

Entrega Final Taller Econometría II

Marco Ramos

Contents

| | |
|-------------------------------------|-----------|
| Ambiente. | 2 |
| I. | 2 |
| II. | 2 |
| III. | 4 |
| IV. | 4 |
| Variable de clasificación | 4 |
| Variables de respuesta | 5 |
| Covariables | 7 |
| V. | 8 |
| VI. | 9 |
| VII. | 10 |
| VIII | 12 |
| IX | 14 |
| X. | 16 |

Ambiente.

```
#carga librerias
library(readr)
library(stargazer)
library(dplyr)
library(ggplot2)
library(ggpubr)
library(RCT)
library(gridExtra)
library(kableExtra)
library(lmtest)

#leeo los datos
data<-read_csv("fake.csv")
```

I.

En este trabajo se intentará explorar el efecto que tiene la percepción subjetiva de credibilidad en el hecho de aceptar una noticia como verdadera y en consecuencia compartirla en redes sociales. Concretamente, esta investigación averiguará el efecto que tiene la exposición a noticias y fuentes falsas y verdaderas en el hecho de que una persona crea que es verdadera y que la comparta. Para ellos se realizará un análisis de inferencia causal sobre los datos de un experimento aleatorizado. Los datos se basan en un experimento de encuesta en el que los individuos fueron asignados aleatoriamente a dos versiones de un informe de noticias en el que solo se varió el nombre de la fuente (y el diseño). Se intentará cuantificar los efectos de la veracidad de la fuente en la credibilidad del contenido.

En particular las hipótesis son ²¹:

- Las personas tienden a creer más en las noticias de una fuente real que de una fuente falsa.
- Las personas serán más propensas a compartir informes de noticias que provienen de fuentes reales que aquellos que provienen de fuentes falsas.

II.

Teoría de cambio

```
DiagrammeR::grViz("digraph {
  graph [layout = dot, rankdir = TB]

  node [shape = rectangle]
  rec1 [label = 'Exposición a noticia',color=grey]
  rec2 [label = 'Percepción de veracidad de la fuente',color=yellow]
  rec4 [label = 'Noticia no creída',color=red]
  rec5 [label = 'Noticia creída',color=limegreen]
  rec6 [label = 'Noticia no compartida',color=red]
  rec7 [label = 'Noticia compartida',color=limegreen]

  # edge definitions with the node IDs

  rec2 -> rec4-> rec6
  rec1 -> rec2 -> rec5-> rec7
```

¹originalmente eran 4 pero dado los datos que tenemos, no eran relevantes ni posible evaluar las últimas dos

```
}",  
height = 500)
```

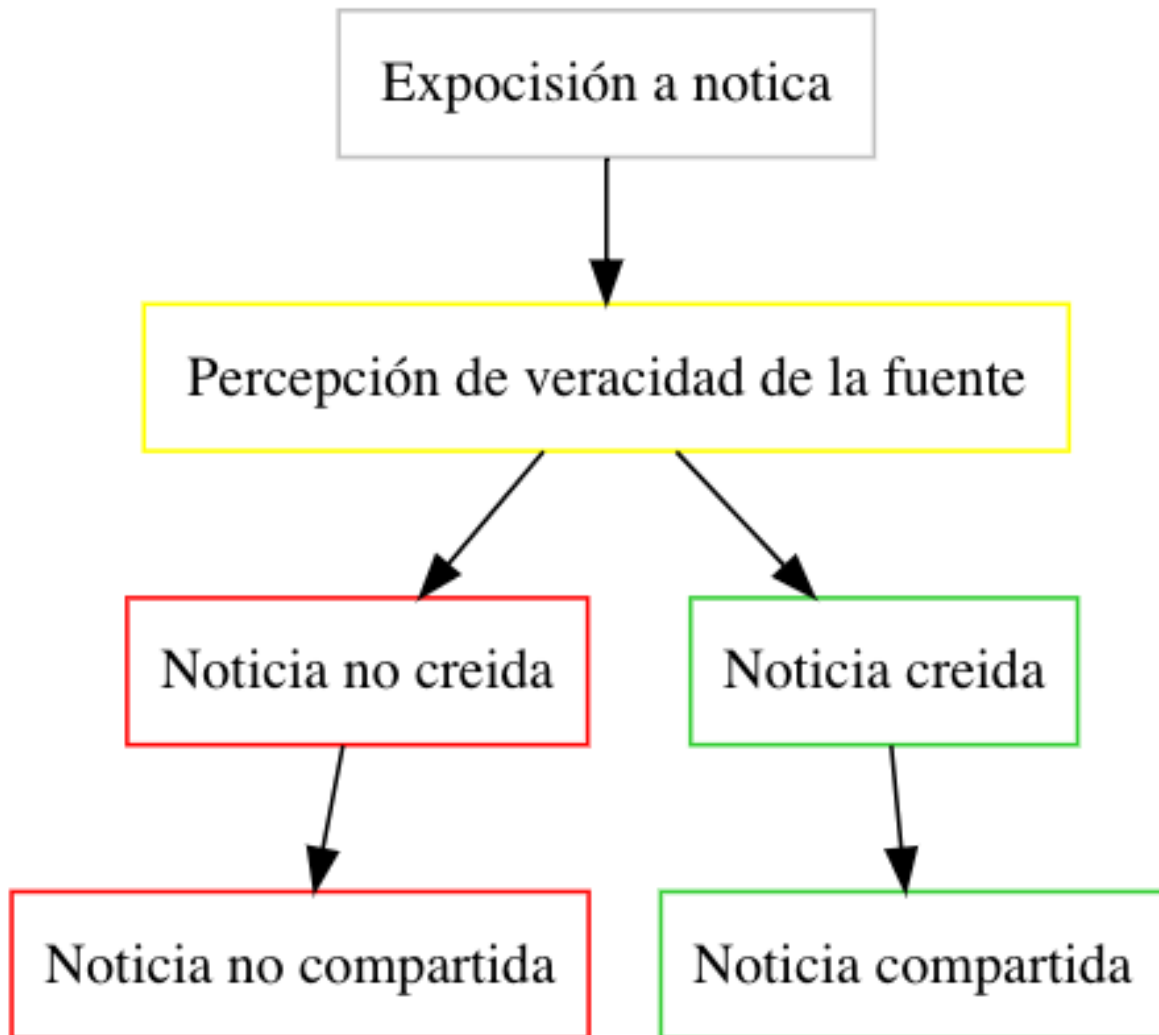


Figure 1: DAG

III.

Participantes

Se trata de 1,980 individuos de Alemania, preseleccionados de una muestra supuestamente aleatoria y representativa (una empresa externa se encargó de eso) que contestaron una encuesta online.

Variables de respuesta

Había dos variables de respuesta; por un lado *y_belief_report* que intenta medir en una escala del 0 al 6 el nivel de credibilidad de la noticia; y por el otro lado, *y_share_report*, que indica un score sobre qué tanto se compartiría dicha noticia en redes sociales.

Clasificación en grupos

Se dividió en grupos de acuerdo al tipo de noticia al que la persona fue expuesto: tratamiento, si la fuente de la noticia era falso; y control, si la fuente era verdadera.

Covariables

Las covariables son la edad, el sexo, el nivel de ingreso (medido en 11 niveles) y el nivel de educación (medido en 5)

¿Cómo se implementó el experimento?

Los individuos fueron asignados aleatoriamente a dos versiones de un informe de noticias en el que solo se varió el nombre de la fuente (y el diseño).

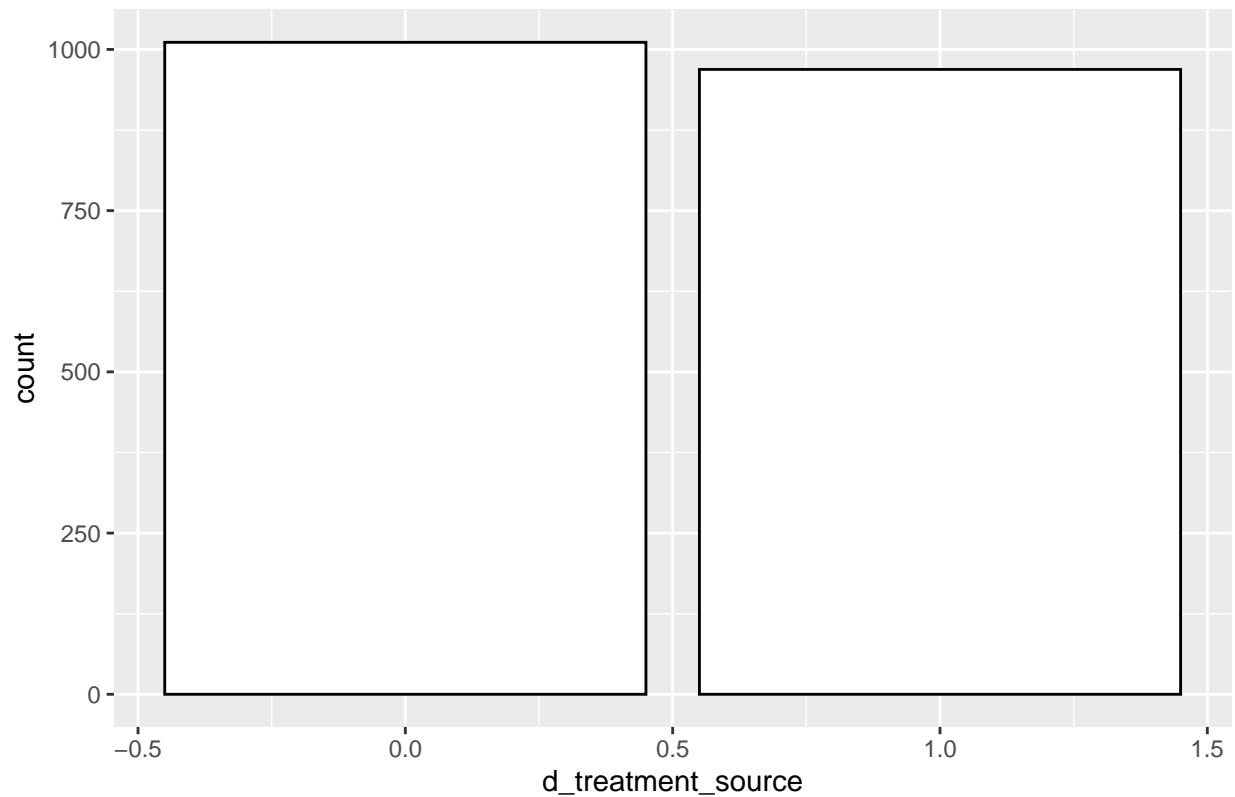
IV.

Para el EDA se optó por realizar un análisis gráfico:

Variable de clasificación

```
#carga librerías
ggplot(data, aes(x=d_treatment_source)) +
  geom_bar(colour="black", fill="white")+ggtitle("Grupos entre control y tratamiento")
```

Grupos entre control y tratamiento

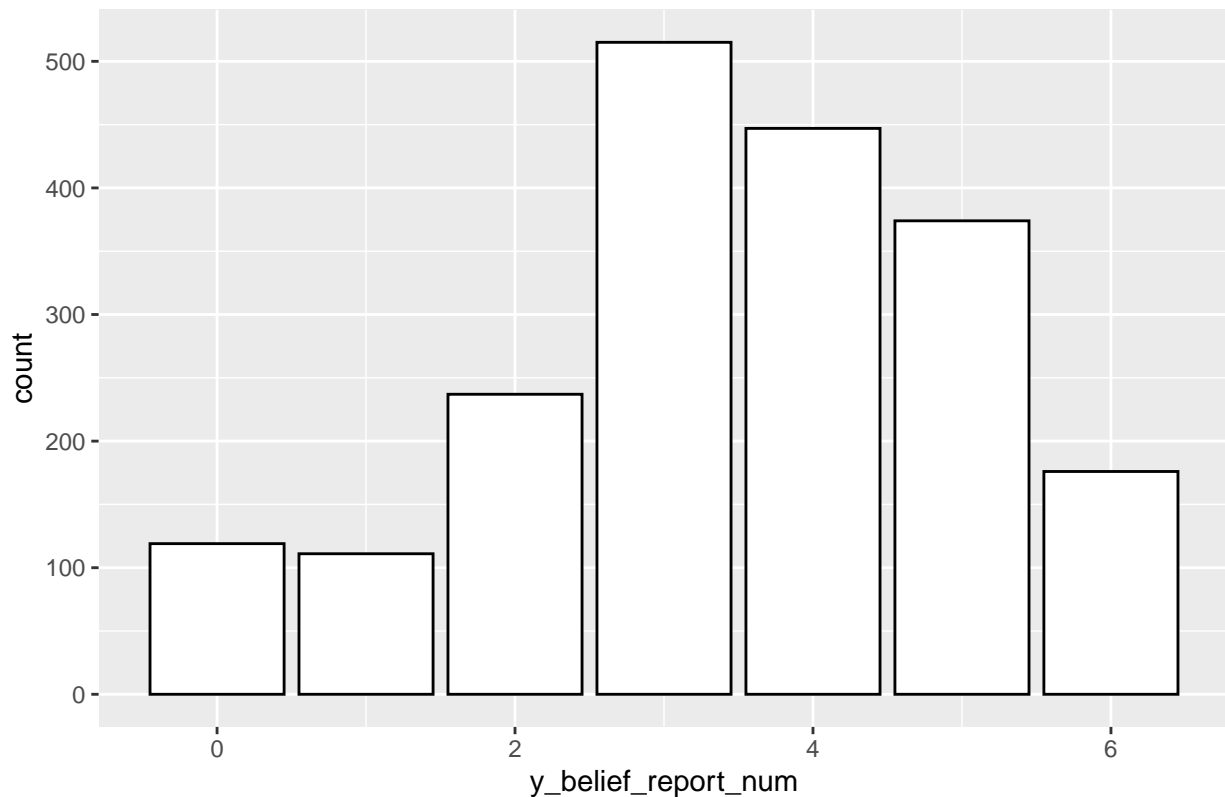


Lo que podemos ver es que la repartición entre los grupos fue muy similar.

Variables de respuesta

```
#carga librerias  
ggplot(data, aes(x=y_belief_report_num)) +  
  geom_bar(colour="black", fill="white")+ggtitle("Frecuencia de indice de credibilidad (0-6)")
```

Frecuencia de índice de credibilidad (0–6)



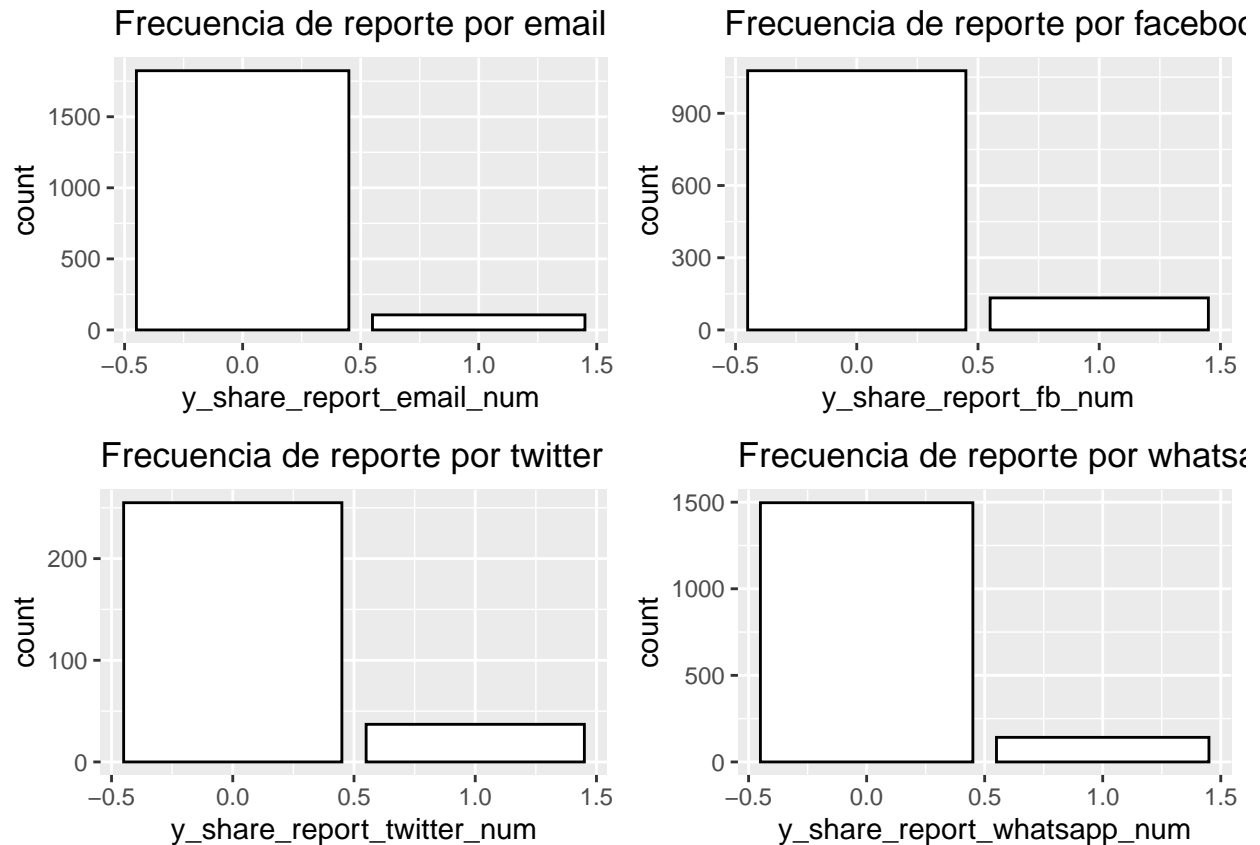
Notamos que el índice de credibilidad está sesgado a la derecha. Dado que los grupos de control y tratamiento son de aproximadamente del mismo tamaño, esto indica que hay una tendencia mayor a creer en la fuente.

```
a<-ggplot(data, aes(x=y_share_report_email_num)) +
  geom_bar(colour="black", fill="white")+ggtitle("Frecuencia de reporte por email")

b<-ggplot(data, aes(x=y_share_report_fb_num)) +
  geom_bar(colour="black", fill="white")+ggtitle("Frecuencia de reporte por facebook")

c<-ggplot(data, aes(x=y_share_report_twitter_num)) +
  geom_bar(colour="black", fill="white")+ggtitle("Frecuencia de reporte por twitter")

d<-ggplot(data, aes(x=y_share_report_whatsapp_num)) +
  geom_bar(colour="black", fill="white")+ggtitle("Frecuencia de reporte por whatsapp")
ggarrange(a,b,c,d,
  ncol = 2, nrow = 2)
```



Asimismo, en cuanto a las variables de compartir la noticia por plataforma, notamos que en general la gente no comparte contenidos. Sin embargo, dado el balance de los grupos, notamos que la gente se atreve a compartir noticias ligeramente más a través de las redes sociales facebook y twitter.

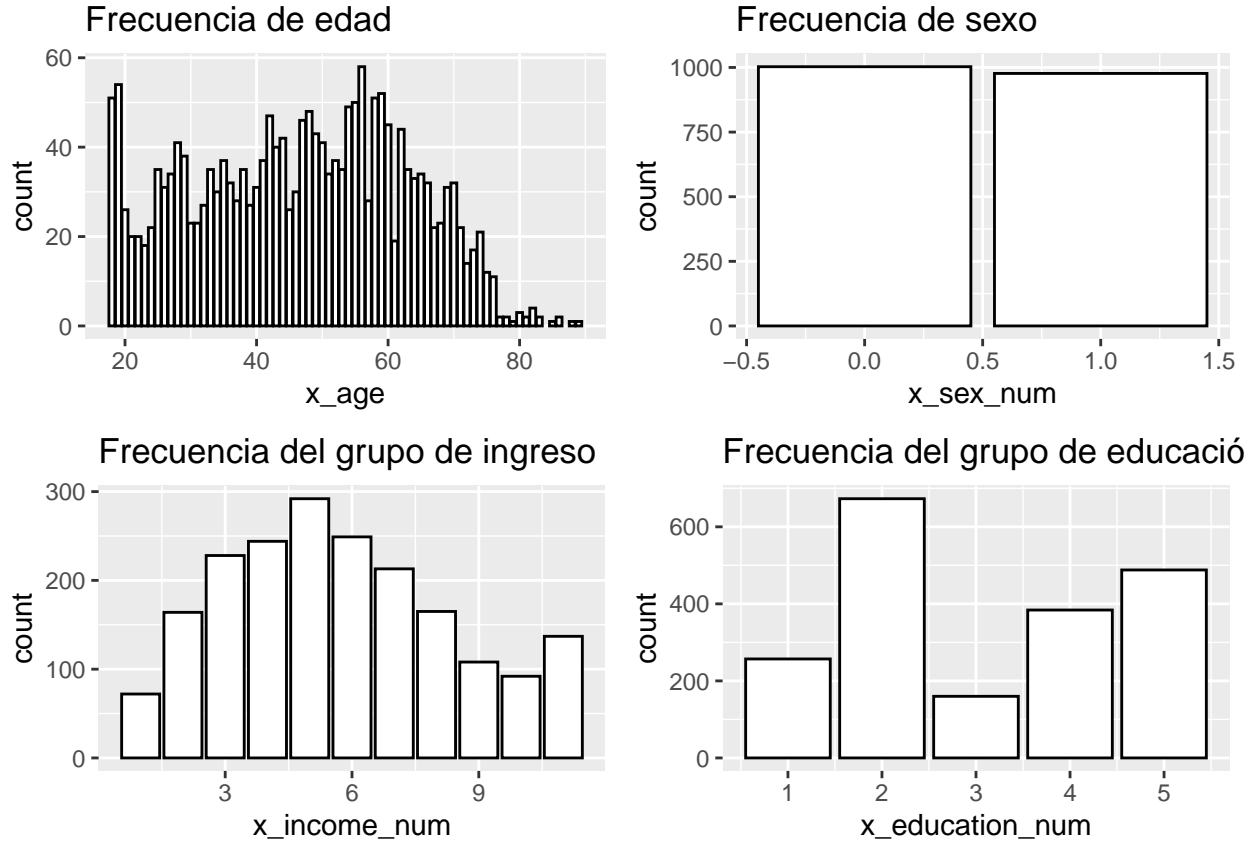
Covariables

```
a<-ggplot(data, aes(x=x_age)) +
  geom_bar(colour="black", fill="white")+ggtitle("Frecuencia de edad")

b<-ggplot(data, aes(x=x_sex_num)) +
  geom_bar(colour="black", fill="white")+ggtitle("Frecuencia de sexo")
c<-ggplot(data, aes(x=x_income_num)) +
  geom_bar(colour="black", fill="white")+ggtitle("Frecuencia del grupo de ingreso")

d<-ggplot(data, aes(x=x_education_num)) +
  geom_bar(colour="black", fill="white")+ggtitle("Frecuencia del grupo de educación")

ggarrange(a,b,c,d,
  ncol = 2, nrow = 2)
```



Finalmente notamos que en cuanto a edades, en general hay buen balance, con dos picos muy claros en menores a 20 y alrededor del 55. En cuanto a sexo, los grupos son virtualmente del mismo tama1o. En cuanto a grupo de ingreso, la media es en el grupo 5 con una distribuci3n con una cola a la derecha larga. Finalmente, en cuanto a grupo de educaci3n, el nivel 2 es el m1s frecuente, seguido del 5, con el grupo 3 como grupo m1s peque1o, lo cual demuestra que es una distribuci3n poco sim3trica y muy peculiar.

V.

Para esta parte no contamos con los datos suficientes para replicar las tablas debido a que no se nos proporcionaron datos de la variable content ni de las de actitud sobre la migraci3n y voto.

```
data_filteres<-data %>% select(x_age,x_sex_num,d_treatment_source)
bt_personal<-balance_table(data_filteres,treatment = "d_treatment_source")

kable(as.data.frame(bt_personal),booktabs=T,caption = "Balance de variables disponibles",
      longtable=T,digits=2)%>%
kable_styling(position = "center",latex_options = "repeat_header")
```

Table 1: Balance de variables disponibles

| variables1 | Media_control1 | Media_trat1 | p_value1 |
|------------|----------------|-------------|----------|
| x_age | 47.15 | 45.89 | 0.08 |
| x_sex_num | 0.51 | 0.47 | 0.06 |

En ambas variables no hay diferencias estadisticamente significativas entre sus medias al 95%. Sin embargo,

sí hay diferencias a un 90% debido a que el p value es muy pequeño (para lo que quisieramos). Sin embargo, las medias nos indican que si bien si puede haber una diferencia real, esta es pequeña.

VI.

Realizo las pruebas t para cada variable de respuesta.

```
#para score de credibilidad
(a<-t.test(y_belief_report_num ~ d_treatment_source, data = data))

##
## Welch Two Sample t-test
##
## data: y_belief_report_num by d_treatment_source
## t = -9.0003, df = 1934.2, p-value < 2.2e-16
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.7594179 -0.4876740
## sample estimates:
## mean in group 0 mean in group 1
## 3.153314 3.776860

#para score de mail
(b<-t.test(y_share_report_email_num ~ d_treatment_source, data = data))

##
## Welch Two Sample t-test
##
## data: y_share_report_email_num by d_treatment_source
## t = -1.8237, df = 1850, p-value = 0.06837
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.039391364 0.001431966
## sample estimates:
## mean in group 0 mean in group 1
## 0.04563895 0.06461864

#para score de facebook
(c<-t.test(y_share_report_fb_num ~ d_treatment_source, data = data))

##
## Welch Two Sample t-test
##
## data: y_share_report_fb_num by d_treatment_source
## t = -2.8176, df = 1163.1, p-value = 0.004921
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.08568860 -0.01533839
## sample estimates:
## mean in group 0 mean in group 1
## 0.08457711 0.13509061

#para score de twitter
(d<-t.test(y_share_report_twitter_num ~ d_treatment_source, data = data))

##
## Welch Two Sample t-test
```

```
##
## data: y_share_report_twitter_num by d_treatment_source
## t = -1.0997, df = 286.87, p-value = 0.2724
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.11927955 0.03376774
## sample estimates:
## mean in group 0 mean in group 1
## 0.1048951 0.1476510

#para score de whatsapp
(e<-t.test(y_share_report_whatsapp_num ~ d_treatment_source, data = data))

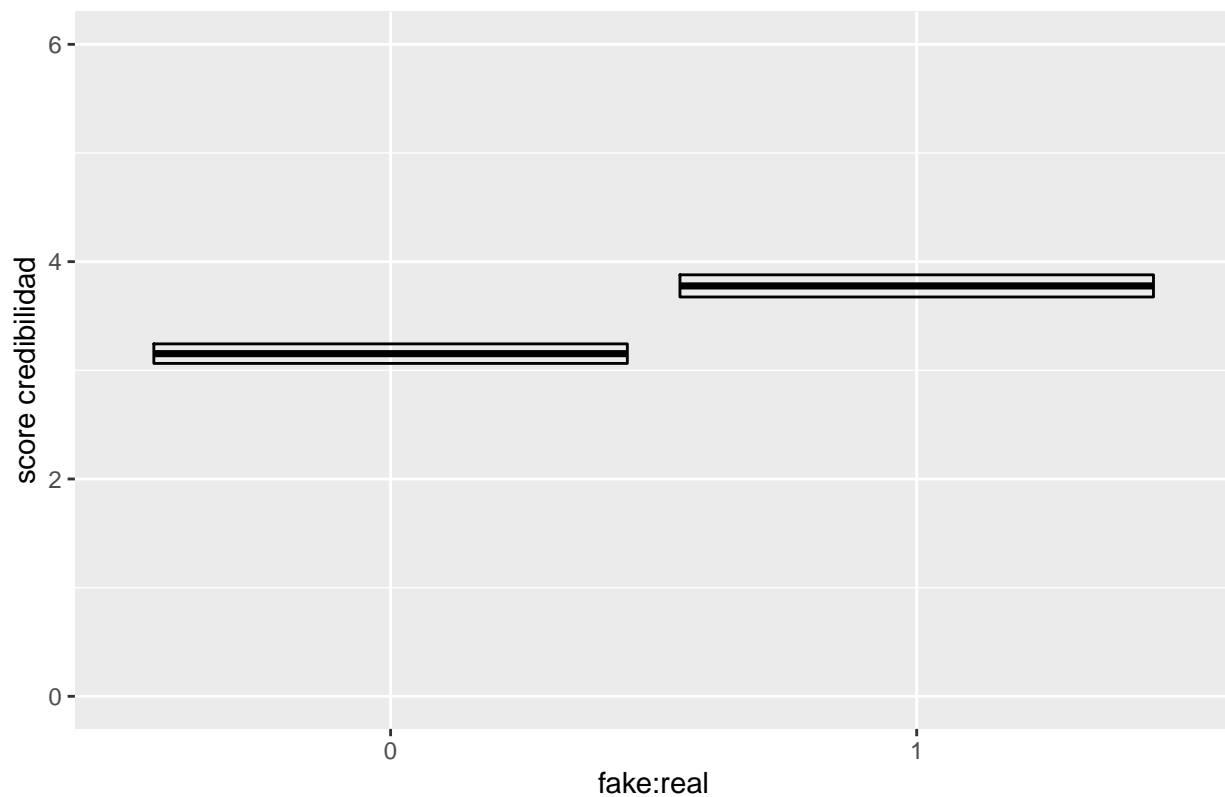
##
## Welch Two Sample t-test
##
## data: y_share_report_whatsapp_num by d_treatment_source
## t = -2.0624, df = 1574.4, p-value = 0.03933
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
## -0.056107974 -0.001407372
## sample estimates:
## mean in group 0 mean in group 1
## 0.07261905 0.10137672
```

VII.

Score credibilidad

```
ggplot(data, aes(factor(d_treatment_source), y = y_belief_report_num)) +
  stat_summary(fun.data = "mean_cl_normal", geom = "crossbar") + ylim(0,6)+
  labs(title = "Credibilidad con intervalos de confianza al 95%")+
  xlab("fake:real")+ylab("score credibilidad")
```

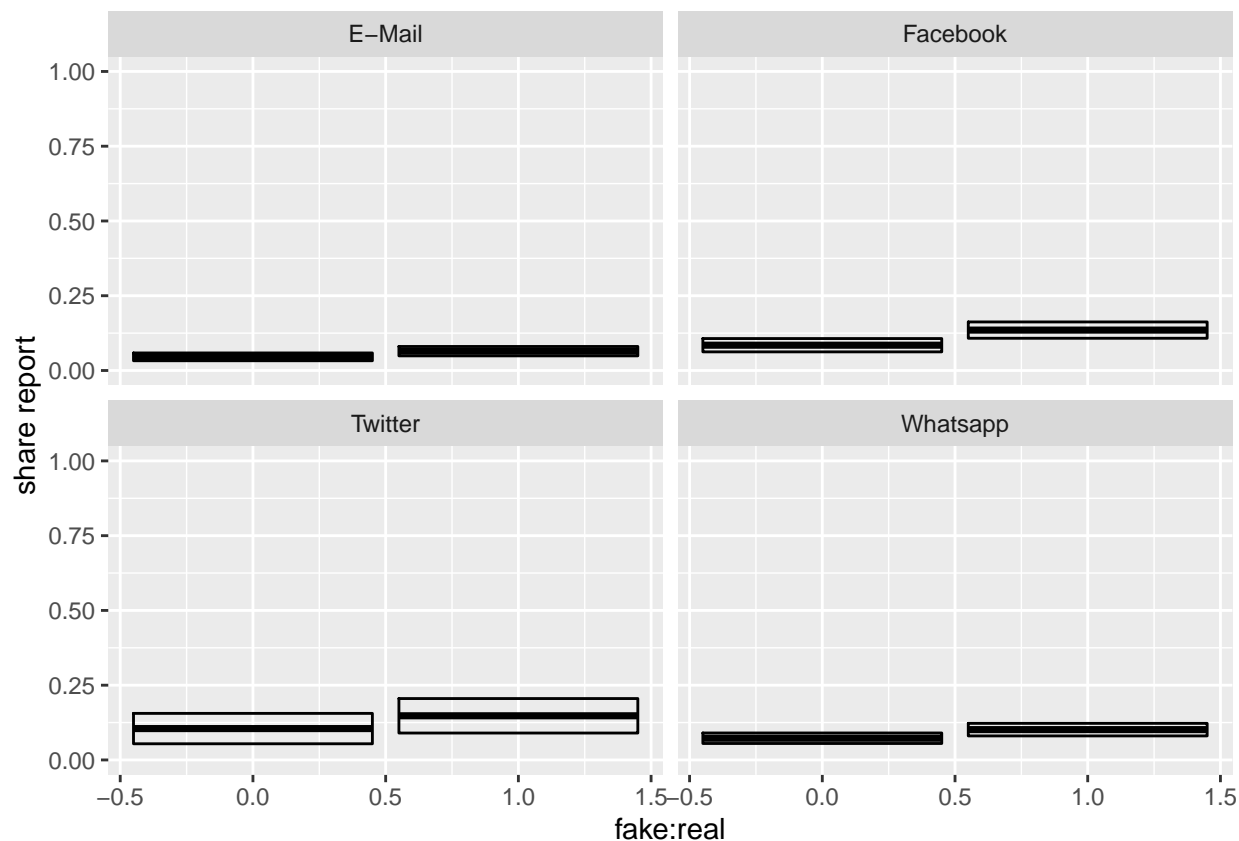
Credibilidad con intervalos de confianza al 95%



Intención de compartir

```
b<-data %>% select(y_share_report_email_num,d_treatment_source) %>% mutate(category="E-Mail")
c<-data %>% select(y_share_report_fb_num,d_treatment_source)%>% mutate(category="Facebook")
d<-data %>% select(y_share_report_twitter_num,d_treatment_source)%>% mutate(category="Twitter")
e<-data %>% select(y_share_report_whatsapp_num,d_treatment_source)%>% mutate(category="Whatsapp")
names(b)<-c("variable","treat","category")
names(c)<-c("variable","treat","category")
names(d)<-c("variable","treat","category")
names(e)<-c("variable","treat","category")
f<-rbind(b,c,d,e)

ggplot(f, aes(treat, y = variable)) +
  stat_summary(fun.data = "mean_cl_normal", geom = "crossbar") +
  xlab("fake:real")+ylab("share report")+ ylim(0,1)+ facet_wrap(~category)
```



De los gráficos podemos concluir que en general, se cumplen las dos hipótesis planteadas:

- Las personas tienden a creer más en las noticias de una fuente real que de una fuente falsa.
- Las personas serán más propensas a compartir informes de noticias que provienen de fuentes reales que aquellos que provienen de fuentes falsas³

Aunque notamos que hay efectos heterógeneos y en algunas plataformas las diferencias son mínimas y no significativas (o con significancias poco convencionales para rechazar la hipótesis nula), como el caso de twitter y de email.

VIII

Ahora se obtienen estos mismos resultados pero implementando sus cálculos mediante la estimación adecuada de mínimos cuadrados ordinarios. Para esta parte no se elaborará el análisis para la congruencia y la interacción debido a que no se tienen los datos.

```
a<-lm(y_belief_report_num~d_treatment_source,data)
b<-lm(y_share_report_email_num~d_treatment_source,data)
c<-lm(y_share_report_fb_num~d_treatment_source,data)
d<-lm(y_share_report_twitter_num~d_treatment_source,data)
e<-lm(y_share_report_whatsapp_num~d_treatment_source,data)
#stargazer(a,b,c,d,e, type="latex", digits = 2)
```

Table 2:

| | Dependent variable: | | | | |
|-------------------------|--------------------------|------------------------|------------------------|--------------------------|---------------------------|
| | Score de creencia (1) | compartir Email (2) | compartir FB (3) | compartir Twitter (4) | compartir Whatsapp (5) |
| d_treatment_source | 0.62*** (0.07) | 0.02* (0.01) | 0.05*** (0.02) | 0.04 (0.04) | 0.03** (0.01) |
| Constant | 3.15*** (0.05) | 0.05*** (0.01) | 0.08*** (0.01) | 0.10*** (0.03) | 0.07*** (0.01) |
| Observations | 1,979 | 1,930 | 1,210 | 292 | 1,639 |
| R ² | 0.04 | 0.002 | 0.01 | 0.004 | 0.003 |
| Adjusted R ² | 0.04 | 0.001 | 0.01 | 0.001 | 0.002 |
| Residual Std. Error | 1.54 (df = 1977) | 0.23 (df = 1928) | 0.31 (df = 1208) | 0.33 (df = 290) | 0.28 (df = 1637) |
| F Statistic | 81.38*** (df = 1; 1977) | 3.35* (df = 1; 1928) | 7.93*** (df = 1; 1208) | 1.20 (df = 1; 290) | 4.29** (df = 1; 1637) |

Note:

*p<0.1; **p<0.05; ***p<0.01

Ahora pasamos a realizar las pruebas de hipótesis correspondientes para cada modelo.

$$H_n : \beta_t = 0$$

$$H_a : \beta_t \neq 0$$

Simplemente observamos la columna de t value y la de significancia.

```
#credibilidad
coeftest(a)
```

```
##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    3.153314   0.048342  65.229 < 2.2e-16 ***
## d_treatment_source 0.623546   0.069121   9.021 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#compartir en mail
coeftest(b)
```

```
##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.045639   0.007253  6.2924 3.858e-10 ***
## d_treatment_source 0.018980   0.010371   1.8301  0.06739 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#compartir en facebook
coeftest(c)
```

```
##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.084577   0.012707  6.6562 4.254e-11 ***
## d_treatment_source 0.050513   0.017940   2.8157  0.004947 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#compartir en twitter
coeftest(d)

##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.104895   0.027856  3.7657 0.0002011 ***
## d_treatment_source 0.042756   0.038995  1.0964 0.2737973
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
#compartir en whatsapp
coeftest(e)

##
## t test of coefficients:
##
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.0726190  0.0096992  7.4872 1.144e-13 ***
## d_treatment_source 0.0287577  0.0138915  2.0702  0.03859 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Notamos que el efecto es significativo a más del 95% en el score de credibilidad y en la incidencia de compartir la nota en facebook y whatsapp. Asimismo el efecto para e mail es signifactivo al 90 % y para twitter no es significativo. En cuanto a la dimensión del efecto, para el score de incidencia es de .62 puntos en la escala del 0 al 6; en cuanto a las variables de compartir, es de 5 puntos porcentuales para facebook, 4 para twitter, 3 para whatsapp y 2 para e-mail.

IX

A continuación se muestran los modelos OLS con las covariables de educación y de ingreso.

```
a<-lm(y_belief_report_num~d_treatment_source+
      factor(x_education_num)+
      factor(x_income_num),data)
b<-lm(y_share_report_email_num~d_treatment_source+
      factor(x_education_num)+
      factor(x_income_num),data)
c<-lm(y_share_report_fb_num~d_treatment_source+
      factor(x_education_num)+
      factor(x_income_num),data)
d<-lm(y_share_report_twitter_num~d_treatment_source+
      factor(x_education_num)+
      factor(x_income_num),data)
e<-lm(y_share_report_whatsapp_num~d_treatment_source+
      factor(x_education_num)+
      factor(x_income_num),data)
stargazer(a,b,c,d,e, type="latex", digits = 2)
```

Table 3:

| | <i>Dependent variable:</i> | | | | |
|--------------------------|----------------------------|------------------------|------------------------|--------------------------|---------------------------|
| | Score de creencia (1) | compartir Email (2) | compartir FB (3) | compartir Twitter (4) | compartir Whatsapp (5) |
| d_treatment_source | 0.61*** (0.07) | 0.02* (0.01) | 0.05*** (0.02) | 0.04 (0.04) | 0.03** (0.01) |
| factor(x_education_num)2 | 0.16 (0.11) | -0.03* (0.02) | -0.09*** (0.03) | -0.04 (0.07) | -0.01 (0.02) |
| factor(x_education_num)3 | 0.45*** (0.15) | -0.04 (0.02) | -0.06 (0.04) | -0.06 (0.09) | 0.003 (0.03) |
| factor(x_education_num)4 | 0.74*** (0.12) | -0.06*** (0.02) | -0.09*** (0.03) | -0.10 (0.07) | 0.01 (0.03) |
| factor(x_education_num)5 | 0.68*** (0.12) | -0.03* (0.02) | -0.05 (0.03) | -0.05 (0.07) | -0.02 (0.03) |
| factor(x_income_num)2 | -0.35 (0.22) | 0.01 (0.03) | 0.05 (0.06) | 0.10 (0.11) | 0.02 (0.05) |
| factor(x_income_num)3 | -0.35* (0.21) | 0.04 (0.03) | 0.01 (0.06) | 0.13 (0.11) | 0.01 (0.04) |
| factor(x_income_num)4 | -0.34 (0.21) | 0.02 (0.03) | -0.0003 (0.05) | 0.14 (0.11) | 0.001 (0.04) |
| factor(x_income_num)5 | -0.11 (0.21) | 0.06* (0.03) | 0.05 (0.05) | 0.10 (0.10) | 0.03 (0.04) |
| factor(x_income_num)6 | -0.16 (0.21) | 0.04 (0.03) | 0.05 (0.06) | 0.13 (0.11) | 0.002 (0.04) |
| factor(x_income_num)7 | -0.15 (0.21) | 0.03 (0.03) | 0.05 (0.06) | 0.20* (0.11) | 0.02 (0.04) |
| factor(x_income_num)8 | -0.003 (0.22) | 0.04 (0.03) | 0.06 (0.06) | 0.18 (0.11) | 0.01 (0.05) |
| factor(x_income_num)9 | -0.16 (0.24) | 0.03 (0.04) | 0.05 (0.06) | 0.25* (0.13) | 0.003 (0.05) |
| factor(x_income_num)10 | -0.14 (0.24) | 0.03 (0.04) | 0.005 (0.06) | 0.11 (0.12) | 0.02 (0.05) |
| factor(x_income_num)11 | -0.13 (0.23) | 0.06* (0.04) | -0.01 (0.06) | 0.12 (0.11) | 0.04 (0.05) |
| Constant | 2.95*** (0.21) | 0.04 (0.03) | 0.12** (0.05) | 0.03 (0.10) | 0.06 (0.04) |
| Observations | 1,945 | 1,901 | 1,196 | 287 | 1,616 |
| R ² | 0.08 | 0.01 | 0.02 | 0.04 | 0.01 |
| Adjusted R ² | 0.07 | 0.003 | 0.01 | -0.02 | -0.003 |
| Residual Std. Error | 1.50 (df = 1929) | 0.23 (df = 1885) | 0.31 (df = 1180) | 0.33 (df = 271) | 0.28 (df = 1600) |
| F Statistic | 11.50*** (df = 15; 1929) | 1.41 (df = 15; 1885) | 1.77** (df = 15; 1180) | 0.67 (df = 15; 271) | 0.65 (df = 15; 1600) |

Note:

*p<0.1; **p<0.05; ***p<0.01

Sin embargo notamos que las variables de income no están aportando mucho. Por lo que probamos los modelos solo con las covariables de educación.

```
a<-lm(y_belief_report_num~d_treatment_source+factor(x_education_num),data)
b<-lm(y_share_report_email_num~d_treatment_source+factor(x_education_num),data)
c<-lm(y_share_report_fb_num~d_treatment_source+factor(x_education_num),data)
d<-lm(y_share_report_twitter_num~d_treatment_source+factor(x_education_num),data)
e<-lm(y_share_report_whatsapp_num~d_treatment_source+factor(x_education_num),data)
stargazer(a,b,c,d,e, type="latex", digits = 2)
```

Table 4:

| | Dependent variable: | | | | |
|--------------------------|--------------------------|------------------------|------------------------|--------------------------|---------------------------|
| | Score de creencia (1) | compartir Email (2) | compartir FB (3) | compartir Twitter (4) | compartir Whatsapp (5) |
| d_treatment_source | 0.62*** (0.07) | 0.02* (0.01) | 0.05*** (0.02) | 0.04 (0.04) | 0.03** (0.01) |
| factor(x_education_num)2 | 0.20* (0.11) | -0.03* (0.02) | -0.09*** (0.03) | -0.02 (0.07) | -0.01 (0.02) |
| factor(x_education_num)3 | 0.49*** (0.15) | -0.03 (0.02) | -0.06 (0.04) | -0.06 (0.08) | 0.01 (0.03) |
| factor(x_education_num)4 | 0.79*** (0.12) | -0.05*** (0.02) | -0.09*** (0.03) | -0.09 (0.07) | 0.01 (0.03) |
| factor(x_education_num)5 | 0.76*** (0.12) | -0.02 (0.02) | -0.05 (0.03) | -0.02 (0.07) | -0.01 (0.02) |
| Constant | 2.71*** (0.10) | 0.08*** (0.02) | 0.15*** (0.03) | 0.14** (0.06) | 0.08*** (0.02) |
| Observations | 1,961 | 1,914 | 1,201 | 287 | 1,627 |
| R ² | 0.08 | 0.01 | 0.02 | 0.01 | 0.004 |
| Adjusted R ² | 0.08 | 0.004 | 0.01 | -0.005 | 0.001 |
| Residual Std. Error | 1.51 (df = 1955) | 0.23 (df = 1908) | 0.31 (df = 1195) | 0.33 (df = 281) | 0.28 (df = 1621) |
| F Statistic | 33.38*** (df = 5; 1955) | 2.45** (df = 5; 1908) | 3.88*** (df = 5; 1195) | 0.72 (df = 5; 281) | 1.26 (df = 5; 1621) |

Note:

*p<0.1; **p<0.05; ***p<0.01

Lo que podemos ver es que son virtualmente los mismos resultados (en cuanto a dimensión y significancia del efecto) con y sin covariables, además de que casi ninguna variable de ingreso resulto significativa para cualquier modelo. De ahí que concluimos que las covariables no son necesarias. En todo caso podríamos agregar las de educación solo con el proposito de mejorar minimamente la eficiencia o si nos queremos enfocar en la inferencia sobre los grupos educativos. Por el contrario, la covariable de ingreso no aporta mucho.

X.

Conclusión

Del análisis anterior podemos concluir que en general, se cumplen las dos hipótesis planteadas:

- Las personas tienden a creer más en las noticias de una fuente real que de una fuente falsa.
- Las personas serán más propensas a compartir informes de noticias que provienen de fuentes reales que aquellos que provienen de fuentes falsas

Aunque notamos que hay efectos heterógeneos y en algunas plataformas las diferencias son mínimas o no significativas (o con significancias poco convencionales para rechazar la hipótesis nula), como el caso de twitter y de email.

Cuantificación del efecto causal

Notamos que el efecto de ser expuesto a una fuente verdadera es significativo a más del 95% en el score de credibilidad y en la incidencia de compartir la nota en facebook y whatsapp. Asimismo el efecto en compartir la noticia por e-mail es significativo al 90 % y para twitter no es significativo. En cuanto a la dimensión del efecto causal, para el score de incidencia es de .62 puntos en la escala del 0 al 6; en cuanto a las variables de compartir, es de 5 puntos porcentuales para facebook, 4 para twitter, 3 para whatsapp y 2 para e-mail.

Utilidad y ventajas de la estrategia

En primer lugar, el diseño experimental aleatorio es una manera muy robusta de argumentar en favor de una hipótesis. En la jerarquía de evidencia, dichos argumentos están en la cima. Además, el diseño hace que la inferencia sea fácil y orgánica en el sentido de que la manipulación de los datos es mínima o inexistente además en dichos diseños es muy fácil e identificar al contrafactual, de esta manera la estimación de los efectos causales es sencilla.

En segundo lugar, más como opinión personal, creo que la pregunta no era la correcta o bien pudo haber sido más ambiciosa, creo que la hipótesis y construcción del experimento es un poco tautológico: “si creamos una fuente falsa, la gente creerá en ella menos y compartirá menos”. Creo que hubiese sido más interesante el preguntarse o evaluar cosas como el porqué la gente cree que una fuente que es falsa o cuales son los determinantes que hacen que alguien crea en una fuente falsa, que si tocan un poco este tema en la revisión de literatura.