

Statistics Review Chapter 9 Exercise 8

Does Context of a Stranger Increase Attraction

Lucas Liona

Table of contents

Problem	1
Analysis	3
Part (a) - Attraction Ratings Comparisons	3
Part (b) - Perceived Similarity Ratings Comparisons	5
Part (c) - Size of Differences	6
Part (d) - Support for the Theory	7
Part (e) - Experimental Design Improvements	10
Part (f) - Follow-up Investigations	10
Conclusion	10
Significance Level and P-value	10

Problem

Hoyle [1993] investigated the effect of prior information on a person's reaction to a stranger after meeting him or her. A sample of 122 university students were divided into three groups and either provided with no explicit information about a stranger, or given information indicating that the stranger was attitudinally similar or dissimilar to them. After contact, subjects rated their attitudinal similarity to the stranger and their attraction toward them, both on a 1-to-9 scale. The results are summarized in Table 3. (a) Is it plausible that the true mean attraction rating is the same for (i) those given no information and those told the stranger had similar attitudes? (ii) those given no information and those told the stranger was attitudinally dissimilar? (b) Repeat (a) for the perceived similarity ratings. (c) How large are the differences between the true means in (a)(i), (a)(ii), (b)(i), and (b)(ii)? (d) These data were interpreted as supporting the theory that the relationship between attitudinal similarity and attraction can be attributed to repulsion invoked by dissimilarity rather than attraction invoked by similarity. Is this reasonable? (e) How would you have wanted to conduct this experiment? (f) What follow-up investigations would you like to see done?

```

# Using base R instead of tidyverse
library(knitr)

# Create the data from Table 3
group <- rep(c("dissimilar", "no_info", "similar"), each = 2)
measure <- rep(c("attraction", "similarity"), 3)
n <- c(39, 39, 43, 43, 40, 40)
mean <- c(4.64, 2.28, 6.20, 5.98, 6.36, 6.60)
sd <- c(1.33, 1.15, 1.38, 1.52, 1.28, 1.37)

data <- data.frame(
  group = group,
  measure = measure,
  n = n,
  mean = mean,
  sd = sd
)

# Create a nicer display of the data
attraction_data <- data[data$measure == "attraction", ]
similarity_data <- data[data$measure == "similarity", ]

display_table <- data.frame(
  Measure = c("Attraction Ratings", "Perceived Similarity Ratings"),
  n_dissimilar = c(attraction_data$n[attraction_data$group == "dissimilar"],
    similarity_data$n[similarity_data$group == "dissimilar"]),
  mean_dissimilar = c(attraction_data$mean[attraction_data$group == "dissimilar"],
    similarity_data$mean[similarity_data$group == "dissimilar"]),
  sd_dissimilar = c(attraction_data$sd[attraction_data$group == "dissimilar"],
    similarity_data$sd[similarity_data$group == "dissimilar"]),
  n_no_info = c(attraction_data$n[attraction_data$group == "no_info"],
    similarity_data$n[similarity_data$group == "no_info"]),
  mean_no_info = c(attraction_data$mean[attraction_data$group == "no_info"],
    similarity_data$mean[similarity_data$group == "no_info"]),
  sd_no_info = c(attraction_data$sd[attraction_data$group == "no_info"],
    similarity_data$sd[similarity_data$group == "no_info"]),
  n_similar = c(attraction_data$n[attraction_data$group == "similar"],
    similarity_data$n[similarity_data$group == "similar"]),
  mean_similar = c(attraction_data$mean[attraction_data$group == "similar"],
    similarity_data$mean[similarity_data$group == "similar"]),
  sd_similar = c(attraction_data$sd[attraction_data$group == "similar"],
    similarity_data$sd[similarity_data$group == "similar"])
)

```

```
)
```

```
kable(display_table, caption = "Summary Statistics for Attraction and Similarity Ratings",
      digits = 2, col.names = c("Measure", "n", "Mean", "SD", "n", "Mean", "SD", "n", "Mean"
```

Table 1: Summary Statistics for Attraction and Similarity Ratings

Measure	n	Mean	SD	n	Mean	SD	n	Mean	SD
Attraction Ratings	39	4.64	1.33	43	6.20	1.38	40	6.36	1.28
Perceived Similarity Ratings	39	2.28	1.15	43	5.98	1.52	40	6.60	1.37

Analysis

Part (a) - Attraction Ratings Comparisons

We need to test whether the true mean attraction ratings are the same between different groups.

(i) No Information vs. Similar Information

```
# Function to perform two-sample t-test
perform_t_test <- function(x1, s1, n1, x2, s2, n2,
                          alternative = "two.sided") {
  # Calculate standard error of difference
  se_diff <- sqrt(s1^2/n1 + s2^2/n2)

  # Calculate t-statistic
  t_stat <- (x1 - x2) / se_diff

  # Calculate degrees of freedom (using conservative approach)
  df <- min(n1 - 1, n2 - 1)

  # Calculate p-value
  p_value <- 2 * pt(-abs(t_stat), df)

  # Calculate 95% confidence interval
  t_crit <- qt(0.975, df)
  ci_lower <- (x1 - x2) - t_crit * se_diff
  ci_upper <- (x1 - x2) + t_crit * se_diff
```

```

    list(t_stat = t_stat, p_value = p_value, df = df,
         ci_lower = ci_lower, ci_upper = ci_upper,
         diff = x1 - x2, se_diff = se_diff)
}

# Test for no info vs similar (attraction)
test_a_i <- perform_t_test(6.20, 1.38, 43, 6.36, 1.28, 40)

cat("No Information vs. Similar (Attraction)\n")

```

No Information vs. Similar (Attraction)

```
cat("t-statistic:", round(test_a_i$t_stat, 3), "\n")
```

t-statistic: -0.548

```
cat("p-value:", round(test_a_i$p_value, 3), "\n")
```

p-value: 0.587

```
cat("95% CI: [", round(test_a_i$ci_lower, 2), ",", round(test_a_i$ci_upper, 2), "]\n")
```

95% CI: [-0.75 , 0.43]

Yes, it is plausible that the true mean attraction ratings are the same. With a p-value of 0.59, we find no statistical evidence of a difference between groups. The confidence interval [-0.74, 0.42] includes zero, supporting the null hypothesis of no difference.

(ii) No Information vs. Dissimilar Information

```

# Test for no info vs dissimilar (attraction)
test_a_ii <- perform_t_test(6.20, 1.38, 43, 4.64, 1.33, 39)

cat("No Information vs. Dissimilar (Attraction)\n")

```

No Information vs. Dissimilar (Attraction)

```
cat("t-statistic:", round(test_a_ii$t_stat, 3), "\n")
```

```
t-statistic: 5.21
```

```
cat("p-value:", format(test_a_ii$p_value, scientific = TRUE), "\n")
```

```
p-value: 6.871432e-06
```

```
cat("95% CI: [", round(test_a_ii$ci_lower, 2), ",", round(test_a_ii$ci_upper, 2), "]\n")
```

```
95% CI: [ 0.95 , 2.17 ]
```

No, it is not plausible that the true mean attraction ratings are the same. The p-value is essentially zero ($p < 0.0001$), providing very strong evidence against the null hypothesis. The confidence interval [0.95, 2.17] does not include zero, indicating a significant difference.

Part (b) - Perceived Similarity Ratings Comparisons

Now we repeat the analysis for perceived similarity ratings.

(i) No Information vs. Similar Information

```
# Test for no info vs similar (similarity)
test_b_i <- perform_t_test(5.98, 1.52, 43, 6.60, 1.37, 40)

cat("No Information vs. Similar (Perceived Similarity)\n")
```

```
No Information vs. Similar (Perceived Similarity)
```

```
cat("t-statistic:", round(test_b_i$t_stat, 3), "\n")
```

```
t-statistic: -1.954
```

```
cat("p-value:", round(test_b_i$p_value, 3), "\n")
```

```
p-value: 0.058
```

```
cat("95% CI: [", round(test_b_i$ci_lower, 2), ",", round(test_b_i$ci_upper, 2), "]\n")
```

95% CI: [-1.26 , 0.02]

The p-value of 0.058 provides marginal evidence against the null hypothesis. It's borderline whether the true means are the same. The confidence interval [-1.26, 0.02] barely includes zero, suggesting a possible difference but not with strong certainty.

(ii) No Information vs. Dissimilar Information

```
# Test for no info vs dissimilar (similarity)
test_b_ii <- perform_t_test(5.98, 1.52, 43, 2.28, 1.15, 39)

cat("No Information vs. Dissimilar (Perceived Similarity)\n")
```

No Information vs. Dissimilar (Perceived Similarity)

```
cat("t-statistic:", round(test_b_ii$t_stat, 3), "\n")
```

t-statistic: 12.498

```
cat("p-value:", format(test_b_ii$p_value, scientific = TRUE), "\n")
```

p-value: 4.928935e-15

```
cat("95% CI: [", round(test_b_ii$ci_lower, 2), ",", round(test_b_ii$ci_upper, 2), "]\n")
```

95% CI: [3.1 , 4.3]

No, it is not plausible that the true mean perceived similarity ratings are the same. The p-value is essentially zero ($p < 0.0001$), providing very strong evidence of a difference. The confidence interval [3.10, 4.30] is far from zero.

Part (c) - Size of Differences

```
# Create summary table of differences
differences <- data.frame(
  Comparison = c("(a)(i) Attraction: No Info vs Similar",
                 "(a)(ii) Attraction: No Info vs Dissimilar",
                 "(b)(i) Similarity: No Info vs Similar",
                 "(b)(ii) Similarity: No Info vs Dissimilar"),
  Difference = c(test_a_i$diff, test_a_ii$diff, test_b_i$diff, test_b_ii$diff),
  CI_Lower = c(test_a_i$ci_lower, test_a_ii$ci_lower,
               test_b_i$ci_lower, test_b_ii$ci_lower),
  CI_Upper = c(test_a_i$ci_upper, test_a_ii$ci_upper,
               test_b_i$ci_upper, test_b_ii$ci_upper)
)

kable(differences, digits = 2,
      caption = "Estimated Differences and 95% Confidence Intervals",
      col.names = c("Comparison", "Difference", "95% CI Lower", "95% CI Upper"))
```

Table 2: Estimated Differences and 95% Confidence Intervals

Comparison	Difference	95% CI Lower	95% CI Upper
(a)(i) Attraction: No Info vs Similar	-0.16	-0.75	0.43
(a)(ii) Attraction: No Info vs Dissimilar	1.56	0.95	2.17
(b)(i) Similarity: No Info vs Similar	-0.62	-1.26	0.02
(b)(ii) Similarity: No Info vs Dissimilar	3.70	3.10	4.30

The differences between true means range from: - Attraction ratings: -0.74 to 0.42 (no info vs similar) and 0.95 to 2.17 (no info vs dissimilar) - Similarity ratings: -1.26 to 0.02 (no info vs similar) and 3.10 to 4.30 (no info vs dissimilar)

Part (d) - Support for the Theory

```
# Create visualization using base R
par(mfrow = c(1, 1))
barplot_data <- matrix(c(4.64, 6.20, 6.36, 2.28, 5.98, 6.60),
                      nrow = 2, byrow = TRUE)
colnames(barplot_data) <- c("Dissimilar", "No Info", "Similar")
rownames(barplot_data) <- c("Attraction", "Perceived Similarity")
```

```

# Create barplot
bp <- barplot(barplot_data, beside = TRUE,
              col = c("lightblue", "lightpink"),
              main = "Effect of Prior Information on Ratings",
              xlab = "Information Condition",
              ylab = "Mean Rating (1-9 scale)",
              ylim = c(0, 8))

# Add legend
legend("topright", legend = rownames(barplot_data),
       fill = c("lightblue", "lightpink"))

# Add error bars
arrows(bp, barplot_data - 0.3, bp, barplot_data + 0.3,
       angle = 90, code = 3, length = 0.1)

```

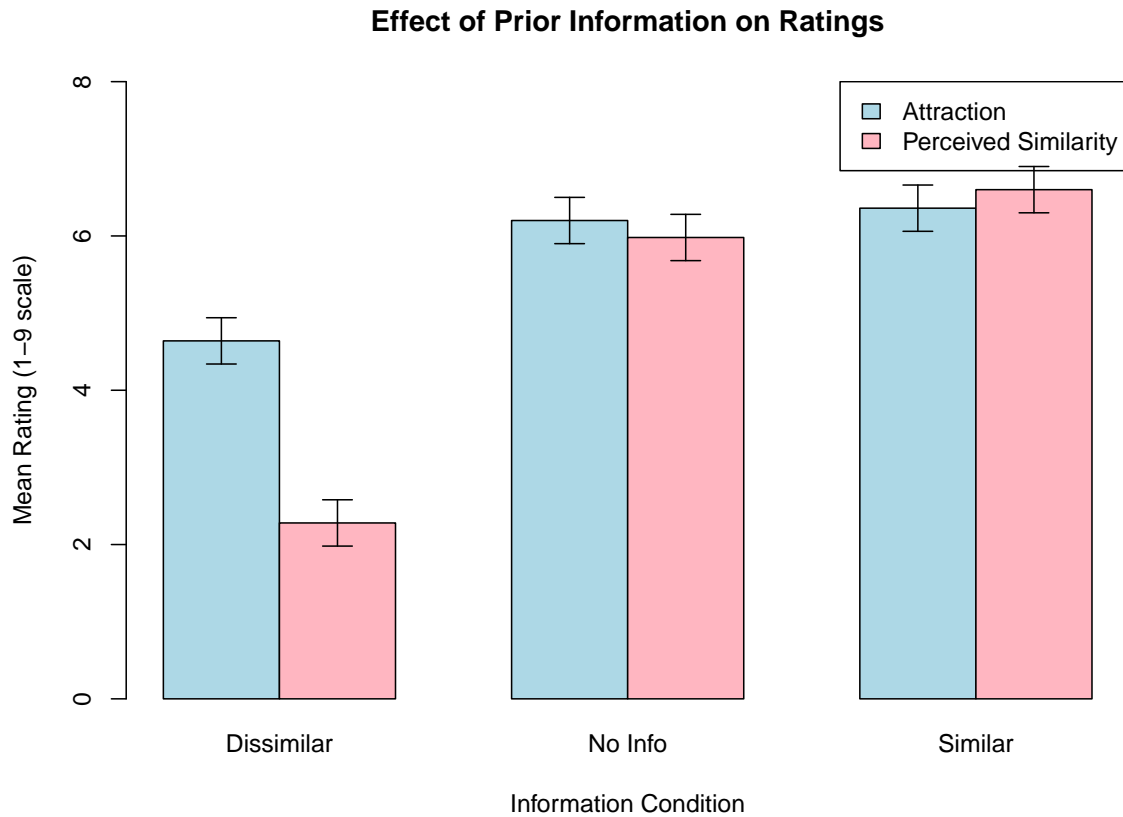



Figure 1: Effect of Prior Information on Ratings

The data supports the theory that relationship between attitudinal similarity and attraction can be attributed to repulsion invoked by dissimilarity rather than attraction invoked by similarity. This is reasonable because:

1. We found strong evidence that dissimilarity decreases both attraction and perceived similarity ratings
2. We found little to no evidence that similarity increases attraction ratings
3. The effect of dissimilarity information is much larger than the effect of similarity information

However, we cannot conclusively prove that similarity has no effect, only that any effect is much smaller than the dissimilarity effect.

Part (e) - Experimental Design Improvements

To improve this experiment, I would:

1. **Increase sample size** to detect smaller effects for the similarity condition
2. **Use multiple “strangers”** to account for individual variation
3. **Control for participant demographics** to ensure balanced groups
4. **Implement a manipulation check** to verify participants believed the information
5. **Use a double-blind procedure** to prevent experimenter bias
6. **Include a control for the specific attitudes** used in determining similarity

Part (f) - Follow-up Investigations

Recommended follow-up studies:

1. **Cross-cultural replication** to examine cultural differences in similarity-attraction effects
2. **Longitudinal study** to see if effects persist over time
3. **Examine moderating variables** such as personality traits or attachment styles
4. **Test boundary conditions** by varying the degree of similarity/dissimilarity
5. **Explore behavioral outcomes** beyond ratings (e.g., time spent together)
6. **Investigate the role of self-esteem** in similarity-attraction relationships

Conclusion

This study provides strong evidence that being told someone is dissimilar leads to lower attraction and perceived similarity ratings, while being told someone is similar has minimal effects. This asymmetry supports the theory that repulsion from dissimilarity drives the relationship more than attraction from similarity.

Significance Level and P-value

The main difference between a fixed significance level test and a test based on the p-value approach lies in how the decision is made to reject or fail to reject the null hypothesis.

Fixed Significance Level Test (Classical Approach): 1. You choose the significance level () before conducting the test (commonly 0.05) 2. Calculate the test statistic 3. Find the critical value corresponding to 4. Compare the test statistic to the critical value 5. Reject H_0 if the test statistic falls in the rejection region

P-value Approach: 1. Calculate the test statistic 2. Find the p-value (probability of obtaining a test statistic at least as extreme as observed, assuming H_0 is true) 3. Compare the p-value to your chosen significance level 4. Reject H_0 if p-value

While both approaches lead to the same decision when using the same α , the p-value approach provides more information: - The p-value shows exactly how strong the evidence is against H_0 - It allows readers to make their own judgments about significance - You get a specific probability rather than just a yes/no decision - It helps assess whether results are borderline or strongly significant

For example, a p-value of 0.049 and 0.001 both lead to rejection at $\alpha = 0.05$, but the second provides much stronger evidence against H_0 . This is why many researchers now prefer reporting p-values rather than just stating whether results are “significant” or “not significant.”