Michael Borsellino and Bob McMahon

Understanding Civic Awareness and Political Activism: A Confirmatory Factor Analysis

Civic awareness is a crucial component of any functional democracy. Therefore, it is important that we understand civic awareness to evaluate the health of our country's political structure. An understanding of this connection is nothing new – there has been a wealth of knowledge produced on indicators of civic awareness. As early as 1835, Alexis de Tocqueville posited a positive correlation between newspaper readership and civic participation (Tocqueville, 1835). This correlation has continued to be studied in an effort to understand democratic participation. For instance, newspapers have also been linked to political participation. The effect of reading the newspaper on political participation is compounded by participation in political discussions (Scheufele, 2002). Similarly, knowledge of current events and political issues is tied to civic engagement and awareness (Galston, 2001; Galston, 2007). Despite this extensive research, little is known about what factors contribute to denotations of civic awareness, like newspaper readership or participation in political discussion. This research seeks to fill that gap by measuring a latent factor representing civic awareness to predict selected measured variables.

Closely related to civic awareness is political activism. Whereas civic awareness is a more passive form of engagement, political activism is an active form of engagement involving behaviors like the signing of petitions. Existing research has shown that activities that are associated with political engagement tend to be correlated. For example, there is evidence that rates of signing petitions, demonstrating, boycotting consumer produces, and occupying buildings have increased concurrently since 1974 (Stolle, Hooghe, & Micheletti, 2005). However, like awareness, much is known about the indicators of engagement but little is known about the factors that influence those indicators. Further, due to the close ties between awareness

and engagement, it stands to reason that civic awareness may predict political engagement (Pasek et al., 2006). As a result, it makes sense to examine political activism measures as both predicted by a factor representing civic awareness and predicted by a factor representing political activism that may or may not be correlated with civic awareness.

As hinted at above, this research is concerned with measuring the effects of civic awareness and political activism on observable variables. In order to carry out this research, we initiate a confirmatory factor analysis (CFA). The CFA is appropriate because it will verify if the unobserved factors, civic awareness and political activism, predict their hypothesized observable indicators. Using this approach, we test two hypotheses: 1) civic awareness underlies the six selected measured variables, and 2) civic awareness underlies three measured variables while political activism underlies the three remaining measured variables. For both hypotheses, the null hypothesis is that the hypothesized structure fits the data well; confirmation of the null hypothesis will indicate a plausible model.

Methodology

General Social Survey

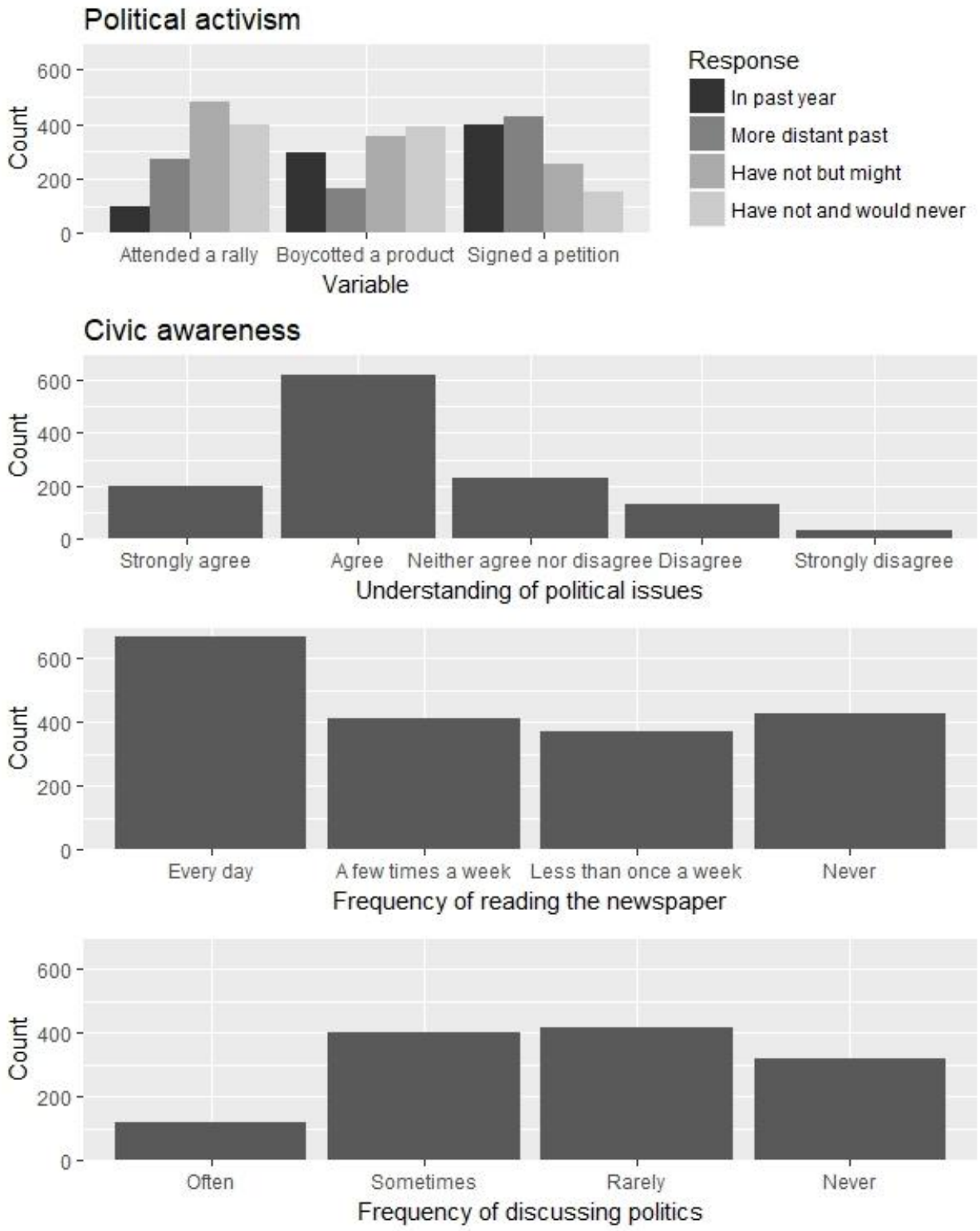
This analysis relies on the General Social Survey (GSS) dataset. The GSS is a sociological survey of United States' residents that originated in 1972. The GSS is unique in that "it is the only full-probability, personal interview survey designed to monitor changes in both social characteristics and attitudes currently being conducted in the United States." It was originally administered every year, though since 1994 it has only been administered in even years for a total of 30 samples to date. The survey has two main goals: 1) "To conduct basic scientific research on the structure and development of American society", and 2) "To distribute up-to-date, important, high-quality data to social scientists, students, policy makers, and others" (GSS – About, 2016).

The GSS sample population consists of all persons aged 18 and over living in households in the United States. The survey is only available in English and Spanish (the addition of Spanish came in 2006), and those requiring a different language are not in the sample population. Further, residents of institutions or group quarters are also not in the sample population (GSS – FAQ, 2016). To select respondents, household addresses are "randomly selected from a scientific sample designed to represent a cross-section of the country." Selected subjects are then asked to complete a ninety minute face-to-face computer assisted personal interview (CAPI) (GSS – For Respondent, 2016). The survey has a minimum target response rate of 75% (NSF, 2016), though the 2014 iteration of the survey, which was used for this analysis, fell short of this target with a rate of 69.2%. Over 30 samples, the GSS has had 59,599 total respondents (GSS – FAQ, 2016). However, due to questions varying from year-to-year, this analysis only incorporates 3,842 respondents from the 2014 data.

Variables

This analysis used two latent variables: Civic Awareness (civaware) and Political Activism (polact). Civic Awareness is assumed to cause the variation and covariation between three measured variables: understanding of political issues (poleff19), frequency of reading the newspaper (news), frequency of discussing politics (discpol). Political Activism is assumed to cause the variation and covariation between the three remaining measured variables: signing of a petition (signdpet), participation in product boycotts (avoidbuy), and attendance of political rallies (attrally). Figure 1, below, present the frequency distributions for these variables. The following paragraphs will explain how these measured variables were attained.

Figure 1. Distributions of observed variables for latent variables


Political activism
Count
600
400
200
0
Attended a rally
Boycotted a product
Signed a petition
Variable
Response
In past year
More distant past
Have not but might
Have not and would never
Civic awareness
Count
600
400
200
0
Strongly agree
Agree
Neither agree nor disagree
Disagree
Strongly disagree
Understanding of political issues
Count
600
400
200
0
Every day
A few times a week
Less than once a week
Never
Frequency of reading the newspaper
Count
600
400
200
0
Often
Sometimes
Rarely
Never
Frequency of discussing politics

Understanding of political issues was determined via the prompt: "To what extent do you agree or disagree with the following statement? I feel I have a pretty good understanding of the important political issues facing America." Responses took the form of a five-point scale from 1-Strongly agree to 5-Strongly disagree; Can't choose, No answer, and Not applicable were other possibilities the interviewer could indicate. There were 2,675 valid responses. Frequency of reading the newspaper was determined by asking "How often do you read the newspaper—every day, a few times a week, once a week, less than once a week, or never?" Responses matched the categories provided in the question. There were 39,249 valid responses. Frequency of discussing politics was attained by asking "When you get together with your friends, relatives or fellow workers, how often do you discuss politics?" Possible responses were Often, Sometimes, Rarely, Never, Can't choose, No answer, and Not applicable. There were 2,722 valid responses.

The following prompt was used for signing a petition, participation in product boycotts, and attendance of political rallies: "Please indicate whether you have [category] in the past year, whether you have done it in the more distant past, whether you have not done it but might do it, or have not done it and would never, under any circumstances, do it." For each case, [category] was replaced by "signed a petition", "boycotted, or deliberately bought, certain products for political, ethical or environmental reasons", or "attended a political meeting or rally", respectively. Responses matched the categories provided in the question. There were 2,697 responses for petition, 2,666 responses for boycott, and 2,710 responses for rally.

Confirmatory Factor Analysis

A confirmatory factor analysis (CFA) was used to test whether a relationship between the measured variables and their respective factors, Civic Awareness and Political Activism, exists. CFA is a type of structural equation model (SEM) that is used for analyzing structural relationships. For this analysis, a CFA is appropriate because it is a measurement model; it tests

whether the hypothesized model fits the data. In Figures 2 and 3, below, the one-factor and two-factor models can be seen. The one-factor model has 6 factor loadings, 6 intercepts, and 6 unique variances. The two-factor model also has 6 factor loadings, 6 intercepts, and 6 unique variances. It also has one factor covariance, signified by the two-sided error between the factors. The dotted lines in the figures below signify factor loadings fixed at 1; this is expanded upon in the next section. Finally, it is important to note that, because of the non-normal distributions of observed variables, this analysis used bootstrapping to estimate the standard errors (circles).

Figure 2. One-Factor Model

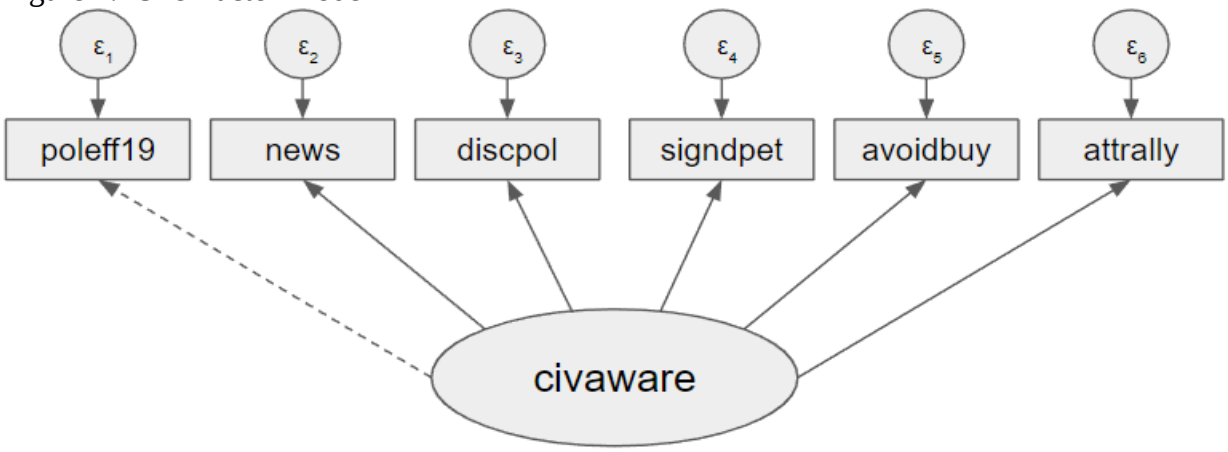
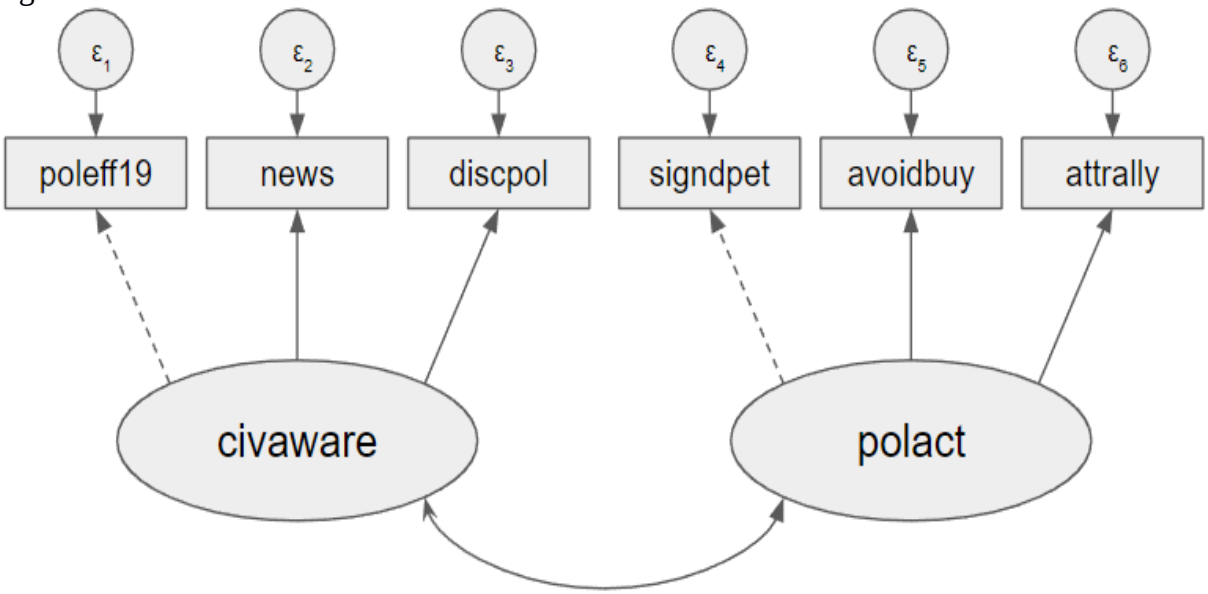


Figure 3. Two-Factor Model



The performance of the model in accounting for the correlations between measured variables was assessed via several tests and indicators of model fit. These included an overall chi-square test of model fit, confirmatory fit index (CFI), standardized root mean square residual (SRMR), and root mean square error of approximation (RMSEA). This analysis was completed with the lavaan package (version 0.5-23.1097) in R (version 3.4.1). The syntax is available in the Appendix. The following section details the results of the CFA.

Results

Table 1, below, shows the output for the one-factor model. In this model, understanding of political issues (poleff19) was constrained to a value of 1 in order to define the scale of the factor variable, Civic Awareness. These coefficients are found labeled Unstandardized Estimates in the table. Notably, all path coefficients are positive, indicating a positive relationship between the factor variable and the measured variables. Further, each coefficient has a p-value of 0.000, indicating that they are statistically significant and different from zero. The residual variances in Table 1 are estimates of error variance for the measured variables. For example, the estimated error variance for frequency of discussing politics (discpol) is 0.527. The test statistics in Table 3 are okay (preferred threshold in parentheses): CFI is 0.894 (>0.90), RMSEA is 0.071 (<0.05), and SRMR is 0.054 (<0.09). However, given the p-value of 0.000, the null hypothesis is rejected. The one-factor model does not adequately fit the data.

Table 1. One-Factor CFA Factor Loadings and Residual Variance

	Unstandardized Estimates	Std. Err	Z-value	P(>|z|)	Residual Variance	Std. Err	Z-value	P(>|z|)
Civic Awareness								
poleff19	1.000				0.741	0.049	15.014	0.000
news	1.306	0.230	5.677	0.000	2.171	0.084	25.904	0.000
discpol	1.632	0.220	7.422	0.000	0.527	0.035	15.267	0.000
signdpet	1.816	0.268	6.784	0.000	0.558	0.036	15.634	0.000
avoidbuy	2.235	0.324	6.901	0.000	0.738	0.059	12.580	0.000
attrally	1.801	0.234	7.709	0.000	0.480	0.034	13.988	0.000

In an effort to improve on the one-factor model, a two-factor model was used. The two factors were Civic Awareness and Political Activism. Table 2, below, shows the output for this two-factor model. In this model, understanding of political issues (poleff19) was constrained to a value of 1 in order to define the scale of Civic Awareness and signing of a petition (signdpet) was constrained to a value of 1 in order to define the scale of Political Activism. These coefficients are found labeled Unstandardized Estimates in the table. Notably, all path coefficients are positive, indicating a positive relationship between the factor variable and the measured variables. Further, each coefficient has a p-value of 0.000, indicating that they are all statistically significant and different from zero. The residual variances in Table 2 are estimates of error variance for the measured variables. For example, the estimated error variance for rally attendance (attrally) is 0.459. The test statistics in Table 3 are promising: CFI is 0.931 (>0.90), RMSEA is 0.061 (<0.05), and SRMR is 0.041 (<0.09). However, given the p-value of 0.000, the null hypothesis is rejected. This two-factor model does not fit the data adequately.

Table 2. Two-Factor CFA Factor Loadings and Residual Variance

	Unstandardized Estimates	Std. Err	Z-value	P(>|z|)	Residual Variance	Std. Err	Z-value	P(>|z|)
Civic Awareness								
poleff19	1.000				0.740	0.047	15.655	0.000
news	1.275	0.208	6.144	0.000	2.135	0.086	24.705	0.000
discpol	1.693	0.238	7.117	0.000	0.417	0.049	8.536	0.000
Political Activism								
signdpet	1.000				0.557	0.037	15.032	0.000
avoidbuy	1.237	0.084	14.660	0.000	0.717	0.057	12.540	0.000
attrally	1.000	0.086	11.646	0.000	0.459	0.037	12.425	0.000

Table 3. Table of Test Statistics

	One-Factor	Two-Factor
Chi-Square P-value	0.000	0.000
CFI	0.894	0.931
RMSEA	0.071	0.061
SRMR	0.054	0.041

Discussion

Overall, the results of this confirmatory factor analysis indicate that the implied correlations of both the one-factor and two-factor models do not match the observed correlations. That is to say, these models are not adequate. While this certainly does not discredit the existing literature on civic awareness and political activism, it does show that there is much to be learned about these social phenomena. Future studies should continue to explore the factors that influence measured variables related to awareness and activism.

There are limitations to this study that must be addressed. First and foremost, there are many more measured variables associated with civic awareness and political activism than the six that this study used. These six were chosen because of their frequency in the literature, however further studies should use a more comprehensive set of variables. A second limitation is the failure to use sample weights. Weights make adjustments to the respondent data that account for how likely each respondent was to be part of the sample. The GSS documentation suggests the use of WTCOMB or WTCOMBNR weights due to the multi-stage sampling design used in 2014 (GSS – FAQ, 2016).

References

Galston, W. (2001). Political knowledge, political engagement, and civic education. Annual *Reviews in Political Science, 4*, 217-234.

Galston, W. (2007). Civic knowledge, civic education, and civic engagement: A summary of recent research. International *Journal of Public Administration*, *30*(6), 623-642.

GSS (2016). About the GSS | NORC.

GSS (2016). For Respondent GSS Survey Participant Information | NORC.

GSS (2016). GSS General Social Survey FAQ | NORC.

Inglehart, R., & Norris, P. (2003). *Rising Tide: Gender Equity and Cultural Change around the World*. Cambridge: Cambridge University Press.

NSF (2008). General Social Survey (GSS) Competition nsf08506.

Pasek, J., Kenski, K., Romer, D., & Jamieson, K. (2006). America's Youth and Community Engagement: How Use of Mass Media is Related to Civic Activity and Political Awareness in 14- to 22-Year-Olds. *Communication Research, 33*(3), 115-135.

Scheufele, D. (2002). Examining differential gains from mass media and their implications for participatory behavior. *Communications Research*, *29*(1), 46-65.

Stolle, D., Hooghe, M., & Micheletti, M. (2005). Politics in the Supermarket: Political Consumerism as a Form of Political Participation. *International Political Science Review, 26*(3), 245-269.

Tocqueville, A. (1835). *Democracy in America*. New York: Vintage Books.

Appendix

```
####### DATA MANAGEMENT #######

#setwd( "C:/R for Class/746/Paper 2" )

####load data

#library(haven)

#GSS14 <- read_sav("C:/R for Class/746/Paper 2/GSS2014merged_R7.sav")

#View(GSS14)

#

#save( GSS14 , file = "GSS14.rda" )

#

#load("GSS14.rda")

#######


#### LIBRARIES ####

library(lavaan)

library(semPlot)

library(Hmisc)

library(reshape2)

library(ggplot2)

library(gridExtra)

library(xlsx)


####### ONE LATENT VARIABLE MODEL #######
```

```
# Set up model string for fit

CFAmodel1 <- '

civaware =~ poleff19 + news + discpol + signdpet + avoidbuy + attrally

'

# Fit the model

fit1 <- cfa(CFAmodel1, data=GSS14, estimator = "WLS", se = "bootstrap")


# Get summary and plot SEM diagram

summary(fit1, fit.measures=TRUE)

semPaths(fit1)


####### TWO LATENT VARIABLE MODEL #######

# Set up model string for fit

CFAmodel2 <- '

civaware =~ poleff19 + news + discpol

polact =~ signdpet + avoidbuy + attrally


#direct effect

polact ~~ civaware

'

# Fit the model

fit2 <- cfa(CFAmodel2, data=GSS14, estimator = "WLS", se = "bootstrap")
```

```
# Get summary and plot SEM diagram

summary(fit2, fit.measures=TRUE)

semPaths(fit2)


####### MODEL COMPARISON #######


anova(fit1,fit2)


####### OUTPUT TABLES #######


# Export summaries

sink("1_Var.txt")

summary(fit1, fit.measures=TRUE)

sink()


sink("2_Var.txt")

summary(fit2, fit.measures=TRUE)

sink()


# Excel export fit measures and loadings

write.xlsx(fitMeasures(fit1), "paper2export.xlsx", sheetName="1 Var Measures",

append=FALSE)

write.xlsx(parameterEstimates(fit2), "paper2export.xlsx", sheetName="1 Var Loadings",
```

```
append=TRUE)
write.xlsx(fitMeasures(fit2), "paper2export.xlsx", sheetName="2 Var Measures",
append=TRUE)
write.xlsx(parameterEstimates(fit2), "paper2export.xlsx", sheetName="2 Var Loadings",
append=TRUE)

####### PLOTS ########

# Create new dataset for setting up data for plots

desc <- GSS14[,c("poleff19","news","discpol","signdpet","avoidbuy","attrally")]

# Format variables as factors with correct labels for plots

desc$poleff19f<-factor(desc$poleff19,levels=c(1,2,3,4,5),
labels=c("Strongly agree","Agree","Neither agree nor
disagree","Disagree","Strongly disagree"))
desc$newsf<-factor(desc$news,levels=c(1,2,3,4),
labels=c("Every day","A few times a week","Less than once a week","Never"))
desc$discpolf<-factor(desc$discpol,levels=c(1,2,3,4),
labels=c("Often","Sometimes","Rarely","Never"))

desc$signdpetf<-factor(desc$signdpet,levels=c(1,2,3,4),
```

```
labels=c("In past year","More distant past","Have not but might","Have not and would never"))

desc$avoidbuyf<-factor(desc$avoidbuy,levels=c(1,2,3,4),
labels=c("In past year","More distant past","Have not but might","Have not and would never"))

desc$attrallyf<-factor(desc$attrally,levels=c(1,2,3,4),
labels=c("In past year","More distant past","Have not but might","Have not and would never"))


### PLOT POLITCAL AWARENESS OBSERVED VARS ###


freq <- table(col(desc[,4:6]), as.matrix(desc[,4:6]))


#Then you need to create a data frame out of it, melt it and plot it:


Variable=c("Signed a petition","Boycotted a product","Attended a rally")     # create list of names

desc.2=data.frame(Variable, cbind(freq))   # combine them into a data frame


# Melting for plotting and setting up labels

desc.2m <- melt(desc.2, id="Variable")


levels(desc.2m$variable)[1] <- "In past year"
```

```
levels(desc.2m$variable)[2] <- "More distant past"
levels(desc.2m$variable)[3] <- "Have not but might"
levels(desc.2m$variable)[4] <- "Have not and would never"

colnames(desc.2m) <- c("Variable","Response","Count")

# Save plot for arranged output
pl1 <- ggplot(desc.2m, aes(Variable, Count)) +
geom_bar(aes(fill = Response), position = "dodge", stat="identity") + scale_fill_grey() +
scale_y_continuous( limits = c(0,700), expand = c(0,0) ) +
labs(title="Political activism",x="Variable",y="Count")

### PLOT CIVIC AWARENESS OBSERVED VARS ###

# Save plots for arranged output
pl2 <- ggplot(data=subset(desc, !is.na(poleff19f)), aes(poleff19f)) + geom_bar() +
scale_y_continuous( limits = c(0,700), expand = c(0,0) ) +
labs(title="Civic awareness",x="Understanding of political issues",y="Count")
pl3 <- ggplot(data=subset(desc, !is.na(newsf)), aes(newsf)) + geom_bar() +
scale_y_continuous( limits = c(0,700), expand = c(0,0) ) +
labs(title=NULL,x="Frequency of reading the newspaper",y="Count")
pl4 <- ggplot(data=subset(desc, !is.na(discpolf)), aes(discpolf)) + geom_bar() +
scale_y_continuous( limits = c(0,700), expand = c(0,0) ) +
```

```
labs(title=NULL,x="Frequency of discussing politics",y="Count")


### GENERATE PLOTS ###


grid.arrange(pl1, pl2, pl3, pl4, nrow = 4)
```