RStudio_090523

Steven Swee

2023-09-05

In this worksheet, we will go over Single-Subject Research Designs. In particular, we will examine multiple baseline designs and possible statistical methods.

Problem Statement

The program you ran to help students study longer was a success! Word gets out and another school district wants you to implement your program. However, this new school district does not have as much funding and wants to know if the timing of the program matters. Since you have evidence that your program works, it may not be ethical to perform a reversal (ABA) design. How can you design a study that tests if your program works? What sort of analysis can be done?

Preliminaries

We will start by importing any libraries that may be useful.

```
# Import libraries
library(ggplot2)
```

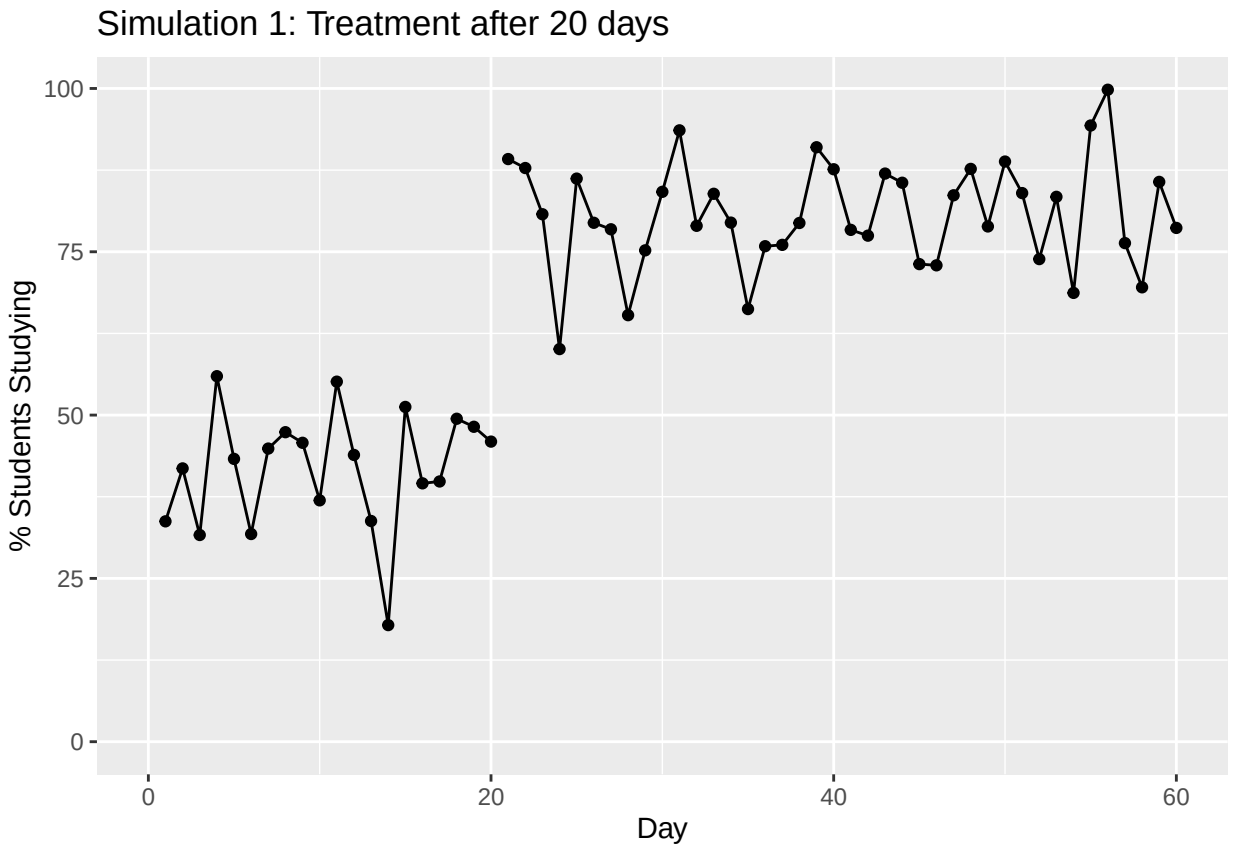
Multiple-Baseline Designs

There may be cases where your treatment/intervention is so effective that it may be unethical to remove it. Additionally, there may also be cases where your independent variable may not return to baseline after a treatment (what would happen if you suddenly stopped giving a child rewards?).

In multiple-baseline designs, treatments are introduced at varying times for each participant. Simulations of what the data may look like are shown below:

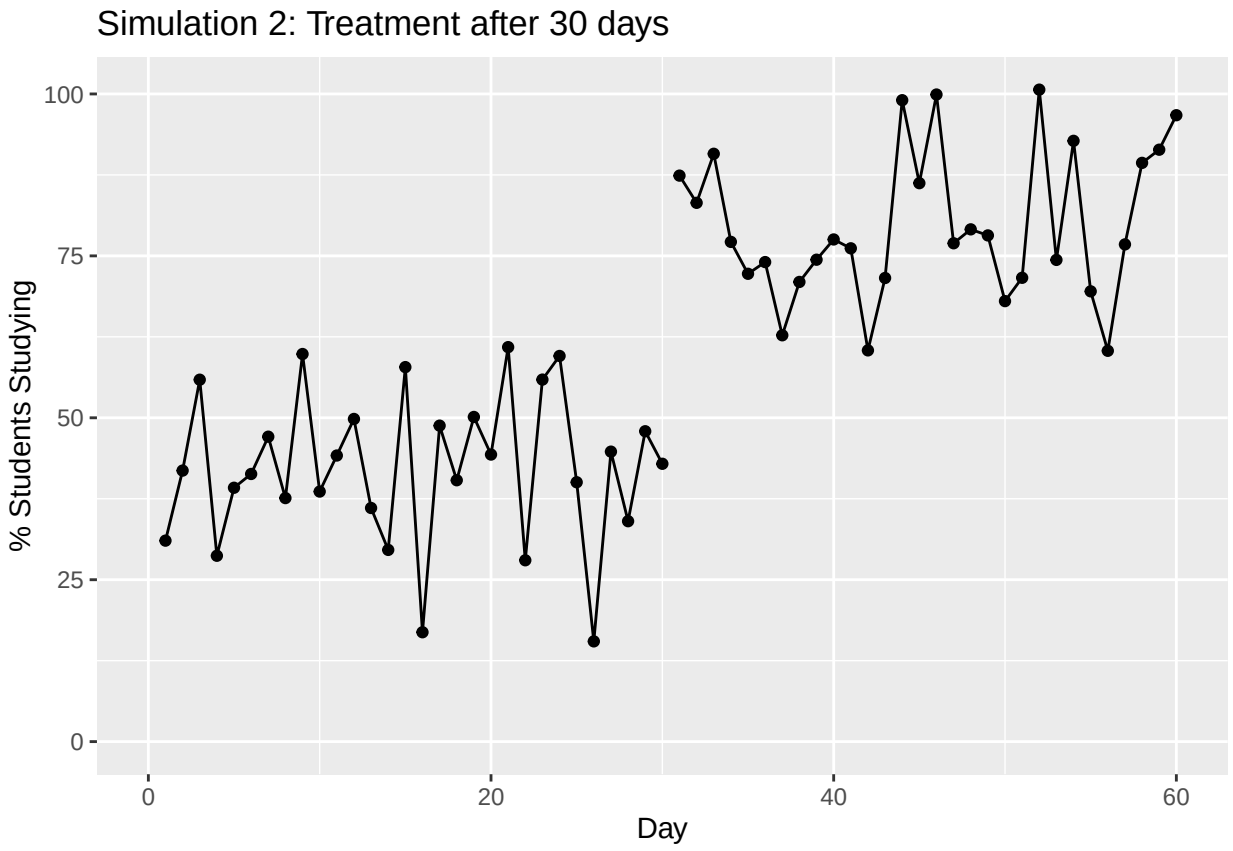
```
# Generate random data
set.seed(1) # Set seed to make data reproducible
n_base = 20 # Number of days for baseline
n_treat = 40 # Number of days for treatment
base = rnorm(n_base, mean = 0.4, sd = 0.1) * 100
treat = rnorm(n_treat, mean = 0.8, sd = 0.1) * 100
data = c(base, treat)
time = seq(from = 1, to = length(data))
condition = c(rep("Baseline", n_base), rep("Treatment", n_treat))
df = data.frame(data, time, condition)
```

```r
# Make plot
ggplot(data = df, aes(x = time, y = data, group = condition)) +
  geom_line() +
  geom_point() +
  expand_limits(x = 0, y = 0) +
  xlab("Day") +
  ylab("% Students Studying") +
  ggtitle("Simulation 1: Treatment after 20 days")
```

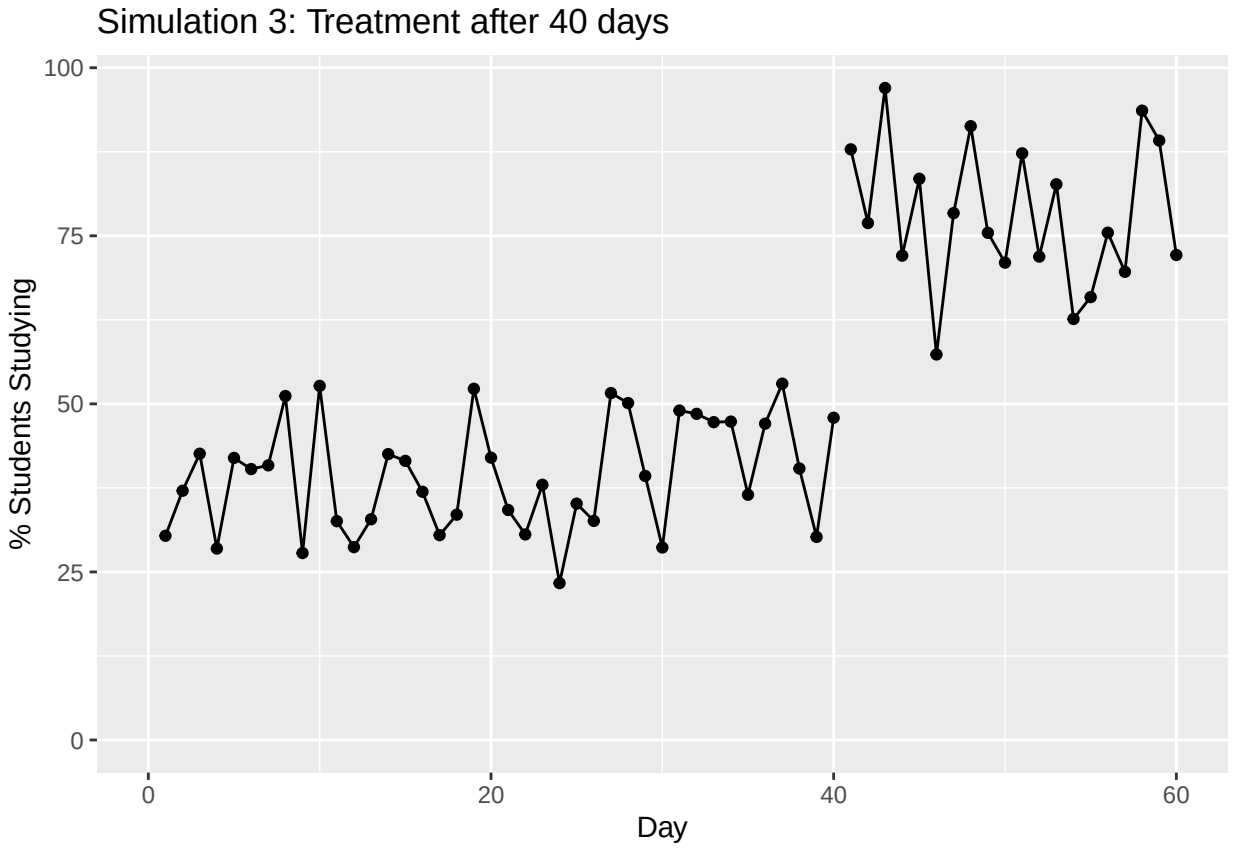



```r
# Generate random data
set.seed(2) # Set seed to make data reproducible
n_base = 30 # Number of days for baseline
n_treat = 30 # Number of days for treatment
base = rnorm(n_base, mean = 0.4, sd = 0.1) * 100
treat = rnorm(n_treat, mean = 0.8, sd = 0.1) * 100
data = c(base, treat)
time = seq(from = 1, to = length(data))
condition = c(rep("Baseline", n_base), rep("Treatment", n_treat))
df = data.frame(data, time, condition)

# Make plot
ggplot(data = df, aes(x = time, y = data, group = condition)) +
  geom_line() +
  geom_point() +
  expand_limits(x = 0, y = 0) +
  xlab("Day") +
```

```
  ylab("% Students Studying") +
  ggtitle("Simulation 2: Treatment after 30 days")
```




```
# Generate random data
set.seed(3) # Set seed to make data reproducible
n_base = 40 # Number of days for baseline
n_treat = 20 # Number of days for treatment
base = rnorm(n_base, mean = 0.4, sd = 0.1) * 100
treat = rnorm(n_treat, mean = 0.8, sd = 0.1) * 100
data = c(base, treat)
time = seq(from = 1, to = length(data))
condition = c(rep("Baseline", n_base), rep("Treatment", n_treat))
df = data.frame(data, time, condition)

# Make plot
ggplot(data = df, aes(x = time, y = data, group = condition)) +
  geom_line() +
  geom_point() +
  expand_limits(x = 0, y = 0) +
  xlab("Day") +
  ylab("% Students Studying") +
  ggtitle("Simulation 3: Treatment after 40 days")
```



Question 1

Why may it be considered unethical to remove a treatment that is working?

Question 2

Design an intervention that uses the reversal design. Then, modify the intervention that uses the multiple-baselines design.

Question 3

Why might the timing of an intervention be a useful feature to consider?

Question 4

You come across two studies. One study introduces the same intervention to three different people at the same time. Another study introduces the same intervention to three different people at three different times. What are the advantages/disadvantages of each study? (Hint: If the dependent variable changes the same way at three different times, then it is unlikely that time is a contributing factor).

Variations of multiple-baseline designs

Like the reversal design, there are ways you can modify the multiple-baseline design. One way is measuring multiple dependent variables. If a treatment is effective for more than one dependent variable, then there is further evidence that the treatment is effective.

Question 5

Using the intervention you came up with in Question 2, determine two variables that you would like to measure. Why is measuring both of these variables better than measuring than just one variable?

Another variation is to use different settings for the treatments. If the treatment is effective at multiple locations, such as at work or at home, then that serves as further evidence that the treatment is effective.

Question 6

Using the intervention you came up with in Question 2, come up with two different settings that you can implement your treatment. Suppose your intervention works in one place but not another. What does this suggest?

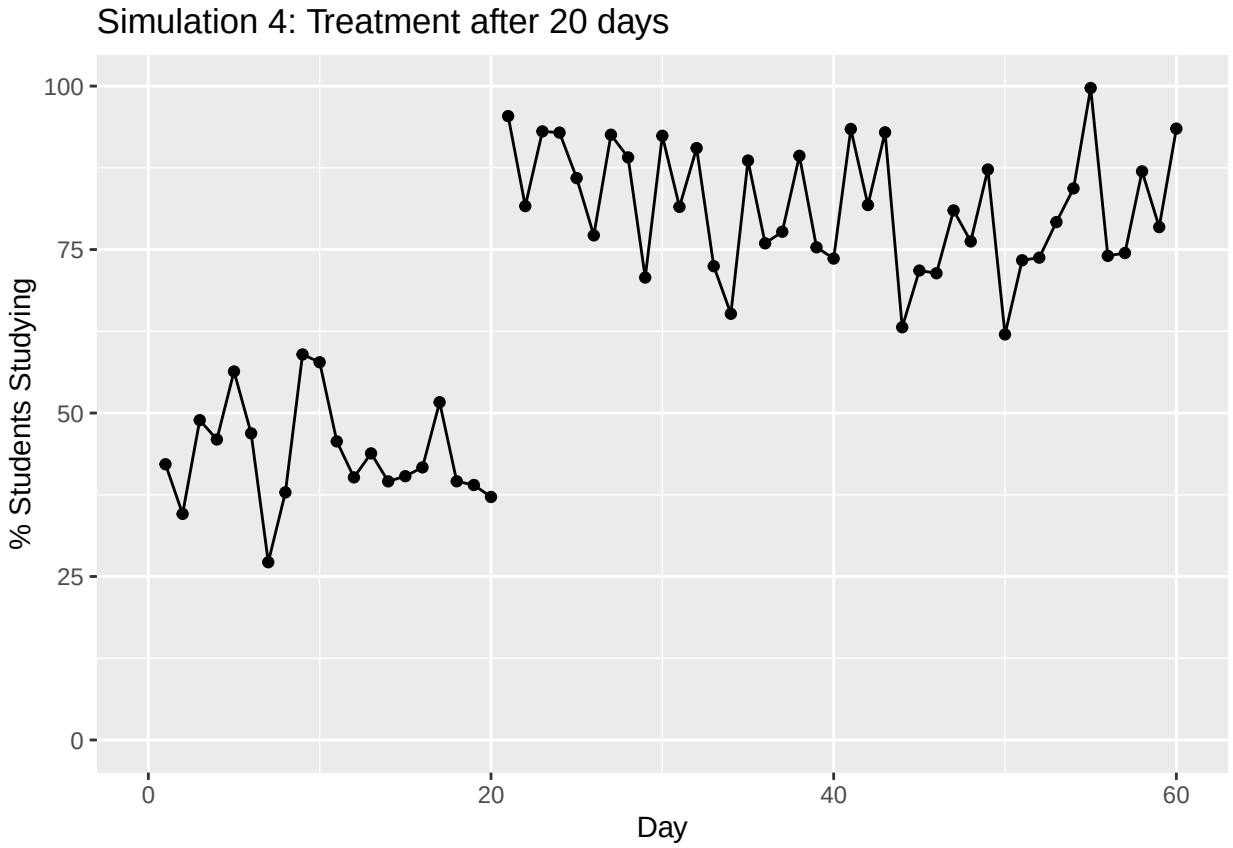
Statistical Analysis in Single-Case Studies

While visual inspection is commonly used in single-subject research, inferential statistics may be used.

The t-test

One possible approach is to use the t-test to see if the averages between the baseline and the treatment are different from each other. Let's run the t-test on the simulated data below to see if there is a significant difference.

```
# Generate random data
set.seed(4) # Set seed to make data reproducible
n_base = 20 # Number of days for baseline
n_treat = 40 # Number of days for treatment
base = rnorm(n_base, mean = 0.4, sd = 0.1) * 100
treat = rnorm(n_treat, mean = 0.8, sd = 0.1) * 100
data = c(base, treat)
time = seq(from = 1, to = length(data))
condition = c(rep("Baseline", n_base), rep("Treatment", n_treat))
df = data.frame(data, time, condition)

# Make plot
ggplot(data = df, aes(x = time, y = data, group = condition)) +
  geom_line() +
  geom_point() +
  expand_limits(x = 0, y = 0) +
  xlab("Day") +
  ylab("% Students Studying") +
  ggtitle("Simulation 4: Treatment after 20 days")
```

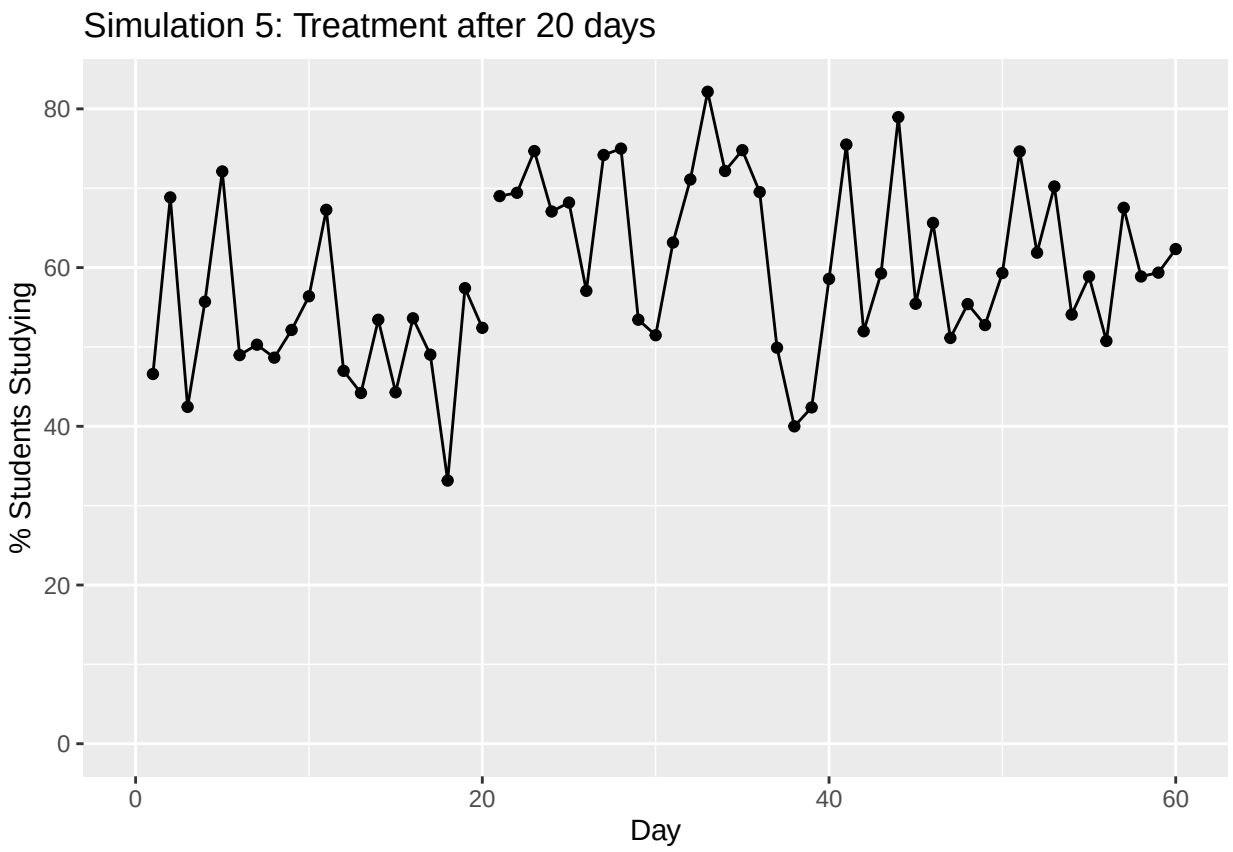



```r
t.test(base, treat, var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  base and treat
## t = -15.109, df = 58, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -42.73494 -32.73623
## sample estimates:
## mean of x mean of y
##  43.76296  81.49855
```

```r
# Generate random data
set.seed(5) # Set seed to make data reproducible
n_base = 20 # Number of days for baseline
n_treat = 40 # Number of days for treatment
base = rnorm(n_base, mean = 0.55, sd = 0.1) * 100
treat = rnorm(n_treat, mean = 0.6, sd = 0.1) * 100
data = c(base, treat)
time = seq(from = 1, to = length(data))
condition = c(rep("Baseline", n_base), rep("Treatment", n_treat))
df = data.frame(data, time, condition)

# Make plot
```

```r
ggplot(data = df, aes(x = time, y = data, group = condition)) +
  geom_line() +
  geom_point() +
  expand_limits(x = 0, y = 0) +
  xlab("Day") +
  ylab("% Students Studying") +
  ggtitle("Simulation 5: Treatment after 20 days")
```




```r
t.test(base, treat, var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  base and treat
## t = -3.868, df = 58, p-value = 0.0002802
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -15.908821  -5.058216
## sample estimates:
## mean of x mean of y
##  52.19442  62.67794
```

Question 7

Using the t-test, both simulations lead to the conclusion that the differences between the baseline and treatment are significant. Should these results be trusted over visual inspection?

Analysis of Variance

When comparing three or more groups, analysis of variance may also be used. However, there are limitations (see Question 7).

Why isn't statistical inference used?

Inferential statistics are limited because they require large sample sizes. What may work for a handful of subjects may not generalize well for the population (by handful, we literally mean a couple of subjects, far less than 30). Additionally, there may conflicting conclusions between statistical tests and visual inspection (see Question 7).

Question 8

What is the difference between statistical significance (determined by t-test or ANOVA) and practical significance (determined by visual inspection)? Which one should you trust more?

Credits

The material for this worksheet was based on this source.

(Click on the word "source" to access the link)