

1                   Reproduction of the Analysis of Holmes, To, & Johnsrude (2021)

2                   Tiffany Hanratty<sup>1</sup>

3                   <sup>1</sup> Brooklyn College

4                   Author Note

5                   Tiffany Hanratty, Department of Psychology, Brooklyn College.

6                   Correspondence concerning this article should be addressed to Tiffany Hanratty, 2900

7 Bedford Ave. E-mail: TIFFANY.HANRATTY66@bcmail.cuny.edu

## Abstract

8  
9 This semester project works to recreate the statistical analysis portion of a study by  
10 Holmes, To, & Johnsrude (2012). This analysis was run on the training portion of the  
11 study. Three paired t-tests were used to analyze the data in order to help determine  
12 significance when comparing different groups. The training portion of this paper analyzes 3  
13 groups: the most familiar condition, the moderately familiar condition, and the least  
14 familiar condition (Holmes, To, and Johnsrude, 2021).

15 *Keywords:* Auditory perception, reproducibility, attention, speech perception,  
16 memory, learning

## Reproduction of the Analysis of Holmes, To, & Johnsrude (2021)

The authors of this study examined how different lengths of voice exposure relate to voice intelligibility and voice recognition (Holmes, To, & Johnsrude, 2021). The study consisted of 50 adult participants, and was broken up into four different parts. This data reanalysis is focused on the training portion of this study. The training portion analyses voice recognition performance in three categories: most familiar, moderately familiar, and least familiar voice. Information on the sections regarding familiarization, the explicit-recognition test, and the speech-intelligibility test can be found in the complete study by Holmes, To, & Johnsrude (2021). The data for this analysis reduction can be found at, and were downloaded from <https://osf.io/7kdyc/>.

## Methods

### Participants

This experiment contained a total of 50 participants which were all aged 18 to 28 (Holmes, To, & Johnsrude, 2021).

### Material

Details pertaining to the acoustic stimuli used in this experiment can be found in the study conducted by Holmes, To, & Johnsrude (2021).

### Procedure

This experiment consisted of familiarization, training, an explicit-recognition test, and a speech-intelligibility test (Holmes, To, and Johnsrude, 2021). This reanalysis pertains to the training portion of this study. The training portion of this study consisted of participants hearing multiple sentences which could be spoken in one of three categories,

39 most familiar, moderately familiar, or least familiar. Some participants ( $N = 25$ ) heard  
40 training sentences alone, while other participants ( $N = 25$ ) heard the sentences with  
41 background noise in the form of babbling. Participants were asked after each sentence to  
42 choose the name of the speaker of the sentence, which corresponds to the familiarity  
43 category. Participants were provided with immediate feedback as to whether or not they  
44 were correct. Detailed information on the remaining three sections of this study can be  
45 found in Holmes, To, & Johnsrude (2021).

## 46 Results

```
47 ## function (dir)
48 ## .Internal(setwd(dir))
49 ## <bytecode: 0x0000000012be9e58>
50 ## <environment: namespace:base>

51 ##

52 ## Paired t-test

53 ##

54 ## data: all_data$fam_1 and all_data$fam_2
55 ## t = 4.2875, df = 49, p-value = 8.447e-05
56 ## alternative hypothesis: true difference in means is not equal to 0
57 ## 95 percent confidence interval:
58 ## 2.036763 5.630437
59 ## sample estimates:
60 ## mean of the differences
61 ## 3.8336

62 ##

63 ## Paired t-test
```

```

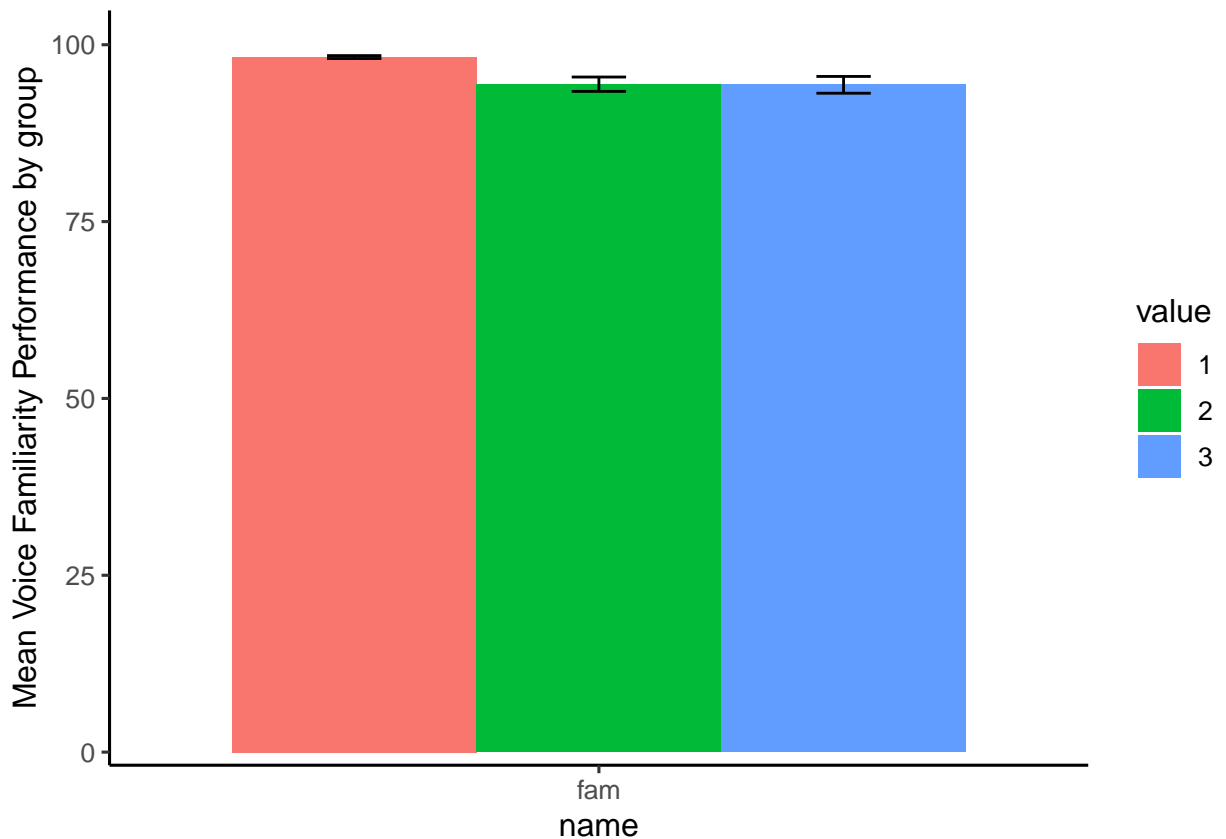
64 ##
65 ## data:  all_data$fam_1 and all_data$fam_3
66 ## t = 3.5619, df = 49, p-value = 0.0008311
67 ## alternative hypothesis: true difference in means is not equal to 0
68 ## 95 percent confidence interval:
69 ##  1.709717 6.136283
70 ## sample estimates:
71 ## mean of the differences
72 ##                               3.923

73 ##
74 ## Paired t-test
75 ##
76 ## data:  all_data$fam_2 and all_data$fam_3
77 ## t = 0.15586, df = 49, p-value = 0.8768
78 ## alternative hypothesis: true difference in means is not equal to 0
79 ## 95 percent confidence interval:
80 ## -1.063256  1.242056
81 ## sample estimates:
82 ## mean of the differences
83 ##                               0.0894

```

84

value	name	meanRT	SEMRT
1	fam	98.2568	0.2078432
2	fam	94.4232	1.0142461
3	fam	94.3338	1.1859128



85

## 86 Hand Reporting

87 A paired t-test was used to compare familiarity conditions with performance. Better  
 88 performance was seen in the most familiar condition compared to the moderately familiar  
 89 condition,  $t(49) = 4.29$ ,  $p < .001$  (Holmes, To, and Johnsrude, 2021). There was also  
 90 better performance seen in the most familiar condition compared to the least familiar  
 91 condition,  $t(49) = 3.56$ ,  $p = .001$  (Holmes, To, and Johnsrude, 2021). There was not a  
 92 significant difference found in the moderately familiar group compared to the least familiar  
 93 group,  $t(49) = 0.16$ ,  $p = .88$ . The confidence interval values were not able to be reproduced  
 94 in this reanalysis.

## Papaja Reporting

A paired t-test was used to compare familiarity conditions with performance. Better performance was seen in the most familiar condition compared to the moderately familiar condition,  $M_d = 3.83$ , 95% CI [2.04, 5.63],  $t(49) = 4.29$ ,  $p < .001$  (Holmes, To, and Johnsrude, 2021). There was also better performance seen in the most familiar condition compared to the least familiar condition,  $M_d = 3.92$ , 95% CI [1.71, 6.14],  $t(49) = 3.56$ ,  $p = .001$  (Holmes, To, and Johnsrude, 2021). There was not a significant difference found in the moderately familiar group compared to the least familiar group,  $M_d = 0.09$ , 95% CI [-1.06, 1.24],  $t(49) = 0.16$ ,  $p = .877$ . The confidence interval values were not able to be reproduced in this reanalysis.

## Discussion

The analysis which was reported by Holmes, To, & Johnsrude (2021) was partially successfully reproduced in this reanalysis. Both the t-test values and the p-values in this reanalysis match the reported values in Holmes, To, & Johnsrude (2021). The confidence interval values were not able to be replicated in this reanalysis. A simulation based power analysis will be completed in the following section.

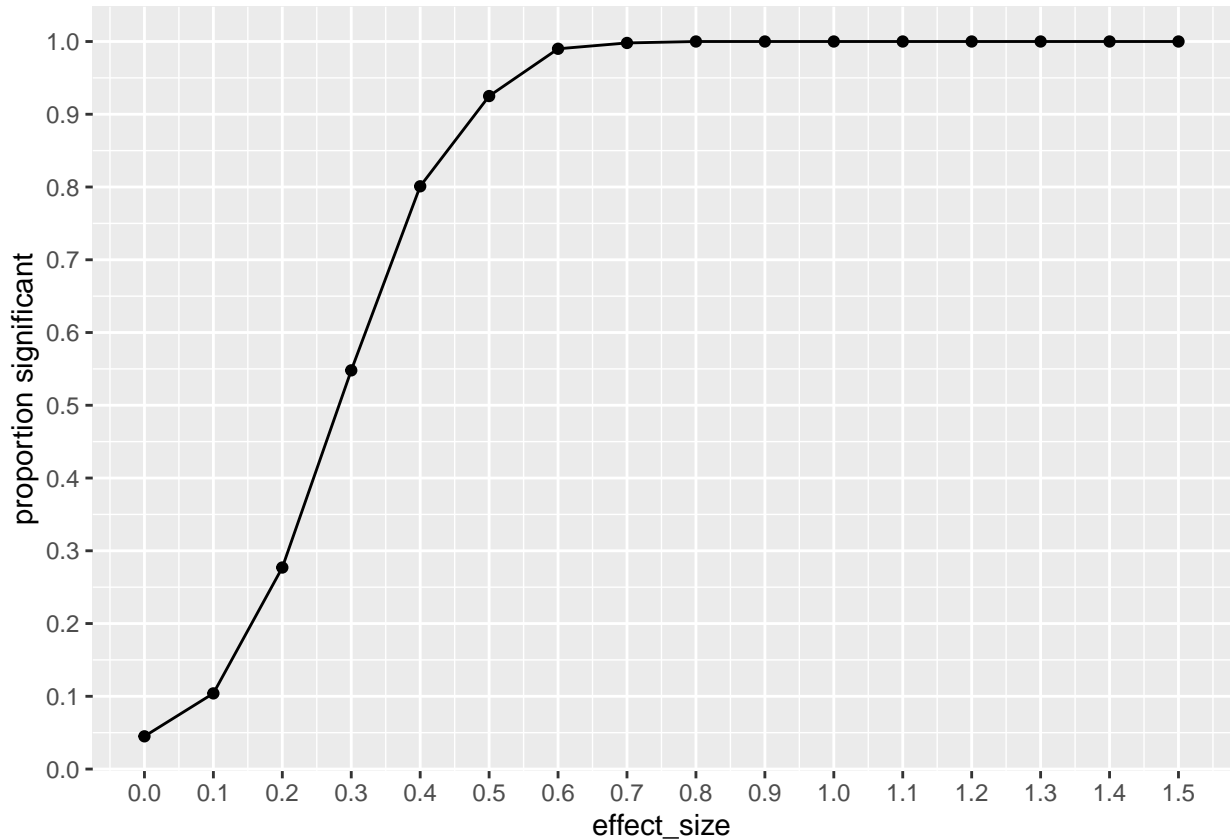
## Simulation-based power analysis

This study ran three t-tests comparing performance between different groups of voice familiarity. The t-tests consisted of performance comparison between the most familiar voice compared to a moderately familiar voice, the most familiar voice compared to the least familiar voice, and the moderately familiar voice compared to the least familiar voice.

This power analysis was conducted in RStudio. For this power analysis a variable was generated in order to indicate the possible occurring effect sizes in this study. A simulation was run for each possible effect size using 1000 simulated experiments. Values were

sampled using an approximately normal distribution. The replicated values' mean is set to be equal to the effect size. Fifty values were run for this distribution to emulate the fifty participants in the experiment. Every time this simulated test is conducted, a p-value is measured. This test will then analyze how many times these simulated experiments produced a p-value at a significance value of  $p < 0.05$ .

Based on the findings of this power curve, when the effect size is at approximately 0.7, a design similar to this studies design will detect an effect at  $p < 0.05$  close to 100% of the time. If the power of a study needs to be increased, a relatively simple way to increase statistical power would be to increase the sample size.





**References**

129

- 130 Holmes, E., To, G., & Johnsrude, I. S. (2021). How Long Does It Take for a Voice to  
131 Become Familiar? Speech Intelligibility and Voice Recognition Are Differentially Sensitive  
132 to Voice Training. *Psychological science*, 32(6), 903–915.  
133 <https://doi.org/10.1177/0956797621991137>