**Instructor Guide: Building Intuition for Two-Sample Inference using Lineups**

**Quick Info**

Level: *Intro undergraduate statistics*

Brief Description: *Students build their intuition about hypothesis testing by answering a research question using a visual test.*

Topics Covered: *exploratory data analysis, stating hypotheses, testing two-sample hypotheses via lineups*

Learning Goals: *Explore the logic of hypothesis tests and their implications. Develop intuition for the logic of hypothesis tests using a visual test. Visualize a reference (null) distribution.*

Software Required: *Optional. Students can use R, JMP, SPSS, etc. to explore the data set prior to viewing the lineups, or the graphics could be included on the student handout.* *Students will need access to the lineup (this can be done using the overhead, via handouts, or by using a* [*Shiny app*](http://shinyapps.its.carleton.edu/multi-sample-quant-app/)*).*

Prerequisites: *descriptive statistics, histograms or boxplots, basic experimental design*

Time: ~*50 minutes in class*

Instructor Resources: *Student handout and instructor guide, sample data set, Shiny app (*[*http://shinyapps.its.carleton.edu/multi-sample-quant-app/*](http://shinyapps.its.carleton.edu/multi-sample-quant-app/)*), tutorial on creating lineups in R (*[*https://aloy.github.io/classroom-vizinf/*](https://aloy.github.io/classroom-vizinf/)*)*

**Why use this lineup activity in your course?**

You can use the lineup protocol to visually introduce the logic behind hypothesis tests with few technical details. Lineups are created by generating a number of decoy plots and randomly situating the data plot into this grid of plots. Using this lineup of plots, students can apply “Sesame Street logic” (i.e., “which one of these is not like the others”), which can then be linked to fundamental statistical ideas. This activity provides the scaffolding to make this connection.

**What type of course is this lineup activity designed for?**

This activity is designed for any undergraduate course that discusses two-sample hypothesis tests.

**When should you use this activity in your course and what are the prerequisites?**

This activity is designed to introduce concepts underlying hypothesis tests, so it should be used on the first day you discuss testing. In my course, this comes after a unit on data collection and EDA. Students should already have working knowledge of descriptive statistics (mean, standard deviation, etc.), statistical graphics (boxplots, etc.), and data collection schemes (basic experimental design). Instructors could assign a reading or give a mini lecture on crafting hypotheses before this activity.

**How should you conduct the activity? How much time should you expect to allocate?**

* **Introduce logic of testing** (5 minutes, depending on the scope).At the beginning of class, take a few minutes to introduce (or review from the reading) the logic behind hypothesis tests—why can’t we simply answer our research question using our one sample. *The logic behind testing could be introduced via reading or a pre-class video. This is also a good time to assign student roles in the groups.*
* **Students discuss questions #1-5 in groups** (10 minutes). These questions review basic experimental design and EDA, and challenge the students to develop appropriate hypotheses.
* **Introduce lineups** (3 minutes)**.** The important point here is that students know one panel is the observed data and the other plots are situations where there is “no difference between groups.” Don’t worry about the technical details and conditions here. *You could also take a few minutes to review the answers to questions #1-5 at this point.*
* **Question #6: Students evaluate the lineup individually** (2 minutes).Be sure to remind students to work individually on this evaluation. Discussion is more interesting and meaningful when each student is an unbiased evaluator. *Saying “take two minutes to evaluate the lineup on your own” is helpful. You could even display a timer if that would be helpful.*
* **Question #7: Groups discuss lineups and come to consensus** (5 minutes).Having students unpack their thoughts about the lineups in groups helps them flesh out why they chose the panel. In addition, discussing one choice per group makes it easier to hear from “everyone” in a short amount of time in a lower stress environment (it’s the group’s choice, not the individual’s). *I recommend announcing when the groups should start their discussion.*
* **Regroup and reveal the data panel.** Have each group briefly report what panel they chose and why (1 minute per group). Display the lineup on a slide and announce and/or highlight the data panel. *I like asking one of the groups who selected the data panel why they made that choice. I have also summarized the rationales for that panel that I just heard.*
* **Question #8: Groups discuss implications of their choice** (3-5 minutes)**.** Students return to their groups to discuss what their choice of plot suggests about the competing claims. *Encourage students to link directly to the claims as they report their answers, this tends to get more focused answers and guides them toward thinking about evidence. I have found the timing for this portion of the activity to be quite variable.*
* **Debrief** (10 minutes)**.** Regroup and have each group report back, or randomly select a couple groups to share their thoughts if time is tight. Once each group has shared their thoughts, summarize how this activity links to each step of the logic behind hypothesis testing, ending with what the class just concluded and how that differs from using a single plot or test statistic.

**Here are some helpful hints for instructors using lineups for the first time.**

1. Spend some time introducing the data set. This can be done while you distribute the handouts and as students get situated in their groups.
2. Have students work in small groups, perhaps 3 or 4. This provides students with the opportunity to discuss their understanding in a lower-stakes environment and to learn from each other.
3. Assign roles to each group member to help groups function efficiently and to avoid the situation where one student does all of the work. We recommend assigning roles such as
   * + Facilitator: makes sure the group stays on task and that each member has room to contribute
     + Spokesperson: reports back to the class, reads from the recorder’s notes
     + Recorder: completes the worksheet for the group, takes coherent notes
     + encourager/questioner: suggests alternatives if the group gets stuck, asks for clarification, poses questions

**Follow-up Activities and Discussion Questions:**

You could follow-up this activity with homework questions or warm-up questions for the next class where students evaluate another lineup or two. This would give you the opportunity to include a lineup for a situation where there is not a discernible difference between the two groups and/or using a different plot type.

* + Do you think plot type influences your ability to detect differences between groups? If so, how? *Plots are test statistics here, so it can be linked to the choice of a test statistic later in the course (e.g., difference in means vs. difference in medians).*
  + If you cannot identify the data plot from the lineup, does that mean there is no difference between groups?
  + If you can identify the data plot from the lineup, does that mean there is an important difference between groups?
  + Consider two situations: (1) you are the only person in class who selected the data plot; (2) twenty people in class select the data plot. What do these two situations suggest about the competing claims?

**What else is in this Instructor Guide?**

In the next section, we provide a commented version of the student activity. We suggest possible alternative formats you can use, questions that you can ask students to facilitate discussion, and possible issues you may encounter.

**References**

This activity was adapted from a case study discussed in chapter 1 of *The Statistical Sleuth*, 3rd ed.

The original study citation is

Amabile, T. M. (1985). Motivation and creativity: Effects of motivational orientation on creative writers. *Journal of Personality and Social Psychology*, 48(2), 393.

The format of this instructor guide was inspired by Shonda Kuiper’s Stat2Labs.

**Building Intuition for Two-Sample Inference using Lineups**

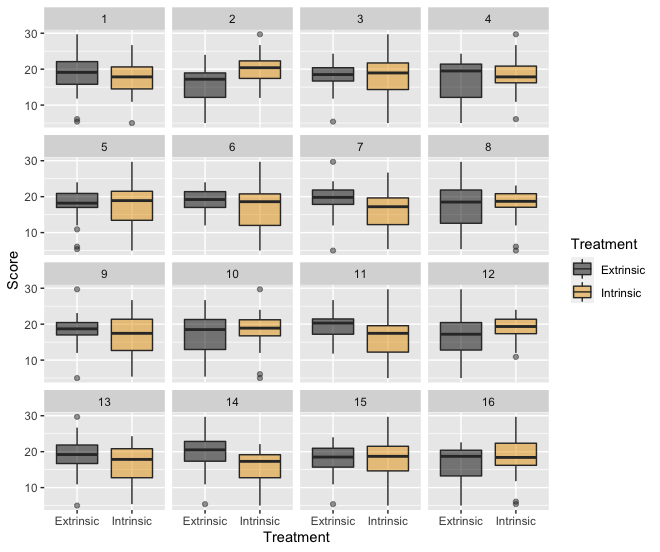
**Background**Evidence suggests that reward systems may operate in the opposite way from what is intended (e.g., ranking systems may decrease productivity; grading systems not stimulate learning). To investigate this phenomenon, Amabile, T. M. (1985) designed a study to explore whether motivation type (intrinsic or extrinsic) impacted creativity scores. In this study, 47 creative writers were randomly assigned to one of two questionnaires where they ranked reasons they write. One questionnaire listed intrinsic motivations and the other listed extrinsic motivations. After completing the questionnaire, all subjects wrote a Haiku about laughter, which was graded for creativity by a panel of poets. The average rating (out of 40 points) for each subject was recorded. In this activity, you will explore whether there are discernible differences in creativity between the two groups.

*Answer questions #1-5 as a group.*

1. Is this an experimental or observational study? How do you know?
2. What research question is being investigated? That is, what is the objective of this study?
3. Write down the competing claims (i.e., hypotheses) being investigated.
4. Identify the response variable and the treatment.
5. What type of plot could you use to compare the two treatment groups? Why did you choose this plot type? How does this plot help you investigate the competing claims?

*Answer question #6 individually. Please do not discuss your answers with your group until you start question #7.*

1. Below is a lineup where one panel is a plot of the observed data and the other 19 plots are generated under the situation where there is no difference between the treatment groups. Please answer the following questions individually. You will discuss your thoughts with your group in the next question.
   1. Which panel is the most different from the others?
   2. What feature of the plot led you to this choice?



*Answer question #7 as a group.*

1. Now that everyone in your group has recorded their thoughts about the lineup, regroup and discuss your thoughts. Come up with a consensus answer to each of the below questions that your group’s spokesperson will share with the class when we regroup.
   1. Which panel is the most different from the others?
   2. What feature of the plot led you to this choice?

*I will reveal the observed data panel to the class soon. Let me know that you have reached this point to help me monitor our time.*

The observed data are in panel # \_\_\_\_\_\_.

1. Now that you know which panel contains the observed data, discuss the following questions with your group. Your spokesperson will share these thoughts when we regroup.
   1. Did your group choose the data panel?
   2. If you chose the data panel, what does this suggest about your competing claims? If you did not choose the data panel, what does this suggest about your competing claims?

*We will regroup shortly to discuss your thoughts. If you have any questions, please ask!*



This work is licensed under a [Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License](http://creativecommons.org/licenses/by-nc-sa/4.0/).