# Think-Alouds for Draft Questions

Some of these draft assessment questions need major revisions, while others only need small changes. We used think-aloud interviews to improve them.

**Instructions:** In small groups, pick a couple draft questions on the handout and interview each other. After completing the interviews, discuss your experience with the group. Is there anything you would change in the questions? How do you think students would respond? Afterward, we'll discuss what we learned from our interviews with students.

Interviewee: Please remember to think aloud: don't pause if you can avoid it, feel free to ask any questions you have about the question (though the interviewer can't respond), and remember to say everything you're thinking.

Interviewer: If you feel your interviewee isn't thinking aloud, feel free to say: "Please remember to think aloud." Remember to not give any feedback—verbal or non-verbal—to the interviewee, regardless of whether they are correct or not.

---

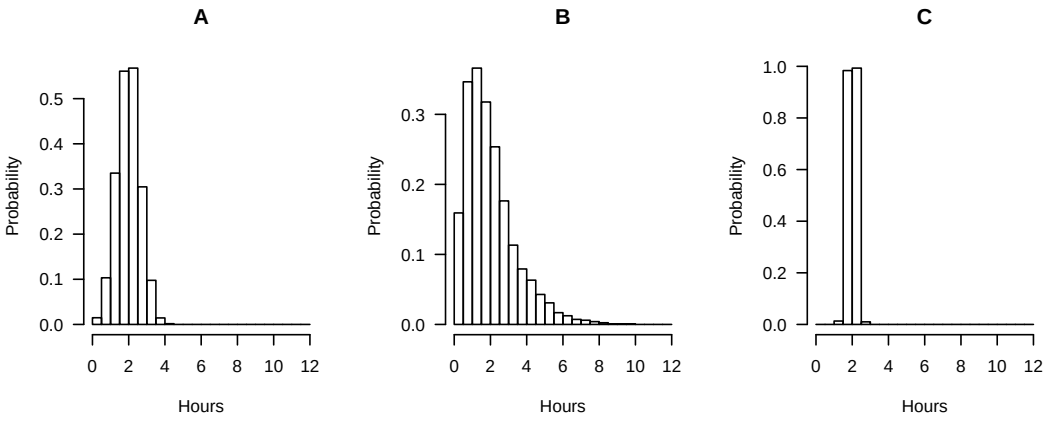
Part of "Using think-aloud interviews to assess student understanding of statistics concepts," with A Reinhart, P Burckhardt, P W Elliott, C Evans, A Luby, M Meyer, J Orellana, R Yurko, G Weinberg, J Wieczorek, and R Nugent. *USCOTS 2019*.

For slides, papers, and references, check out the TeachStat Research Group website: `http://www.stat.cmu.edu/teachstat/`

We'd also appreciate your feedback on our draft questions in our Instructor Survey! `https://isle.heinz.cmu.edu/surveys/instructor/`

Interested in using our assessment in your course or participating in our research? Contact Alex Reinhart: `areinhar@stat.cmu.edu`

`study-time` To estimate the average number of daily hours that students study at a large public college, a researcher randomly samples some students, then calculates the average number of daily study hours for the sample. Pictured below (in scrambled order) are three histograms: One of them represents the population distribution of study hours; the other two are sampling distributions of the mean $\bar{X}$, one for sample size $n = 5$, and one for sample size $n = 50$.




Circle the most likely distribution for each description.

(a) Population distribution. A B C

(b) Sampling distribution for $n = 5$. A B C

(c) Sampling distribution for $n = 50$. A B C

How confident are you in your answer? ☐ Guessed ☐ Somewhat sure ☐ Confident

`plot-matching`

(a) What graphic would be most appropriate to visualize the number of students in each major at Carnegie Mellon University?

i. Scatterplot
ii. Mosaic plot
iii. Bar chart
iv. Box plots

(b) What graphic would be most appropriate to visualize the weights of burgers served at three different restaurants?

i. Histogram
ii. Scatterplot
iii. Mosaic plot
iv. Box plots

(c) What graphic would be most appropriate to visualize the final exam grades for a large introductory course?

i. Box plots
ii. Mosaic plot
iii. Scatterplot
iv. Histogram

(d) What graphic would be most appropriate to visualize the SAT scores and GPAs of freshman students?

i. Scatterplot
ii. Histogram
iii. Box plots
iv. Mosaic plot

How confident are you in your answer? ☐ Guessed ☐ Somewhat sure ☐ Confident

`vitamin-c` A clinical trial randomly assigned subjects to receive either vitamin C or a placebo as a treatment for a cold. The trial found a statistically significant negative correlation between vitamin C dose and the duration of cold symptoms. Which of the following can we conclude?

(a) Recovering faster from a cold causes subjects to take more vitamin C

(b) Taking more vitamin C causes subjects to recover faster from a cold.

(c) We cannot draw any conclusions because correlation does not imply causation.

(d) We cannot draw any conclusions because assignment was random instead of systematic

How confident are you in your answer? ☐ Guessed ☐ Somewhat sure ☐ Confident

`horse-races` You have 25 ducks. You want to know which duck is the lightest and which is the heaviest, so you begin to weigh the ducks in a random order. However, after you have weighed five ducks, the scale breaks.

Which of the following events is most likely?

(a) The heaviest duck you weighed is also the heaviest of the 25

(b) You weighed the lightest duck of the 25

(c) The heaviest duck you weighed is the heaviest of the 25 **and** you weighed the lightest of the 25

(d) Either the heaviest duck you weighed is the heaviest of the 25 **or** you weighed the lightest of the 25

How confident are you in your answer? ☐ Guessed ☐ Somewhat sure ☐ Confident

`books` A survey of Californians found a statistically significant positive correlation between number of books read and nearsightedness.

Which of the following can we conclude about Californians?

(a) Reading books causes an increased risk of being nearsighted.

(b) Being nearsighted causes people to read more books.

(c) We cannot determine which factor causes the other, because correlation does not imply causation.

(d) We cannot draw any conclusions because Californians aren't a random sample of people.

How confident are you in your answer? ☐ Guessed ☐ Somewhat sure ☐ Confident

`wacky-alpha` Farmer Brown is about to conduct a study and will use the $p$-value of a hypothesis test to reach her conclusion.

- If $p < 0.01$, she will reject the null hypothesis at the 1% significance level.
- If $p < 0.05$, she will reject the null hypothesis at the 5% significance level.
- If $p < 0.10$, she will reject the null hypothesis at the 10% significance level.

What is the probability that Farmer Brown will falsely reject a true null hypothesis?

(a) 1%

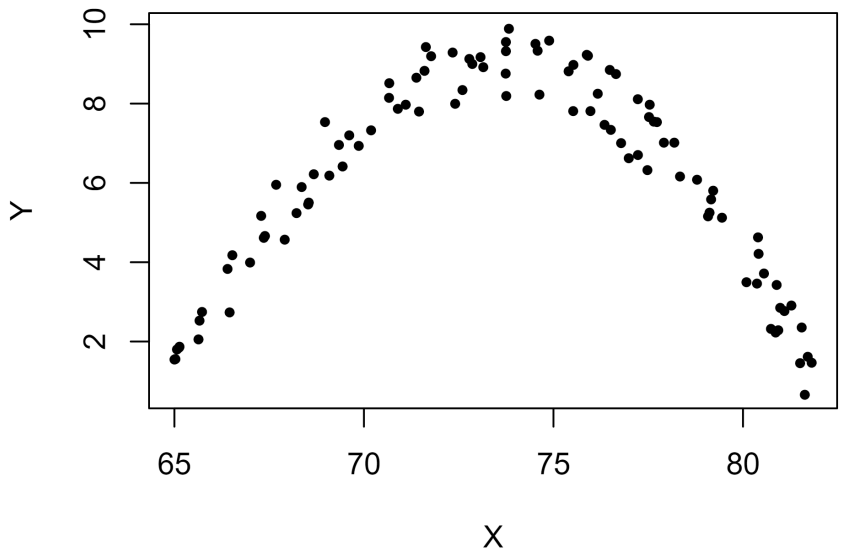
(b) 5%

(c) 10%

(d) 1%, 5% or 10% depending on the $p$-value.

How confident are you in your answer? ☐ Guessed ☐ Somewhat sure ☐ Confident

`u-correlation` You create a scatterplot of two continuous variables $X$ and $Y$:




Which of the following values is the best correlation $r$ between $X$ and $Y$ to report?

(a) $-0.5$
(b) 0.5
(c) $-0.9$
(d) 0.9
(e) Correlation is not an appropriate measure for these two variables
(f) Need more information

How confident are you in your answer? ☐ Guessed ☐ Somewhat sure ☐ Confident

## Want to try think-aloud interviews with your students?

There are several great resources you can use while you design a think-aloud interview study:

- W. K. Adams and C. E. Wieman, "Development and validation of instruments to measure learning of expert-like thinking," *International Journal of Science Education*, vol. 33, no. 9, pp. 1289–1312, 2011. https://doi.org/10.1080/09500693.2010.512369
- M. W. van Someren, Y. F. Barnard, and J. C. Sandberg, *The Think Aloud Method: A practical guide to modelling cognitive processes*, Academic Press, 1994. http://hdl.handle.net/11245/1.103289

Just remember that this is human subjects research, so if you want to publish your work, get IRB approval first! We'd be happy to talk with you about our experiences and share any wisdom we might have.