The Effect of Gridlines on Comprehension of Trends in a Data Visualization

Thejas Raju Ravi Govind Raju

1 Proposed Hypothesis

Question: Does the presence of gridlines help in a better comprehension of the trend-line present in the data visualization?

Hypothesis: If gridlines are added to the data visualization that contains a trend-line, it leads to a better comprehension of the trends being presented in the chart.

My experiment would investigate the importance of using gridlines when representing a data visualization in which some trends exist that are depicted using one or more trend-line(s). The core aspects related to the comprehension of trend-lines and the importance of gridlines would be discussed. A reader's ability to conceive the angle of a trend-line is limited. This can be observed especially while comparing charts whose trend-lines have a very minute slope difference. *The minute difference may be difficult to interpret but if gridlines are added, the reader will have a point of reference and can make more accurate interpretations of the data visualization.* Gridlines in a chart may be in the form of axes ticks, major horizontal and vertical lines, and minor horizontal and vertical lines. The experiment would establish if the usage of gridlines would help in interpreting the trends present in a data visualization.

The following are some related work on the usage and effect of trend-lines and gridines on the interpretation of data visualizations:

Milo Schield presented his findings *'Three Kinds of Statistical Literacy: What Should We teach?' (December 31, 2001)*. This study reports a 1985 survey conducted by Peter Holmes on *'The Statistical Needs of Non-Specialist Young Workers'* (Fig. 1a), according to which, only 19% of the employees surveyed draw trend lines. This shows that the usage of trend-lines is relatively low when compared to the other aspects of statistics to represent the data. Some work to improve the accuracy of interpretation of trend-lines may encourage the usage of trend-lines for visualization and this is what my experiment aims to do

60% draw up tables of data	19% draw trend lines
54% read and interpret tables of data	19% read and interpret histograms
53% assess the accuracy of someone else's data	17% calculate median and quartiles
52% decide what data to collect	*17% assign probabilities to events*
51% calculate the mean	15% allow for non-response to questionnaires
40% detect and estimate trends	14% select the questions on questionnaires
38% simplify tabulated data	*13% use statistical tests to compare sets of data*
37% allow for variability in data	*13% use probability to measure uncertainty*
37% make decisions using data	12% read and interpret results of simulations
35% make projections	9% calculate correlation coefficients
27% draw bar charts and time series graphs	8% calculate moving averages
20% use words such as likely and uncertain	*6% use a statistical test of significance*
19% calculate variance or standard deviation	4% use the normal distribution
19% use logarithm or other specialist scales	2% calculate index numbers

Fig. 1a

In his ground-breaking book *'Visual Display of Quantitative Information (1983)'*, Edward Tufte explained various methods of effectively visualizing data by depicting inaccuracies in the examples of data visualizations which were published by other companies and organizations and showing the improvements that can be made. In one such example, it can be seen that the chart depicting *'Import and Export of England to and from all of North America from 1770 to 1782'* (Fig. 1b) can be seen populated with gridlines whereas the chart depicting *'Exports and Imports to and from Denmark and Norway from 1700 to 1780'* (Fig. 1c) consists of lesser gridlines and appears to be more clear and elegant than the previous chart. The book suggests that the Data-Ink ratio which refers to ratio of data represented to the ink used must be maximized for efficiency. The graph (Fig. 1d) consists of only 10-20 percent non-data-ink which makes up the grid ticks and frames. However, through the chart depicting *'Relationship of Actual Rates of Prediction to Actual Rates (104 cities 1980)'* (Fig. 1e), it can be observed that the ink is devoted to non-data features such as the gridlines overwhelm the data points being visualized.

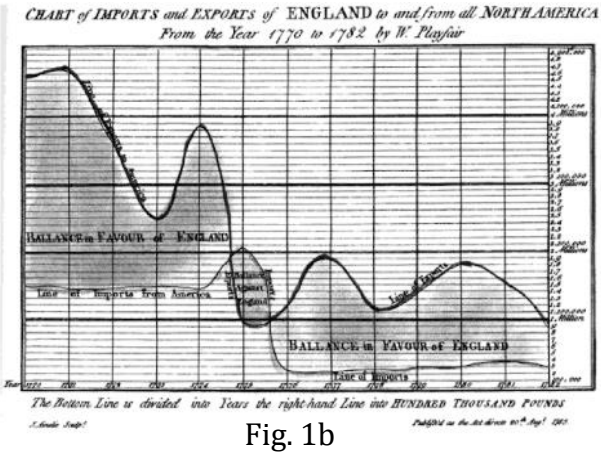


Fig. 1b

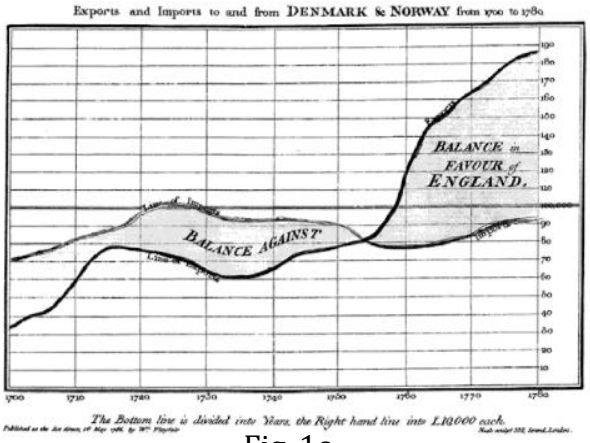


Fig. 1c

Fig. 1d

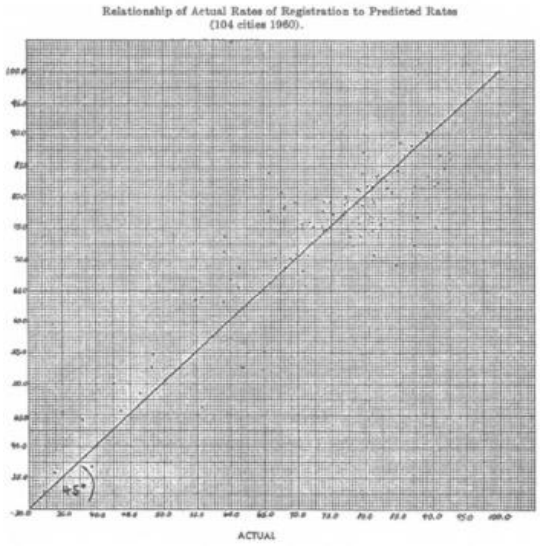


Fig. 1e

In the paper *'The Shape Parameter of a Two-Variable Graph'* by Cleveland et al. (June 1988), a hypothesis about orientation resolution and judgement of slope ratio is presented, and an experiment about slope perception is carried out. This paper proposed that the average slope in a line chart should be 45 degrees. According to an article titled *'Aspect Ratio and Banging to 45 Degrees'* by Robert Kosara, the paper was specific to comparing slopes between two lines and the slope is the average between those two lines.

2 Experimental Method

2.1 Overview

The overall experiment design can be classified into two categories, namely, within-group and between-group design. The within-group design can be performed with a small pool of participants where all the participants are exposed to every condition present in the experiment. In the between-group design, two or more groups are tested with different factors that are present in the experiment. This design requires more participants than the within-group design. The within-group design reduces the risk of errors whereas the differences present in the between-group design can increase the errors obtained. Hence, the within-group design is adopted in my experiment.

-**Dependent Variable:** The accuracy with which the data visualizations are comprehended.
-**Independent Variable:** The grid-lines that are used to serve as a reference for interpreting trends present in the data visualization. These variables have a direct effect on dependent variables.
-**Confounding Variable:** They are known as 'extra variables' or 'third variables' and must be taken into account to obtain maximum accuracy in an experiment. If they are not considered, they may suggest the presence of a correlation when there isn't and may also introduce a bias in the experiment. Hence, it is important to resolve the confounding variables during experiment design. The confounding variables in this experiment can be:

i) Subject's eye-sight
ii) Ambient lighting of the room
iii) Age
iv) Gender
v) Environment conditions i.e. without distraction to ensure concentration.

The experiment will be run by choosing 50 participants who will be asked to answer a questionnaire which will help me draw conclusions about the hypothesis. The participants chosen make up the adult population. A group of subjects with similar background and characteristics are shortlisted. These shortlisted candidates are then chosen in a random sampling method to downplay the effect of confounding variables. While choosing the subjects, it is ensured that the selection is restricted to subjects with no eye-sight related issues. By requesting the participants to disclose beforehand any eye-sight problems, a method of restriction can be applied such that only subjects who do not have eye-sight related problems can be selected. The experiment will be performed in a controlled environment with minimal noise and visual distractions. It is ensured that the room where the experiment is being performed has adequate

ambient lighting in all areas and the print quality of the questionnaire upholds standards. The answers obtained will be analysed by comparing them to the ground truth that is established by me beforehand. The majority voting or the option chosen by subjects is used to check the accuracy of the experiment.

2.2 Data collection

Data will be collected in terms of accuracy, time taken to answer, and objective measurement. The accuracy is established by comparing the participant's answers in the questionnaire and the ground truth that is established beforehand. The time taken to answer each part of the questionnaire is calculated where one part consists of questions without gridlines and the other part consists of questions with gridlines. This will help gain an understanding about the speed at which participants could answer with or without the help of gridlines. The accuracy values are calculated by the means of majority voting which will report the votes that coincide with ground-truth. Objective measurement establishes how accurately the questions were answered irrespective of what was experienced while performing the task. However, subjective measurement takes into account what the participant experienced while answering the questionnaire. In the experiment, only objective measurements are considered as questions such as comprehending the trends present in data visualizations will be asked and the time taken to interpret the data is being recorded.

The data collected will help me understand as to how accurately people interpret the presence of trends in a chart and while comparing charts with similar trends, if gridlines help in interpreting trends in a more effective way. The time taken to answer the questionnaire will help to analyse if gridlines make it easier or difficult to interpret trends in a data visualisation. The hypothesis will be substantiated by performing the experiment and recording the accuracy obtained in answering questions and the speed at which they were answered.

2.3 Selected subjects

The subjects that take part in the experiment will be chosen such that they make up all possible age groups, genders, socio-economic status and education-level. The University consists of a population that are of ranging age groups and gender. The students from undergraduate to PhD level as well as professors and other staff would provide a holistic view of the whole population. Hence, the subjects would be chosen from the people at UCD itself.

The subjects will be sourced by sending out e-mails to students, professors, and service staff who are a part of UCD. There will also be pamphlets handed out to the people at university to take part in the experiment. Flyers will be put up at common-places such as library, cafeteria and the sports complex. This will ensure maximum awareness and participation in the experiment.

The aim of obtaining 50 subjects to take part in the experiment where male and females make up equal proportion of the sample. To ensure that the sample is representative, the probabilistic sampling technique is adopted. The fifty subjects are chosen from a pool of 500 applicants in a manner such that each subject has an equal and fair chance of being chosen for the final experiment. In order to enforce this standard, a random sampling approach is adopted where fifty subjects are chosen at

random from the pool of 500 participants. The participants are labelled according to their gender and a table of random numbers is used to select the fifty participants who will eventually be the subjects of the experiment.

A representative sample refers to a sample of subjects that closely resemble if not all most of the characteristics of the population as a whole. The sample obtained using the above method is considered to be representative as the fifty people are selected such that they represent all the age groups, genders, have different socio-economic status and education-level. This will help in generalizing the experimental results to the population as a whole.

2.4 Data analysis

The results of the experiment will be analysed to establish the following:

i) Accuracy of answering the questions by comparing the participant's response to the ground truth that was established.
ii) Time taken to answer each set of visualizations.

As the questionnaire consists of questions with only one set of correct answers, the responses of the participants can be evaluated against this set of correct answers to establish the accuracy with which the questions were answered. The average accuracy is calculated by dividing the sum each response's accuracy with the total number of questions.

The experiment will contain two questionnaires, where one set contains questions with no gridlines, and the second set consists of the same questions but accompanied by gridlines. The time taken to answer each questionnaire by each participant will aid in establishing which set took a minimum time to answer.

The accuracy and the time taken to answer the questionnaires will help us in answering the question of whether the presence of gridlines help in a better comprehension of trend-lines present in a data visualization.

2.5 Practical setup

The experiment will be conducted offline within a controlled environment. The participants must physically travel to the premises where the experiment will be performed. It is ensured that the room where the experiment is being conducted has adequate lighting in all areas and is distraction-free.

Once the participants arrive, they will be asked to disclose any sight-problems that they may possess. Participants who do not have any eye-sight problems are allowed to continue with the experiment. They are then given a brief introduction about data visualization and other key words that are used in the questionnaire. After this, they will be instructed to complete two sets of questionnaire where the first questionnaire will consists of data visualizations without grid-lines and the second one will consist of visualizations with gridlines. The individual time taken to answer each set will be noted. The questionnaire will be pen and paper based where the visualization will be displayed on a printed paper and the participant will be asked to indicate his/her response using a tick-mark placed below the option.

The following instructions will be given to the participants at the beginning of the experiment:

a) The survey will be conducted using two questionnaires namely Part-A and Part-B where Part-A should be answered first.
b) Both parts consist of three questions each. Please ensure your responses are indicated by using a clear tick mark below the option. Kindly use only one response for each question.
c) Kindly ensure that I am notified when you complete each part of the questionnaire so that the individual time taken to answer the particular section can be noted down.

The experiment will last approximately ten minutes in total and the questionnaires which have been answered will be collected at the end of the experiment to analyse the results.

3 Data Visualisations

The charts A.1 and A.2 displayed below are examples of charts which contain a trend-line but do not contain a gridline. The participant will be asked to compare these two charts and answer a question based on the comparison.

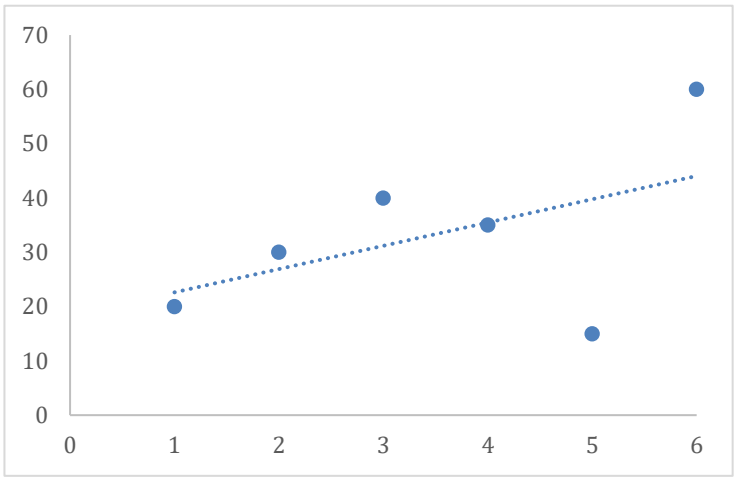


Chart A.1

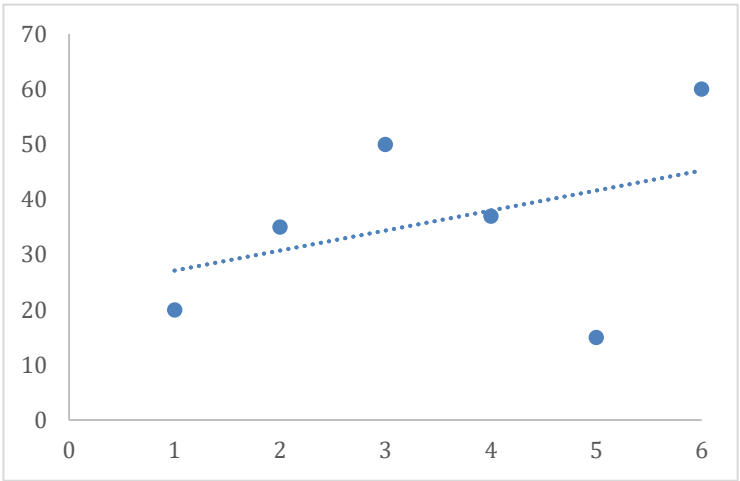


Chart A.2

The charts B.1 and B.2 displayed below are examples of charts which contain a trend-line and are accompanied with gridlines. The participant will be asked to compare these two charts and answer a question based on the comparison.

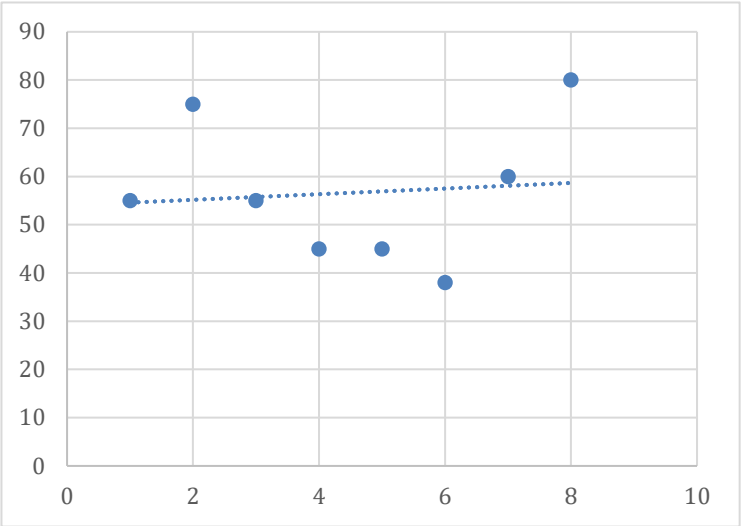


Chart B.1

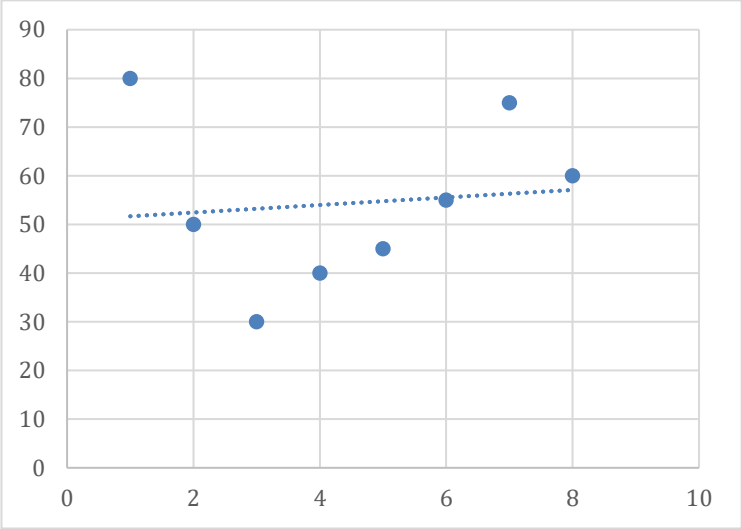


Chart B.2

4 Pilot Experiment

4.1 Data Analysis

The participant's response to the questionnaire and the time taken by each participant is recorded and analysed. The following table is used to summarize the data obtained during the experiment. The correct and incorrect values are established by comparing the participant's response with the ground truth that was established beforehand.

The tables displayed below summarize the data that has been collected during the pilot experiment.

QUESTIONNAIRE PART-A				
	QUESTION-1	**QUESTION-2**	**QUESTION-3**	**AVERAGE**
PARTICIPANT-1	Incorrect	Correct	Incorrect	
PARTICIPANT-2	Incorrect	Incorrect	Correct	
PARTICIPANT-3	Incorrect	Incorrect	Incorrect	
ACCURACY	0	0.33	0.33	0.22
TIME TAKEN (in minutes)	1.44	1.49	1.32	1.42

TABLE 1.1: PART-A Data Summary

QUESTIONNAIRE PART-B				
	QUESTION-1	**QUESTION-2**	**QUESTION-3**	**AVERAGE**
PARTICIPANT-1	Correct	Correct	Incorrect	
PARTICIPANT-2	Incorrect	Correct	Incorrect	
PARTICIPANT-3	Correct	Incorrect	Correct	
ACCURACY	0.66	0.66	0.33	0.55
TIME TAKEN (in minutes)	1.09	1.1	0.5	0.89

TABLE 1.2: PART-B Data Summary

Table 1.1 consists of the data which summarises the result of responses obtained by participants answering the Part-A of the questionnaire which consists of data visualizations without gridlines.

Table 1.2 consists of the data which summarises the result of responses obtained by participants answering the Part-B of the questionnaire which consists of data visualizations with gridlines.

	AVERAGE ACCURACY	**AVERAGE TIME TAKEN (in minutes)**
PART-A (Without Gridlines)	0.22	1.42
PART-B (With Gridlines)	0.55	0.89

TABLE 1.3: Results Summary

From Table 1.3, we can compare the results obtained by analysing the response of participants in each part of the questionnaire. It can be observed that the accuracy obtained while comprehending data visualizations with gridlines is higher at about 0.55 whereas the accuracy in interpreting the data visualizations without gridlines stands at 0.22. It can also be observed that the average time taken to answer the questionnaire consisting of visualizations without gridlines was higher when compared to the average time taken to answer the questions that accompanied the visualizations with gridlines.

The above observations suggest that presence of gridlines help in a better comprehension of the trend-line present in a data visualization.

4.2 Reflections

The pilot experiment was carried out with three participants. During the experiment, it was found to be difficult to keep track of time for all three participants simultaneously. Feedback from the participants showed that the gridlines did not appear to be clear on print. Hence, the experiment can be conducted individually instead of simultaneously or the questions can be presented using an online form to all the participants simultaneously which will have a timer enabled to report the individual time taken to complete each part of the questionnaire. Using an online form will also help display the gridlines present in the data visualizations in a clearer manner. The language in which the questions appear also affect the time taken to answer the questionnaire as the proposed design assumed the participants to be fluent in English. This can be corrected by ensuring the selection of participants who are fluent in English or by supplying questionnaires in the language preferred by the participant. The addition of more questions with trendlines of significant difference for comparison will help in assessing the cognitive ability of the participant with a reduced error rate. Considering these factors will help improve the outcome of the experiment and reduce any errors that may occur.

The proposed experiment design attempts and succeeds in establishing that if gridlines are added to data visualizations that contain a trend-line, it leads to a better comprehension of the trends being presented in the chart.

References

Schield, M. (December 2001). Three Kinds of Statistical Literacy: What Should We Teach. ICOTS-6.
Available at: http://web.augsburg.edu/~schield/milopapers/2002icots6web.pdf (Accessed: 20 November 2018)

Tufte, E. Visual Display of Quantitative Information (1983). Available at: https://books.google.ie/books?hl=en&lr=&id=XPpDDQAAQBAJ&oi=fnd&pg=PA219&dq=edward+tufte+grid&ots=jPhcHuCXlk&sig=Z9WBQqxQFbBuLgS3I_hZgi96z-E&redir_esc=y#v=onepage&q&f=false (Accessed on: 20 November 2018)

William S. Cleveland, Marylyn E. McGill and Robert McGill. The Shape Parameter of a Two-Variable Graph. Journal of the American Statistical Association. Vol. 83, No. 402 (Jun., 1988), pp. 289-300 (12 pages). Available at: https://www.jstor.org/stable/2288843?seq=8#metadata_info_tab_contents (Accessed on: 20 November 2018)

Kosara, R. Aspect Ratio and Banking to 45 Degrees (June 2013). Available at: https://eagereyes.org/basics/banking-45-degrees (Accessed on: 20 November 2018)

Available at: https://www.verywellmind.com/what-is-a-within-subjects-design-2796014 (Accessed on: 20 November 2018)

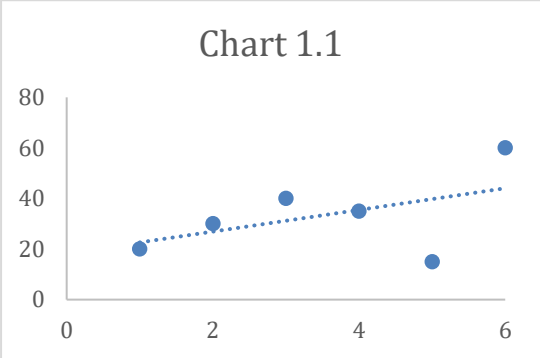
Available at: https://www.statisticshowto.datasciencecentral.com/experimental-design/confounding-variable/ (Accessed on: 20 November 2018)

Appendix

Participant Questionnaire- Part-A

Please use a ✓ below the given options to indicate your response to the question. Use only one attempt per question.

1) Among Charts 1.1 and 1.2 , which chart shows the highest increase in trends?



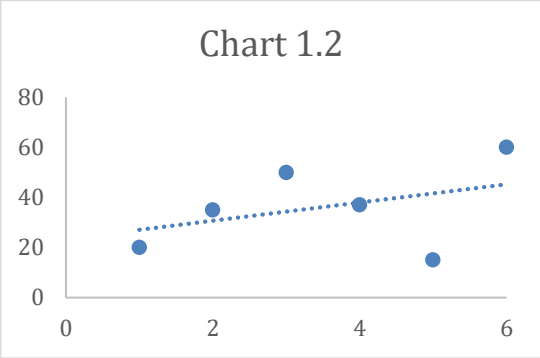
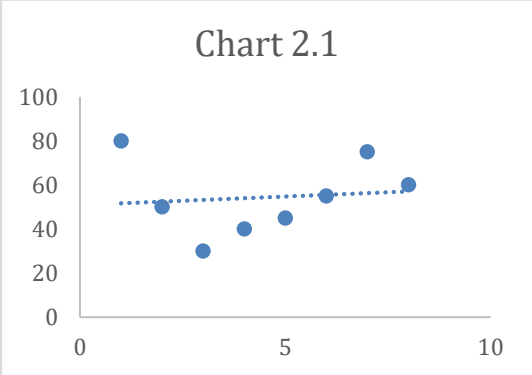
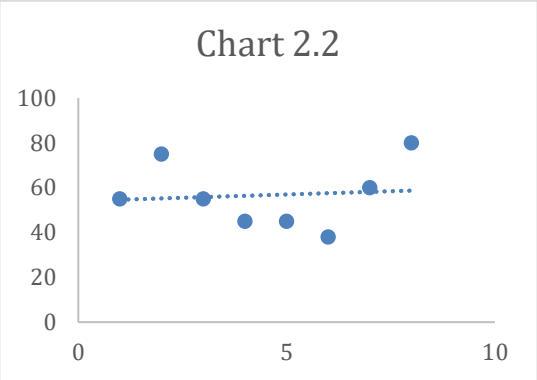


Chart 1.1	Chart 1.2	Both have same trends	Difficult to Say

2) Is the following statement true or false?

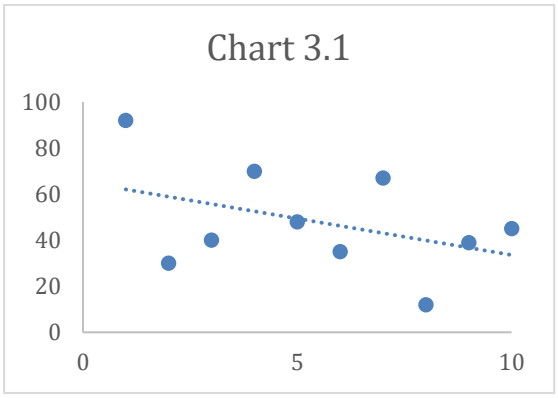
Statement: The increase in trends observed in chart 2.1 is lower than the increase observed in chart 2.2.





True	False	Difficult to Say

3) Among charts 3.1 and 3.2, which do you think represents a higher decrease in trends?



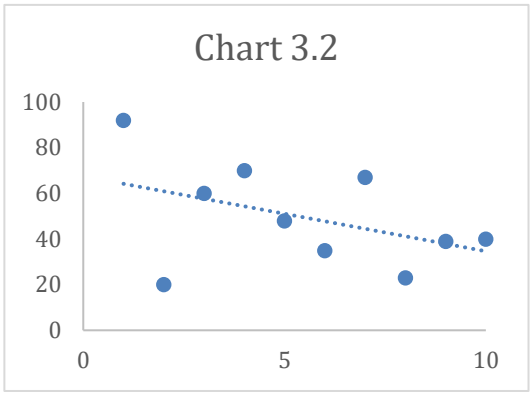
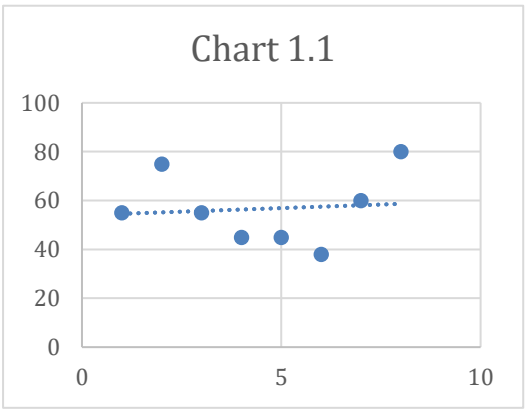


Chart 3.1	Chart 3.2	Both have same trends	Difficult to Say

Participant Questionnaire- Part-B

Please use a ✔ below the given options to indicate your response to the question. Use only one attempt per question.

1) Among Charts 1.1 and 1.2 , which chart shows the highest increase in trends?



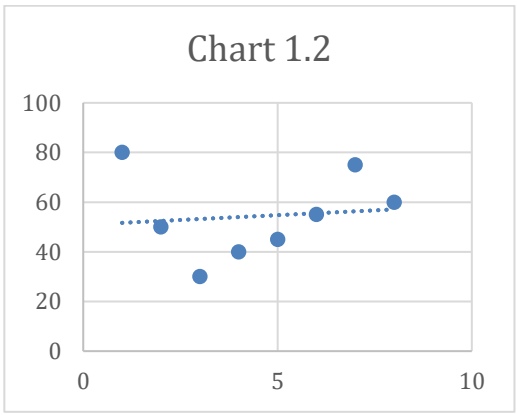
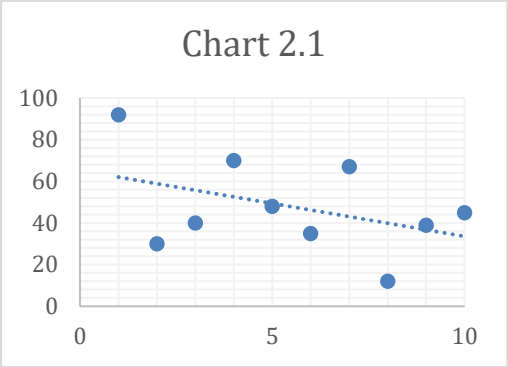


Chart 1.1	Chart 1.2	Both have same trends	Difficult to Say

2) Among charts 2.1 and 2.2, which do you think represents a steeper decrease in trends?



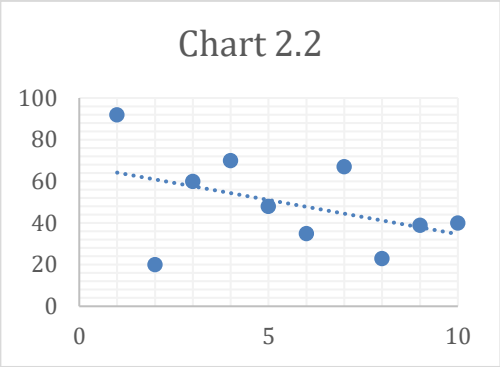
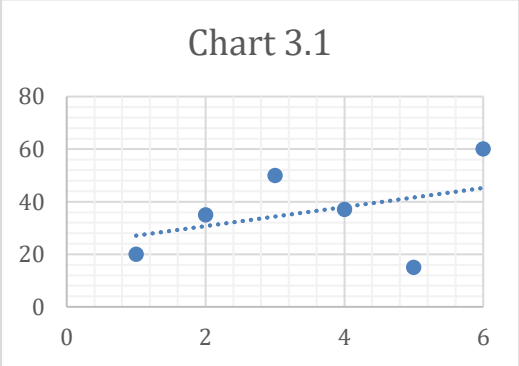


Chart 2.1	Chart 2.2	Both have same trends	Difficult to Say

3) Among charts 3.1 and 3.2, which do you think represents a decrease in trends?



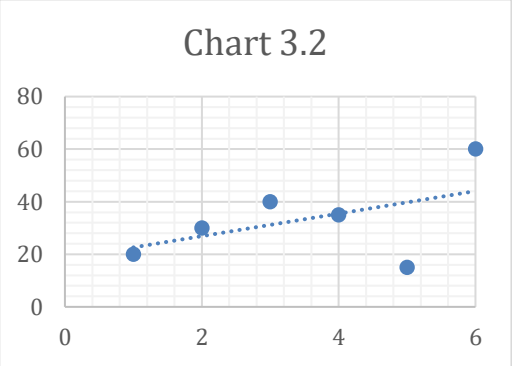


Chart 3.1	Chart 3.2	Both have same trends	Difficult to Say