:

  - Interpreting colors and encodings

  - Understanding what values actually represent (counts, impact, cost, success, etc.)

2. Data & Value Clarity Issues

Common Problems

- Exact numbers or quantities were often unclear

  - Users wanted to know *how many projects*, not just relative patterns

- Counting across multiple columns or categories was confusing

- The meaning of values (e.g. cost, impact, outcome) was not always explained

Examples from Feedback

- "The amount of projects is not very clear; better to know the specific number."

- "Not sure what I am looking at; the combination is not clear."

- "Better have the number of different types of dots."

- "Less visible of the specific projects."

3. Color & Visual Encoding Confusion

Common Problems

- Color meanings were often implicit or ambiguous

- Users did not always understand:

    - What red vs blue means

    - Whether red indicates "high", "negative", or "important"

- Some color schemes suggested unintended interpretations

Examples

- "Don't understand the meaning of the color."

- "Red is interpreted as negative or unsuccessful."

- "Implicit color coding without explanation."

- "Blue takes too much space and dominates the view."

4. Chart Type & Representation Feedback

Heatmaps

- Frequently mentioned as effective and intuitive

- Users liked:

    - Clear overview

    - Easy identification of high / low areas

- However:

    - Interpretation still depends on clear legends and explanations

Examples:

- "Heatmap looks best."

- "Focusing on heatmap: red is highest, yellow lowest."

Scatter Plots & Line Charts

- Scatter plots were sometimes preferred over:

    - Too many discrete categories

    - Overloaded bar or type-based charts

- Lines were suggested when trends over time or comparison were important

Examples:

- "Scatter plot looks better than different types."

- "Better to use lines; too many plots are overwhelming."

Matrix

- Mixed reactions:

    - Some users liked matrix views

    - Others found them difficult to read

Key issues:

- Full matrices were considered too complex

- Highlighting or filtering was suggested

- Splitting positive vs negative outcomes could improve clarity

Examples:

- "Matrix looks good, but full matrix is not necessary."

- "Matrix is difficult to understand; split positives and negatives."

5. Interaction & User Flow Issues

Common Problems

- Users were unsure:

    - Where to start

    - What to click next

    - What effect their interaction had

Examples:

- "Clicking on map → city → project; user is lost."

- "The dashboard looks clear but takes a long time to read."

- "The starting point is not clear."

Suggested Improvements

- Clear interaction cues

- More guided flow

- Progressive disclosure instead of showing everything at once

6. Categories, Outcomes & Interpretation Problems

Issues

- Users struggled to understand:

    - Differences between categories

    - Why certain outcomes differ

    - What "impact" or "success" actually means

Examples:

- "Can't understand the difference between categories."

- "The aspect and outcome are not very clear."

- "What defines success? Multiple metrics are mixed."

From Michael

Key Points

- Size encoding was interpreted as "successfulness"

- Red color was interpreted as negative

- Success was defined across multiple metrics, which caused confusion

- Users wanted clearer answers to:

    - What does this graph actually mean?

    - What are companies spending vs achieving?

    - Are low-cost but high-impact projects visible?

Suggestions

- Make comparisons relative, not absolute

- Highlight projects with:

    - Lower cost but strong impact

- Use the map as a powerful comparison tool

- Reduce matrix complexity:

    - Highlight selected categories

    - Avoid showing full matrix by default