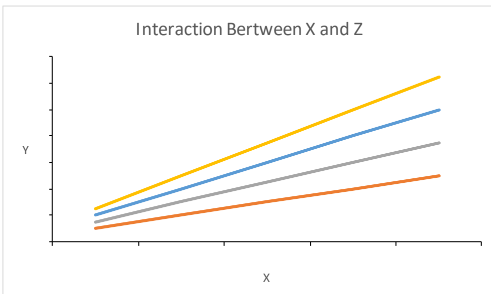
Interactions with Logistic Regression

An interaction occurs if the relation between one predictor, X, and the outcome (response) variable, Y, depends on the value of another independent variable, Z (Fisher, 1926). Z is said to be the moderator of the effect of X on Y, but a $X \times Z$ interaction also means that the effect of Z on Y is moderated by X. There are a number of synonyms for statistical interactions, including moderator, modifier, qualifier, magnifier, antagonistic, effect modifier, or buffering effect. An interaction represents a synergistic or multiplicative effect tested by adding a product variable, XZ to the model, implying a non-additive effect that is over and above the effect of the linear effects of X and Y entered together in the model. The regression coefficient for the product term represents the degree to which there is an interaction between the two variables. The effect of X on Y is not the same for all values of Z, which, in linear regression, is graphically represented by non-parallel slopes.



If slopes are parallel, the effect of X on Y is the same at all levels of Z, and there is no interaction. Variable X and Z may be binary or continuous. For the special case in which X and Z are both binary, the regression model with continuous response is equal to an analysis of variance (ANOVA).

Interactions are similarly specified in logistic regression if the response is binary. The right hand side of the equation includes coefficients for the predictors, X, Z, and XZ.

$$\ln\left(\frac{\pi}{1-\pi}\right) = \alpha + \beta_1 X + \beta_2 Z + \beta_3 XZ$$

If the interaction coefficient β_3 is significant, we conclude that the association between X and the probability that $Y = 1$ depends on the values of Z. X and Z may be binary or continuous. For the special case in which X and Z are both binary, then the analysis corresponds with the 2 × 2 × 2 contingency table analysis.

Statistical Tests

The test of the interaction may be conducted with the Wald chi-squared test or a likelihood ratio test comparing models with and without the interaction term. In this particular case, the Wald test appears to perform better than the likelihood ratio test (Allison, 2014). Note that it is always important to have the lower order (main effects) in the model with the interaction variable. Without them, the interaction is potentially confounded with the overall (and additive) effects of the two variables.

Multiplicative Effects on Proportions and the Linear Link Function

Most researchers testing interactions with logistic regression use the above describe method for determine that there is a multiplicative increase in the odds, which involves an effect of X on the logit that depends on the value of Z. The logit transformation of the predicted probabilities, however, is by nature a nonlinear transformation, so an interaction of X and Z on the logit will not necessarily mean that there is an interaction effect of these variables when considering the predicted probabilities. The reverse is also true, that when there is a multiplicative increase in the predicted probabilities, it is possible not to have a multiplicative increase in the odds (Hosmer et al., 2013). This counterintuitive phenomenon is illustrated in a figure from Hosmer and colleagues in the overhead “Logistic Interaction Figures” on the class site, who refer to the odds interaction as a "multiplicative interaction" and the interaction in the probabilities as an "additive interaction." A method that can be used to test for the additive interaction is to use a binomial linear link regression which uses a linear link to the probabilities akin to that for OLS regression but special computations of the standard errors based on the binomial distribution (more on link functions in the section of this course on generalized linear models). The binomial linear link regression may have estimation complications (particularly with continuous predictors) and can lead to negative predicted probabilities or probabilities greater than 1.0. When the focus is on the enhanced risk of an outcome and thus the predicted probabilities, the binomial linear link approach may be the desirable approach.

Centering
Rescaling the predictors is often recommended (Aiken & West, 1991) to improve the interpretation of the lower order effects, β_1 and β_2, sometimes referred to as *main effects*. *Centering* or creating deviation scores, which involves subtracting the mean from each predictor's original value, $x = X - \bar{X}$ and $z = Z - \bar{Z}$, before computing the interaction term, $xz = \left(X - \bar{X}\right)\left(Z - \bar{Z}\right)$, reduces multicollinearity among the predictors when these new variables are used in the model. The interaction coefficient and its significance are not affected by the rescaling but the standard errors of the main effects are improved because the non-essential multicollinearity is reduced. Coding of binary predictors may either be 0 and 1 (dummy coding) or they may be centered. Although centering binary variables may seem odd, the interpretation of the intercept when X and Z are at their mean may be more desirable than interpretation of their values at 0 (i.e., the logit for Y for the 0 group). With a set of g -1 dummy variables for three or more groups, centering in the usual manner is not likely to make sense, because the set of dummy variables must be orthogonal and considered together. Should a centering interpretation of the intercept and lower order effects be desirable in this case, you can consider using effect coding (see Hardy, 1993, for effect coding systems).

Centering of predictors with interaction tests will not impact the interaction coefficient or its significance tests, but the lower order terms (i.e., main effects for X and Z) will be affected. In addition to remedying the inflation of the standard errors for these lower-order terms, centering X an Z improves their interpretation. Each lower order effect represents a conditional effect on the outcome for particular values of the other variable. For example, if testing an interaction between life stress (X) social support (Z) and life stress in predicting clinical depression, the odds ratio for life stress is the increment in odds for the average level of social support. Similarly, the odds for social support is the odds for the average level of life stress.

Probing Significant Interactions
A significant interaction indicates that the effect of X is not the same for all values of Z, but neither the value nor the sign of the coefficients gives us clear information about the nature of the interaction. It is, therefore, a good idea to follow a significant interaction with some further probing of the nature of the interaction. Most commonly, the effect of X on Y for two or three specific values of Z are computed, known as *simple effect coefficients* or simple slopes. A little algebra can show the derivation of simple effect coefficients, $\beta_1|Z_i$, for Y regressed on X at some particular values of the moderator, Z_i, from the interaction equation above (Jaccard, 2001).

$$\beta_1 \mid Z_i = \beta_1 + \beta_3 Z_i$$

At the least, this involves a graph of lines representing several simple effects, but, preferably, graphing is accompanied by some follow-up significance tests. Standard errors for these simple effect coefficients can be estimated and a Wald test of their significance from zero. Simple slopes of the effect of X provides information about whether X would be significantly related to Y if Z was equal to some chosen value. Choice of values for Z depends on the researcher's goals and/or the data. If Z is binary, it usually makes sense to plot and test the effect of X for the two groups of Z, when $Z = 0$ and when $Z = 1$. If Z is continuous, values for Z can be theoretically based or arbitrary values (often one standard deviation above and below the mean).

Simple slopes can be tested using a macro (e.g., Andrew Hayes' macros for SPSS and SAS (https://afhayes.com/spss-sas-and-r-macros-and-code.html and http://processmacro.org/index.html) or a computer method. In the computer method, the logistic model with the interaction is tested multiple times using different scalings for the Z variable (Jaccard, 2001). This method capitalizes on the fact that when Z is centered, the main effect for the X variable, β_1, from the interaction model is a simple effect coefficient. It represents the effect of X when Z is equal to its mean, because the rescaled value z has a value of 0 when Z is equal to its mean. Rescaling Z again, where the standard deviation (s_z) is ***subtracted***, $z_{\text{high}} = z - s_z$, gives the simple effect coefficient and its significance test for X at one standard deviation ***above*** the mean of Z. The x variable scaling is unchanged, but the interaction must be computed anew, so that $xz_{\text{high}} = z_{\text{high}}*x$. The low-$Z$ slope simple slope can be tested by rescaling Z again, this time ***adding*** one standard deviation from the mean of Z, where $z_{\text{low}} = z + s_z$, and then recalculating the interaction term.

Plots for interactions in logistic regression can use the logit, $\ln(\pi/1\text{-}\pi)$ or the estimated probabilities, $P(Y=1)$, on the y-axis using the logistic transformation. After the simple effect coefficients have been found, these values can be plotted with a statistical package or a spreadsheet (e.g., Jeremy Dawson's excel sheet, http://www.jeremydawson.com/slopes.htm).

Interactions with Multiple Category Predictors

Interactions tests when one or more of the predictors involved in the interaction has three or more categories becomes complicated quickly. For nominal category variables (e.g., religion or type of company), the predictor variable must be split into a set of dummy variables (or effect coded variables). With three groups, two dummy variables are needed, using one of the groups as a referent group. The referent group has a code of 0 for both of the dummy variables. For example, if comparing Christian, Jewish, and Muslim, two dummy variables might be constructed with Christian as the referent group. In that case, there are two comparison variables, with Jewish vs. Christian and Muslim vs. Christian. To test an interaction between the religion of a job applicant and, say, years experience in the field in predicting whether or not the applicant is hired, two product variables must be constructed and entered into the equation together. Each interaction coefficient represents the multiplicative effect on the logit of the continuous variable with the particular dummy comparison, such as the Jewish-Christian comparison with years experience. For a significant interaction in this case, the interaction indicates that the odds of being hired for each increment in years experience is not the same in these two religious groups. In fact, if the odds of hiring for an additional year's experience is calculated separately for the three groups, the odds ratio for the interaction is simply the ratio of the two odds ratios (e.g., $OR_{\text{Jewish}} = 1.4$, $OR_{\text{Muslim}} = .9$, $OR_{\text{Jewish-ChristianxMuslim}} = 1.4/.9 = 1.56$; Jacccard, 2001).

The interpretation of the lower order effects is a function of the choice of referent. For the above example with Christian as the referent group, the odds ratio for the years experience main effect is the odds of being hired for a year increase in experience for the Christian group. The main effect for years experience is really a simple (or conditional) effect. One can easily obtain the simple effects of years experience for the other two groups by recoding the dummy variables, first making Jewish as the referent group and then making Muslim as the referent group. Other types of follow-up analyses are also possible, such as considering group differences for certain levels of the continuous variable (see Jaccard, 2001 for a more complete discussion), sometimes referred to as the Johnson-Neyman test (Hayes & Mathes, 2009).

Software Examples

I first use the Quinnipiac data[1] to reanalyze the three-way contingency table using logistic regression, where the three binary variables are response (candidate choice), independent party identification, and state (OH =0, GA = 1).

Simple Binary × Binary Interaction (same as 2 × 2 × 2 contingency analysis) Example
SPSS

```
compute inter=ind*state.
logistic regression vars=response with ind state inter.
```

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	ind	.106	.136	.601	1	.438	1.111
	state	-.117	.107	1.195	1	.274	.889
	inter	-.035	.205	.029	1	.864	.966
	Constant	.264	.076	12.128	1	.000	1.302

a. Variable(s) entered on step 1: ind, state, inter.

The test of the interaction is the Wald chi-squared for the variable INTER (which is the XZ coefficient). The chi-squared values is .029 with p = .864. Looking back at the handout "Three-way Contingency Tables" you should see more than a resemblance to the results of the Breslow-Day test.

If you wanted to center the variables (which I did not do above), you could use the following syntax.

```
AGGREGATE  /meanstate=MEAN(state).
AGGREGATE  /meanind=MEAN(ind).
compute cstate=state - meanstate.
compute cind=ind - meanind.
compute inter=cstate*cind.
logistic regression vars=response with cstate cind inter.
```

R

```
> d$inter <- d$ind * d$state
> logmod <- glm(response ~ state + ind + inter, data = d, family = "binomial")
> summary(logmod)
Call:
glm(formula = response ~ state + ind + inter, family = "binomial",
    data = d)

Deviance Residuals:
   Min      1Q  Median      3Q     Max
-1.338  -1.271   1.025   1.087   1.116

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.26386    0.07577   3.483 0.000497
state       -0.11726    0.10729  -1.093 0.274416
ind          0.10561    0.13621   0.775 0.438119
inter       -0.03491    0.20456  -0.171 0.864481

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 2690.6  on 1959  degrees of freedom
Residual deviance: 2687.6  on 1956  degrees of freedom
AIC: 2695.6
Number of Fisher Scoring iterations: 4

#obtain odds ratios
> exp(cbind(OR=coef(logmod), confint(logmod))

                   OR     2.5 %   97.5 %
(Intercept) 1.3019481 1.1227452 1.511213
state       0.8893556 0.7205862 1.097422
ind         1.1113882 0.8516247 1.452967
inter       0.9656903 0.6466785 1.442303
```

[1]Data source: https://poll.qu.edu/georgia/release-detail?ReleaseID=3679. Note that the data extrapolated cell sample sizes and used some rounding, so the results should be taken as only approximate.

For centering in R

```
> mydata$cstate <- mydata$state - mean(mydata$state)
> mydata$cind <- mydata$ind - mean(mydata$ind)
> mydata$inter <- mydata$cind * mydata$cstate
```

SAS

```
proc logistic data=one order=data descending;
model response=state ind state*ind;
run;
```

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	0.2639	0.0758	12.1283	0.0005
state	1	-0.1173	0.1073	1.1945	0.2744
ind	1	0.1056	0.1362	0.6012	0.4381
state*ind	1	-0.0349	0.2046	0.0291	0.8645

For centering

```
proc means data=one;
var state ind;
output out = meandata mean = mstate mind;
data two;
if _N_ = 1 then set meandata;
set one;
cstate = state - mstate;
cind = ind - mind;
inter = cstate*cind;
run;
```

Continuous × Continuous Interaction

In this illustration, I explore whether there is an interaction between negative social exchanges and depression in predicting heart disease among older adults. Does depression have a greater effect on heart disease when there are more negative social exchanges?

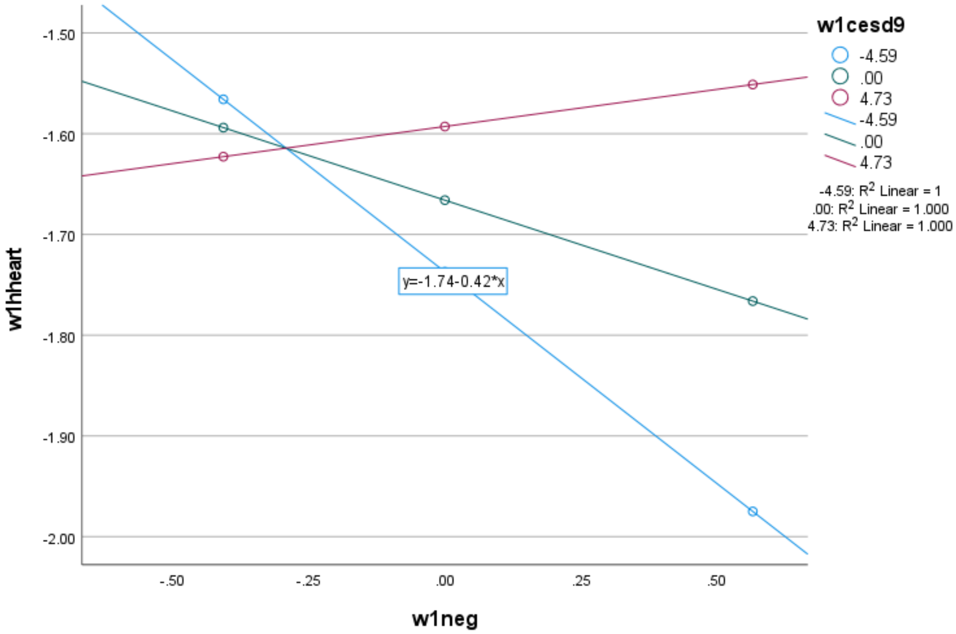
These interactions can be tested following the same procedure as above, but the process macro will compute the interaction and the simple (conditional) effects too. Download the Hayes' process macro, http://processmacro.org/index.html, and save it on your harddrive in a location you can specify in the `insert file=` command. The PROCESS macro version 3.1 and above allows for logistic regression when the dependent variable is binary. No special instructions are needed, the macro recognizes binary variables when specified on `y=`. The macro centers the predictors for you when you use `/center=1`.

```
*2-way interaction with logistic using Hayes' process macro.
* download the macro from http://processmacro.org/index.html and store it in a known location.
*replace my path with yours.

AGGREGATE  /meanneg=MEAN(w1neg).
AGGREGATE  /meandep=MEAN(w1cesd9).
compute cw1neg=w1neg - meanneg.
compute cw1cesd9=w1cesd9 -meandep.

insert file='C:\Jason\SPSSWIN\macros\process.sps'.
execute.

process y=w1hheart
/x=w1neg
/w=w1cesd9
/model=1
/center=1
/plot=1
/moments=1.
```

```
***************** PROCESS Procedure for SPSS Version 3.4 *****************

          Written by Andrew F. Hayes, Ph.D.       www.afhayes.com
    Documentation available in Hayes (2018). www.guilford.com/p/hayes3

**************************************************************************
Model  : 1
    Y  : w1hheart
    X  : w1neg
    W  : w1cesd9

Sample
Size:  692

**************************************************************************
OUTCOME VARIABLE:
 w1hheart

Coding of binary Y for logistic regression analysis:
  w1hheart  Analysis
       .00       .00
      1.00      1.00

Model Summary
       -2LL    ModelLL         df          p   McFadden   CoxSnell   Nagelkrk
   619.1451     6.5729     3.0000      .0868      .0105      .0095      .0159

Model
              coeff         se          Z          p       LLCI       ULCI
constant    -1.6659      .1074   -15.5184      .0000    -1.8763    -1.4555
w1neg        -.1771      .2238     -.7914      .4287     -.6158      .2615
w1cesd9       .0155      .0233      .6657      .5056     -.0301      .0611
Int_1         .0531      .0256     2.0715      .0383      .0029      .1032
```

```
These results are expressed in a log-odds metric.

Product terms key:
 Int_1    :        w1neg    x        w1cesd9

Likelihood ratio test(s) of highest order
unconditional interactions(s):
        Chi-sq         df          p
X*W     4.5505     1.0000      .0329
----------
    Focal predict: w1neg    (X)
          Mod var: w1cesd9  (W)

Conditional effects of the focal predictor at values of the moderator(s):

    w1cesd9     Effect         se          Z          p       LLCI       ULCI
    -4.5898     -.4206      .2969    -1.4167      .1566    -1.0026      .1613
      .0000     -.1771      .2238     -.7914      .4287     -.6158      .2615
     4.7259      .0736      .1995      .3689      .7122     -.3173      .4645

Data for visualizing the conditional effect of the focal predictor:
Paste text below into a SPSS syntax window and execute to produce plot.

DATA LIST FREE/
   w1neg      w1cesd9    w1hheart   prob       .
BEGIN DATA.
     -.4070    -4.5898    -1.5658      .1728
      .0000    -4.5898    -1.7370      .1497
      .5657    -4.5898    -1.9749      .1219
     -.4070      .0000    -1.5939      .1688
      .0000      .0000    -1.6659      .1590
      .5657      .0000    -1.7661      .1460
     -.4070     4.7259    -1.6227      .1648
      .0000     4.7259    -1.5928      .1690
      .5657     4.7259    -1.5511      .1749
END DATA.
GRAPH/SCATTERPLOT=
 w1neg    WITH     w1hheart BY       w1cesd9  .
GRAPH/SCATTERPLOT=
 w1neg    WITH     prob     BY       w1cesd9  .

*********************** ANALYSIS NOTES AND ERRORS ************************

Level of confidence for all confidence intervals in output:
  95.0000

W values in conditional tables are the minimum, the mean, and 1 SD above the mean.

NOTE: One SD below the mean is below the minimum observed in the data for W,
      so the minimum measurement on W is used for conditioning instead.

NOTE: The following variables were mean centered prior to analysis:
          w1cesd9  w1neg
```

The commands generated from the plot request start with DATA LIST FREE/ and can be pasted into a syntax file and run. The graphs do not include fit lines for the groups, so click on the chart in the output, go to Elements, and choose Fit Line at Subgroups. The y-axis is in logit metric.



I used the simple effect coefficients ("conditional effects") to calculate the simple odds ratios for the effect of depression at low (OR = $e^{-.4206}$ = .657, mean (OR = $e^{-.1771}$ = .838), and high values of negative social exchanges (OR = $e^{.0736}$ = 1.076).

Note: R and SAS code were updated on 4/26/21

R

Using the PROCESS macro in R

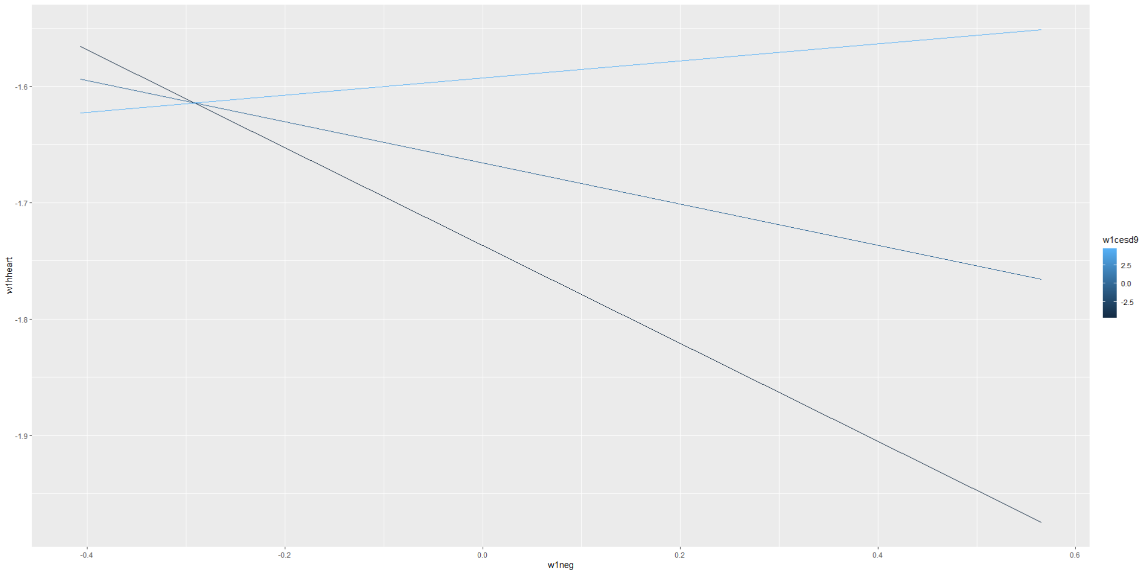
```
> source('c:/jason/R/macros/process.R',echo=FALSE)

***************** PROCESS for R Version 3.5.3 beta0.6 *****************

           Written by Andrew F. Hayes, Ph.D.  www.afhayes.com
   Documentation available in Hayes (2018). www.guilford.com/p/hayes3

***********************************************************************

PROCESS is now ready for use.
Copyright 2020 by Andrew F. Hayes ALL RIGHTS RESERVED

Distribution of this beta release of PROCESS is prohibited
without written authorization from the copyright holder.

> process(data=d,y="w1hheart",x="w1neg",w="w1cesd9",model=1,center=1,moments=1,plot=1)

***************** PROCESS for R Version 3.5.3 beta0.6 *****************

           Written by Andrew F. Hayes, Ph.D.  www.afhayes.com
   Documentation available in Hayes (2018). www.guilford.com/p/hayes3

***********************************************************************

Model : 1
    Y : w1hheart
    X : w1neg
    W : w1cesd9

Sample size: 692


***********************************************************************
Outcome Variable: w1hheart

Coding of binary Y for logistic regression analysis:
   w1hheart  Analysis
     0.0000    0.0000
     1.0000    1.0000
```

```
Model Summary:
       -2LL    ModelLL         df          p  McFadden  CoxSnell  Nagelkrk
   619.1451     6.5729     3.0000     0.0868    0.0105    0.0095    0.0159

Model:
             coeff         se          Z          p       LLCI       ULCI
constant   -1.6659     0.1074   -15.5184     0.0000    -1.8763    -1.4555
w1neg      -0.1771     0.2238    -0.7914     0.4287    -0.6158     0.2615
w1cesd9     0.0155     0.0233     0.6657     0.5056    -0.0301     0.0611
Int_1       0.0531     0.0256     2.0715     0.0383     0.0029     0.1032

These results are expressed in a log-odds metric.

Product terms key:
Int_1  :  w1neg  x  w1cesd9

Likelihood ratio test of highest order
unconditional interaction(s):
       Chi-sq        df         p
X*W    4.5505    1.0000    0.0329
----------
Focal predictor: w1neg (X)
      Moderator: w1cesd9 (W)

Conditional effects of the focal predictor at values of the moderator(s):
    w1cesd9     effect         se          Z          p       LLCI       ULCI
    -4.5898    -0.4206     0.2969    -1.4167     0.1566    -1.0026     0.1613
     0.0000    -0.1771     0.2238    -0.7914     0.4287    -0.6158     0.2615
     4.7259     0.0736     0.1995     0.3689     0.7122    -0.3173     0.4645

Data for visualizing the conditional effect of the focal predictor:
      w1neg    w1cesd9   w1hheart       prob
    -0.4070    -4.5898    -1.5658     0.1728
     0.0000    -4.5898    -1.7370     0.1497
     0.5657    -4.5898    -1.9749     0.1219
    -0.4070     0.0000    -1.5939     0.1688
     0.0000     0.0000    -1.6659     0.1590
     0.5657     0.0000    -1.7661     0.1460
    -0.4070     4.7259    -1.6227     0.1648
     0.0000     4.7259    -1.5928     0.1690
     0.5657     4.7259    -1.5511     0.1749

******************** ANALYSIS NOTES AND ERRORS ************************

Level of confidence for all confidence intervals in output: 95

W values in conditional tables are the minimum, the mean, and 1 SD above the mean.

NOTE: One SD below the mean is below the minimum observed in the data for W,
      so the minimum measurement on W is used for conditioning instead.

NOTE: The following variables were mean centered prior to analysis:
         w1cesd9 w1neg

NOTE: Some cases with missing data were deleted. The number of deleted cases was: 32
```

Plotting in R

First save the data under the section above “Data for visualizing the conditional effect of the focal predictor” as a text file (.txt). Then create a new R script file like the following.

```
> rm(d)
>
> d <- read.table(file=paste('c:/jason/R/cdaclass/heartplotdata.txt',sep="/"),header=TRUE)
>
> library(ggplot2)
> ggplot(data = d, aes(x=w1neg, y=w1hheart,group=w1cesd9, color=w1cesd9)) +
+   geom_line()
```



SAS

Using the PROCESS macro in SAS

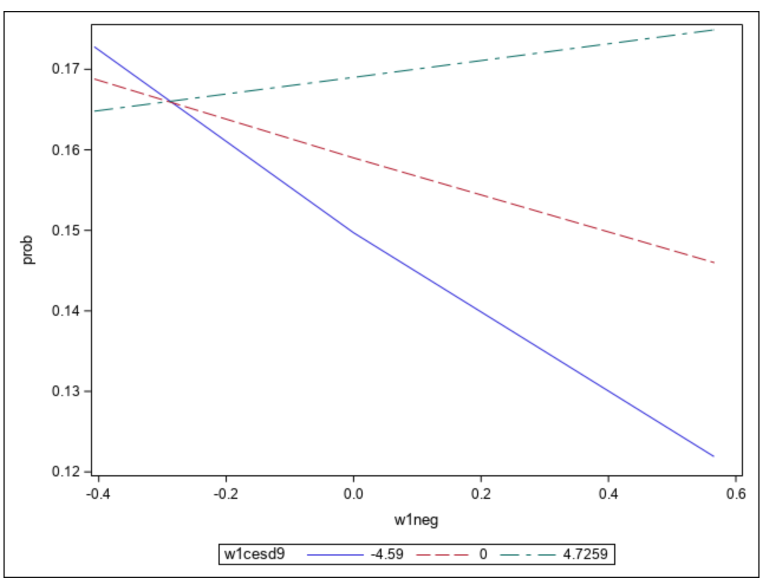
```
dm "output;clear;log;clear";
ods results off; ods listing;
options ps=60 ls=120;  *ls=240;

title 'Heart disease data from the LLSSE';

proc import datafile="c:\jason\spsswin\cdaclass\heart.sav" out=one dbms = sav replace;
run;

OPTIONS MSTORED SASMSTORE=macros;
%include"c:\jason\sas\macros\process.sas";
%process(data=one, y=w1hheart,x=w1neg,w=w1cesd9,model=1,center=1,plot=1,moments=1);
run;
```

```
                    Heart disease data from the LLSSE            14:17 Monday, April 26, 2021   3
          ******************************   PROCESS v3.5.3 for SAS   ******************************
                        Written by Andrew F. Hayes, Ph.D.  http://www.afhayes.com
                       Documentation available in Hayes (2018) www.guilford.com/p/hayes3
         ********************************************************************************************
                                      PROCESS is now ready for use.
                              Copyright 2020 by Andrew F. Hayes. ALL RIGHTS RESERVED.
```

```
*****************************   PROCESS v3.5.3 for SAS   ******************************
         Written by Andrew F. Hayes, Ph.D.  http://www.afhayes.com
      Documentation available in Hayes (2018) www.guilford.com/p/hayes3
*****************************************************************************************
                            Model and Variables
                            Model:     1
                                Y:     W1HHEART
                                X:     W1NEG
                                W:     W1CESD9
                                Sample size:
                                         692
*****************************************************************************************
                             OUTCOME VARIABLE:
                             W1HHEART
            Coding of binary Y for logistic regression analysis:
                             W1HHEART Analysis
                                 0.00     0.00
                                 1.00     1.00
                                Model Summary
       -2LL    ModelLL         df          p   McFadden   CoxSnell   Nagelkrk
   619.1451     6.5729     3.0000     0.0868     0.0105     0.0095     0.0159
                                   Model
                coeff         se          Z          p       LLCI       ULCI
  constant    -1.6659     0.1074   -15.5184     0.0000    -1.8763    -1.4555
  W1NEG       -0.1771     0.2238    -0.7914     0.4287    -0.6158     0.2615
  W1CESD9      0.0155     0.0233     0.6657     0.5056    -0.0301     0.0611
  Int_1        0.0531     0.0256     2.0715     0.0383     0.0029     0.1032
              These results are expressed in a log-odds metric
                            Product terms key:
         Int_1   :       W1NEG   x       W1CESD9
         LR test(s) of highest order unconditional interactions:
                      Chi-sq              df               p
        X*W           4.5505           1.0000           0.0329
                      --------------------------
                           Focal predict: W1NEG    (X)
                              Mod var: W1CESD9  (W)
  Conditional effects of the focal predictor at values of the moderator(s):
    W1CESD9     Effect         se          Z          p       LLCI       ULCI
    -4.5898    -0.4206     0.2969    -1.4167     0.1566    -1.0026     0.1613
     0.0000    -0.1771     0.2238    -0.7914     0.4287    -0.6158     0.2615
     4.7259     0.0736     0.1995     0.3689     0.7122    -0.3173     0.4645
     Data for visualizing the conditional effect of the focal predictor:
                        W1NEG    W1CESD9   W1HHEART       prob
                      -0.4070    -4.5898    -1.5658     0.1728
                      -0.0000    -4.5898    -1.7370     0.1497
                       0.5657    -4.5898    -1.9749     0.1219
                      -0.4070     0.0000    -1.5939     0.1688
                      -0.0000     0.0000    -1.6659     0.1590
                       0.5657     0.0000    -1.7661     0.1460
                      -0.4070     4.7259    -1.6227     0.1648
                      -0.0000     4.7259    -1.5928     0.1690
                       0.5657     4.7259    -1.5511     0.1749
******************************** ANALYSIS NOTES AND ERRORS ********************************
         Level of confidence for all confidence intervals in output:
                                                             95.0000
    W values in conditional tables are the minimum, the mean, and 1 SD above the mean
       NOTE: One SD below the mean is below the minimum observed in the data for W,
               so the minimum measurement on W is used for conditioning instead.
           NOTE: The following variables were mean centered prior to analysis: W1CESD9
W1NEG
```

Plotting in SAS
Using the above values under the "Data for visualizing the conditional effect of the focal predictor" section. Note that you must comment out the ods options lines and I needed to restart SAS before running this syntax to get the plot.

```
dm "output;clear;log;clear";
* ods results off; *ods listing;  *comment out both commands in this line to get
figure;
options ps=60 ls=120;  *ls=240; *use to avoid wrapping altogether;

*enter data from Process macro output;
data one;
  input w1neg w1cesd9 w1heart prob;
  cards;
 -0.4070    -4.5898    -1.5658     0.1728
 -0.0000    -4.5898    -1.7370     0.1497
  0.5657    -4.5898    -1.9749     0.1219
 -0.4070     0.0000    -1.5939     0.1688
 -0.0000     0.0000    -1.6659     0.1590
  0.5657     0.0000    -1.7661     0.1460
 -0.4070     4.7259    -1.6227     0.1648
 -0.0000     4.7259    -1.5928     0.1690
 0.5657     4.7259    -1.5511     0.1749
;
run;

*specifies the location of the file and file format (I used png graphics file);
ods listing gpath="c:\jason\sas\cdaclass\";
ods graphics / imagename="interaction" imagefmt=png;

proc sgplot data=one;
series x=w1neg y=prob /group=w1cesd9
  legendlabel='Values of CESD';
run;
```



You can also plot manually using Jeremy Dawson's excel sheet,
http://www.jeremydawson.com/slopes.htm

Sample Write-Up

Logistic regression was used to investigate whether depression might moderate the effects negative social exchanges on self-reported heart disease. Both predictors were centered around their means (Aiken & West, 1991) before computing the interaction term, and all terms were entered into the model together. The results indicated a significant interaction, *b* = .053, *SE* = .026, *p* = .04, 95%CI[.002, .10]. To probe the interaction, simple effects coefficients were computed or three values of depression, 1 SD below the mean, at the mean, and 1 SD above the mean. At high levels of depression, more negative social exchanges were associated with slightly higher, but nonsignificant, increase of odds of heart disease, *b* = .074, SE = .200, *OR* = 1.076, p = .71, 95%CI[-.32, .46]. At the mean depression level, negative social exchanges were associated with a slight nonsignificant decrease in odds of heart disease, b = -.177, SE = .224, OR = .838, p = .43, 95%CI[-.62,.26]. Negative social exchanges had slightly stronger negative relation to hear disease for high levels of depression, although this effect was also nonsignificant, *b* =-.421, SE = .270, OR = .657, p = .157, 95%CI[-1.002,.16]. Figure 1 graphs the interaction, showing the change in the expected probability of heart by negative social exchanges for depression at -1 SD and +1 SD from the mean. Overall, the significant interaction suggests that negative social exchanges have slightly different relationship to self-reported heart disease depending on the level of depression also experienced, but the magnitude of the increased risk or decreased risk associated with negative social exchanges appears to be fairly minimal.

References and Further Readings

Aiken, L. S., & West, S. G. (1991). Multiple regression: Testing and interpreting interactions. Thousand Oaks, CA: Sage Publications Inc.
Allison, P. D. (2014). Measures of fit for logistic regression. *SAS Global Forum, Washington, DC*.
Fisher, R.A. (1926). The Arrangement of Field Experiments. *Journal of the Ministry of Agriculture of Great Britain, 33,* 503–513.
Hardy, M. A. (1993). *Regression with dummy variables (Vol. 93)*. Sage.
Hayes, A. F., & Matthes, J. (2009). Computational procedures for probing interactions in OLS and logistic regression: SPSS and SAS implementations. *Behavior Research Methods, 41*(3), 924-936.
Jaccard, J. (2001). *Interaction effects in logistic regression* (Vol. 135). Thousand Oakes, CA: Sage.
Johnson, P.O., & Neyman J. (1936). Tests of certain linear hypothesis and their application to some educational problems. *Statistical Research Memoirs, 1*, 57–93.