Анализ результатов эксперимента (ГМ)

## 0. Подготовка

Ставим пакет:

library(tidyverse)

Загружаем данные и смотрим их структуру:

my\_df <- read\_csv2("../data/final results.csv")  
head(my\_df)

## # A tibble: 6 x 7  
## id sex example.audio example.video answer order v\_a  
## <int> <chr> <chr> <chr> <chr> <int> <chr>  
## 1 1 f p k p 1 v  
## 2 1 f k t k 1 v  
## 3 1 f t t t 1 v  
## 4 1 f k k k 1 v  
## 5 1 f p t p 1 v  
## 6 1 f t p t 1 v

Переведем некоторые переменные в факторы, для корректного отображения:

my\_df$example.audio <- factor(my\_df$example.audio, levels = c("p", "t", "k"))  
my\_df$example.video <- factor(my\_df$example.video, levels = c("p", "t", "k"))  
my\_df$answer <- factor(my\_df$answer, levels = c("-", "p", "t", "k"))

## 1. Визуализация

Сперва проанализируем данные только по тем участникам, которые и смотрели видео, и слушали аудио:

my\_df %>%  
 filter(v\_a == "v") %>%   
 count(sex, example.audio, example.video, answer, sort = TRUE) %>%   
 ggplot(aes(example.audio, example.video, size = n))+  
 geom\_point(aes(size = n), alpha=0.8, color="darkblue", show.legend =FALSE) +  
 geom\_text(aes(label = n), color="white", size = 5) +  
 scale\_size(range = c(4,16)) +  
 facet\_grid(sex~answer)+  
 theme\_bw()+  
 labs(x = "аудио стимулы",  
 y = "видео стимулы",  
 title = "Результаты сгрупированные по ответу и полу")



График разделен на 4 столбца (ответы респондентов) и 2 строки (пол респондентов). Как видно из графика, распределение ответов, видимо, связано с аудио стимулом, а не видео стимулом. Единственное отличие --- многие респонденты ответили, что слышат k, в случае аудио стимула t и видео стимула p. Данный эффект не является эффектом МакГурка, однако нужно показать его статистическую значимость.  
Посмотрим на аналогичные результаты для тех респондентов, которые лишь слушали аудио стимулы, но не смотрели видео стимулы.

my\_df %>%  
 filter(v\_a == "a") %>%   
 count(sex, example.audio, example.video, answer, sort = TRUE) %>%   
 ggplot(aes(example.audio, example.video, size = n))+  
 geom\_point(aes(size = n), alpha=0.8, color="darkblue", show.legend =FALSE) +  
 geom\_text(aes(label = n), color="white", size = 5) +  
 scale\_size(range = c(4,16)) +  
 facet\_grid(sex~answer)+  
 theme\_bw()+  
 labs(x = "аудио стимулы",  
 y = "видео стимулы",  
 title = "Результаты сгрупированные по ответу и полу")



Появляется ощущение, что все носители отвечают одинаково вне зависимости от того, видят ли они видео стимул или нет. Из этого мы можем сделать вывод, что наблюдаемый эффект связан с особенностью нашего стимула, а не вызван каким-то перцептивным эффектом связанным с взаимодействием аудио и видео стимулов. Перечислим некоторые факты, которые мы хотим доказать:

* разница в ответах между респондентами разного пола не является статистически значимой
* разница в ответах между респондентами смотревших видео стимулы и не смотревших оных не является статистически значимой
* ответы респондентов в нашем эксперименте скорее совпадают с аудио стимулом, а не с видео стимулом.

## 2.1 Влияет ли пол на ответы?

Начнем с того, что покажем, что пол не влияет на распределение ответов. Для этого составим таблицу с уникальными значениями видео стимулов, аудио стимулов, ответов и переменной, кодирующей смотрели ли данный респондент видео стимул, в строках и полом респондента в столбцах.

my\_df %>%  
 na.omit() %>%   
 count(example.audio, example.video, answer, v\_a, sex) %>%   
 spread(key = sex, value = n) %>%   
 mutate(f = ifelse(is.na(f), as.integer(0), f),  
 m = ifelse(is.na(m), as.integer(0), m),  
 type = paste(example.audio, example.video, answer, v\_a)) %>%   
 ungroup() %>%   
 select(type, f, m) ->  
 my\_df\_sex  
head(my\_df\_sex)

## # A tibble: 6 x 3  
## type f m  
## <chr> <int> <int>  
## 1 p p p a 15 15  
## 2 p p p v 60 60  
## 3 p t - a 1 0  
## 4 p t p a 14 14  
## 5 p t p v 55 53  
## 6 p t t a 0 1

Применим к нашим данным тест Фишера:

fisher.test(my\_df\_sex[, -1], simulate.p.value=TRUE,B=1e7)

##   
## Fisher's Exact Test for Count Data with simulated p-value (based  
## on 1e+07 replicates)  
##   
## data: my\_df\_sex[, -1]  
## p-value = 0.926  
## alternative hypothesis: two.sided

Как видно из очень выского p-value (0.93 > 0.05), у нас нет оснований отбросить нулевую гепотизу о том, что пол респондента не влияет на ответы.

## 2.2 Влияет ли просмотр видео стимула на ответы?

Cоставим таблицу с уникальными значениями видео стимулов, аудио стимулов и ответов в строках и переменной, кодирующей смотрели ли данный респондент видео стимул, в столбцах.

my\_df %>%  
 na.omit() %>%   
 count(example.audio, example.video, answer, v\_a) %>%   
 spread(key = v\_a, value = n) %>%   
 mutate(v = ifelse(is.na(v), as.integer(0), v),  
 a = ifelse(is.na(a), as.integer(0), a),  
 type = paste(example.audio, example.video, answer)) %>%   
 ungroup() %>%   
 select(type, v, a) ->  
 my\_df\_v\_a  
head(my\_df\_v\_a)

## # A tibble: 6 x 3  
## type v a  
## <chr> <int> <int>  
## 1 p p p 120 30  
## 2 p t - 0 1  
## 3 p t p 108 28  
## 4 p t t 9 1  
## 5 p t k 3 0  
## 6 p k - 2 0

Применим к нашим данным тест Фишера:

fisher.test(my\_df\_v\_a[, -1], simulate.p.value=TRUE,B=1e7)

##   
## Fisher's Exact Test for Count Data with simulated p-value (based  
## on 1e+07 replicates)  
##   
## data: my\_df\_v\_a[, -1]  
## p-value = 0.7243  
## alternative hypothesis: two.sided

Как видно из очень выского p-value (0.66 > 0.05), у нас нет оснований отбросить нулевую гепотизу о том, что просмотр видео стимулов влияет на ответы респондентов. В результате, мы убеждаемся, что наблюдаемый эффект (ответы k, при видео стимуле p и при аудио стимуле t) связан с особенностью нашего стимула, а не вызван каким-то перцептивным эффектом связанным с взаимодействием аудио и видео стимулов.

## 2.3 Каке стимулы больше влияют на ответы: видео или аудио?

Создадим два столбца eq\_a, если ответ респондента совпадает с аудио стимулом, но не с видео стимулом, и eq\_v, если ответ респоднта совпадает с видео стимулом, но не с аудио стимулом:

my\_df <- read\_csv2("../data/final results.csv")  
my\_df %>%  
 na.omit() %>%   
 count(example.audio, example.video, answer) %>%   
 filter(example.audio != example.video) %>%   
 mutate(answer = ifelse(example.audio == answer, "eq\_a", answer),  
 answer = ifelse(example.video == answer, "eq\_v", answer)) %>%   
 filter(answer == "eq\_a"|answer == "eq\_v") %>%   
 spread(key = answer, value = n) ->  
 my\_df\_equalities  
head(my\_df\_equalities)

## Source: local data frame [6 x 4]  
## Groups: example.audio, example.video [6]  
##   
## # A tibble: 6 x 4  
## example.audio example.video eq\_a eq\_v  
## <chr> <chr> <int> <int>  
## 1 k p 123 18  
## 2 k t 118 15  
## 3 p k 139 8  
## 4 p t 136 10  
## 5 t k 148 1  
## 6 t p 43 7

В результате, получилось 6 контекстов, в каждом из которых случаев совпадений ответа респондента с аудио стимулом в более чем пять раз превосходит количество случаев совпадений ответа респондента с видео стимулом. Данное различие для любых пар чисел больше 1 будет давать статистически значимое различие при использовании метода .

## 2.4 Влияние порядка

Чтобы показать, что порядок не влияет необходимо составить таблицу совсеми возможными порядками и одним столбцом со всеми возможными комбинациями ответов и стимулов.

my\_df %>%  
 na.omit() %>%   
 count(example.audio, example.video, answer, order) %>%   
 spread(key = order, value = n) %>%   
 mutate(`1` = ifelse(is.na(`1`), as.integer(0), `1`),  
 `2` = ifelse(is.na(`2`), as.integer(0), `2`),  
 `3` = ifelse(is.na(`3`), as.integer(0), `3`),  
 type = paste(example.audio, example.video, answer)) %>%   
 ungroup() %>%   
 select(type, `1`, `2`, `3`) ->  
 my\_df\_groups  
head(my\_df\_groups)

## # A tibble: 6 x 4  
## type `1` `2` `3`  
## <chr> <int> <int> <int>  
## 1 k k k 49 49 49  
## 2 k k p 0 0 1  
## 3 k k t 1 1 0  
## 4 k p - 4 0 1  
## 5 k p k 36 43 44  
## 6 k p p 7 6 5

fisher.test(my\_df\_groups[, -1], simulate.p.value=TRUE,B=1e7)

##   
## Fisher's Exact Test for Count Data with simulated p-value (based  
## on 1e+07 replicates)  
##   
## data: my\_df\_groups[, -1]  
## p-value = 0.6489  
## alternative hypothesis: two.sided

И опять точный Тест Фишера (p-value = 0.64 > 0.05) показал, что у нас нет оснований отвергать нулевую гипотезу о том, что порядок не влияет на рспределение ответов.