

Activity 1 – Initial Impressions (15 mins)

- 1. What insights are the researchers hoping to gain by collecting this data? What research questions are they hoping to answer? (In journal 1-2 sentences)**

Researchers aim to see whether LinguaLift improves users' language proficiency and how participants perceive the software (ease of use, engagement, satisfaction).

The core questions are about effectiveness, usability, engagement features, and user experience.

- 2. Participant 13's data differs from the others in the data set.**

- a) What do you think could have happened and how might this affect the overall data analysis? (In journal 1-2 sentences)**

Participant 13 gave extremely low survey ratings despite showing some learning improvement.

This outlier may skew averages and misrepresent overall satisfaction trends.

- b) What strategies could you use to handle outliers in this data set? (In journal 1-2 sentences)**

Strategies include examining context for data entry error, reporting results with and without the outlier, or using median and robust statistical methods to reduce distortion.

Activity 2 – Quantitative Data (30 mins)

- 3. What are some simple ways researchers might summarize the survey results to present to a team of software developers? (In journal 1-2 sentences)**

Researchers could summarize mean and median scores for each survey question and show the frequency distribution of ratings.

Simple bar charts or summary tables would give developers a quick overview of user satisfaction and key issues.

- 4. Using Datawrapper (<https://www.datawrapper.de/>), Excel, Python (matplotlib.pyplot) or any other software you are familiar with, create a plot that shows**

- a) Participant satisfaction with the app**

- b) How participant satisfaction is related to learning outcomes**

- c) The relationship between actual learning outcome and perceived learning outcomes of the participants**

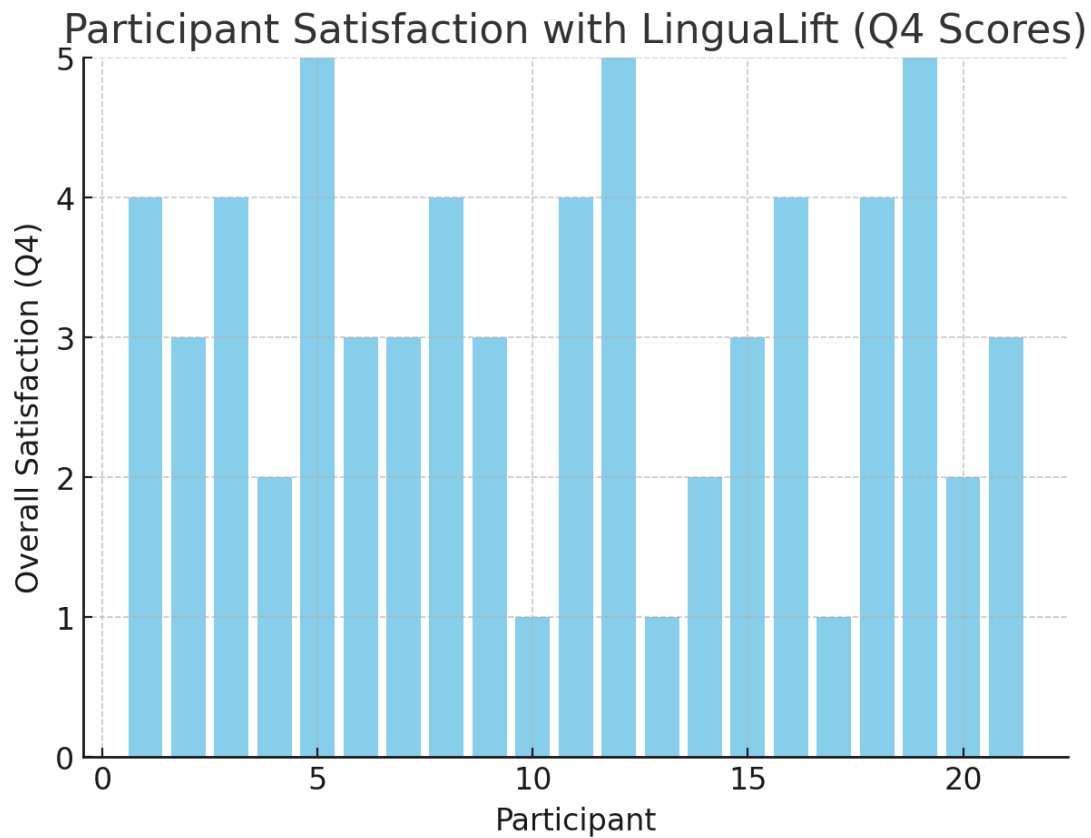
a) Bar chart was used because it clearly shows satisfaction levels across participants. Most participants reported moderate to high satisfaction (3–5).

b) Scatter plot was used to show correlation. Higher test score gains generally align with

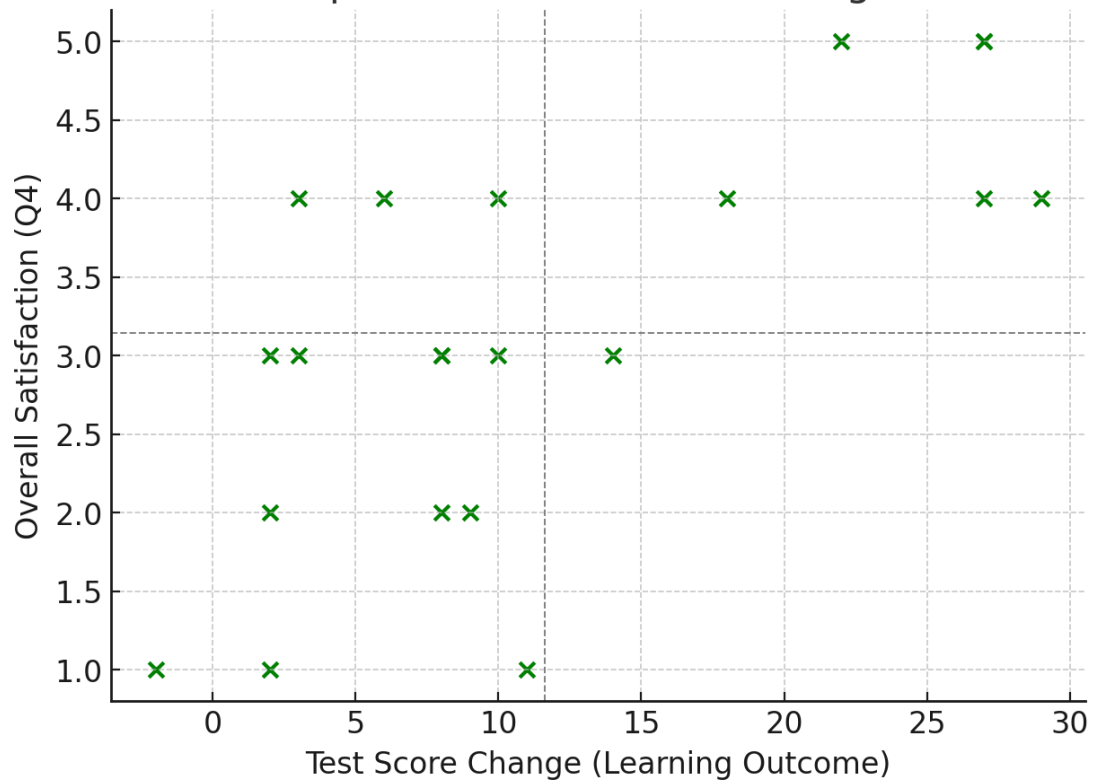
higher satisfaction, though some exceptions exist.

c) Scatter plot was used to compare actual vs. perceived outcomes. In general, perceived effectiveness follows actual improvement, but a few participants misjudge their progress.

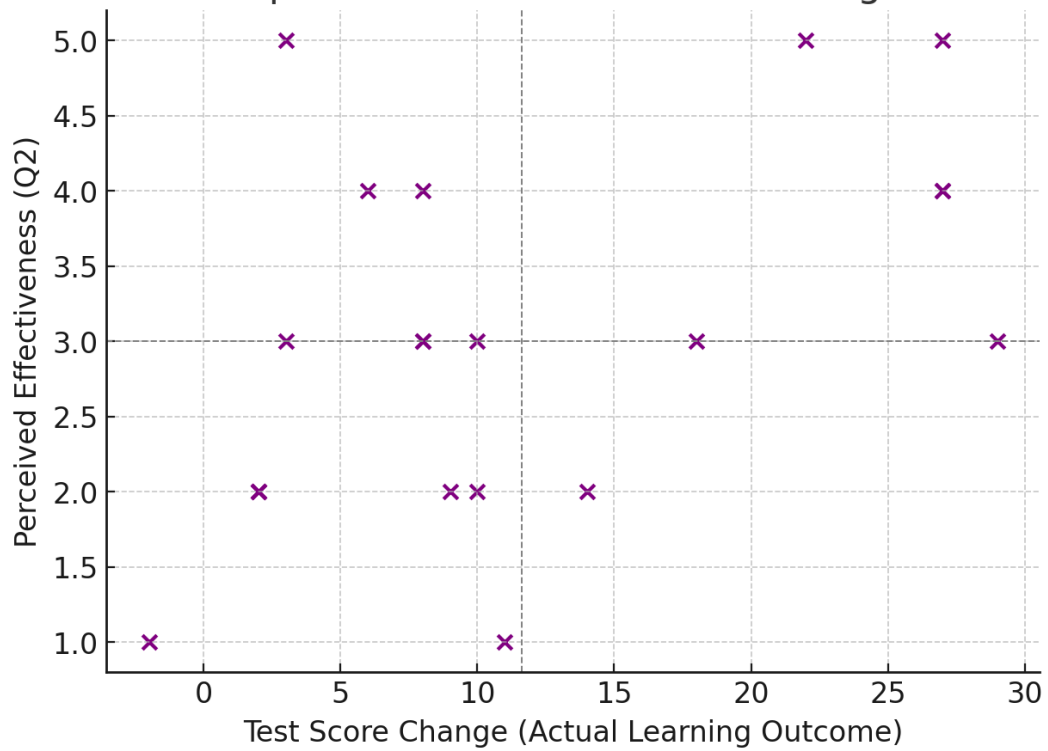
Paste an image of your plots in your journals. Include a one sentence justification for the visualisation/plot used and one sentence describing the trend in the data revealed by the plot.



Relationship: Satisfaction vs. Learning Outcomes



Relationship: Actual vs. Perceived Learning Outcomes



Activity 3 – Qualitative Data (30 mins)

5. The research team has identified the following codes from the qualitative data:

- 1) Usability Issues**
- 2) Technical Issues**
- 3) Engagement Features**
- 4) Learning Effectiveness**
- 5) AI Interaction**
- 6) Content Limitations and Diversity**

In your groups, determine how you would code each of the participant comments (i.e., which code does each comment relate to?)

6. One theme that could be developed from the qualitative data is “User Experience”. The User Experience theme encompasses both the issues that hinder experience as well as features that enhance it, including codes such as Usability Issues, Technical Issues, AI Interaction and Engagement Features.

Discuss in your groups other themes that could be developed from the data. (In journal describe the theme and codes it relates to in 2-3 sentences).