Paper RI-06

Don't Avoid It, Exploit It: Using Annotate to Enhance Graphical Output

Sarah Mikol, Rho, Inc., Chapel Hill, NC

ABSTRACT

SAS/GRAPH® is widely used for displaying the results of an analysis or simply creating graphical summaries of the data. Often times it is necessary to customize these graphics to either meet client needs or clarify the output. SAS/GRAPH's customization tool, ANNOTATE, provides the ability to either alter the default graphical elements that are displayed or add new elements to the output. However, many users are intimidated by the intricate syntax of the Annotate facility and may settle for graphics that do not provide a clear, comprehensive presentation of the results. This paper will first provide an introduction to the cornerstone of the Annotate facility, the Annotate dataset, and will explain the details of its construction. A summary of the SAS/GRAPH annotation coordinate system, as well as Annotate macros, which offer a more compact way to produce the annotation dataset, will be given. Lastly, two examples considering both data driven and predetermined modifications will show how to incorporate custom text, symbols, and line segments into graphics produced by both GPLOT and GCHART.

INTRODUCTION

It is often necessary in clinical research to customize graphs beyond the default outputs. You may need to display study data more completely or just differently in order to meet client needs and/or wishes. This paper is geared towards either beginning or experienced SAS® users with an understanding of the DATA step and The GPLOT and GCHART Procedures in SAS/GRAPH. The Annotate facility provides the opportunity to customize graphics with text, lines, and objects of any color and format and at any location. This paper aims to ease readers into the Annotate facility and provide examples of several modifications that can make a display more comprehensive and aesthetically pleasing to a client.

THE ANNOTATE DATA SET

The Annotate data set is the cornerstone of the Annotate facility. The data set and the ANNOTATE= option in a SAS/GRAPH procedure are all that are needed to add annotations to your graphics. The Annotate data set is a standard SAS data set with predetermined variable names that specify either annotation functions or modifications to those functions. The variables are instructions for creating annotations on a graph, while the observations in the data set are annotation commands.

ANNOTATE VARIABLES

The Annotate variables used in this paper are listed in Table 1 and described as they relate to the examples. These variables can be either a parameter in an Annotate macro or defined separately in the Annotate data set. The key variable, FUNCTION, tells SAS which action to perform. The remaining variables modify that particular function.

A complete list of variables and the possible values for each can be found in the SAS/GRAPH Annotate Dictionary.

Task	Variable	Description
Define an action	FUNCTION	Specifies a drawing or programming action
Specify a position	X, Y	Numeric horizontal and vertical coordinates
	XSYS, YSYS	Coordinate system for X coordinates, Y coordinates
Assign attributes	COLOR	Color of text, line, or object
	LINE	Line type to use in drawing
	POSITION	Placement and alignment for text strings
	ANGLE	Angle of text label
	ROTATE	Angle at which to place individual characters in a text string
	SIZE	Height or line thickness
	STYLE	Font or pattern
	TEXT	Text to use in a label

Task	Variable	Description
	WHEN	Whether an element is drawn before or after procedure graphics output

Table 1. Subset of Annotate Data Set Variables

COORDINATE SYSTEMS

Coordinate systems are used in the Annotate facility to identify the location where annotations will be placed, or 'drawn.' The coordinate system variables XSYS and YSYS are required in the Annotate data set and are assigned based on combinations of drawing areas and units, which are described in Table 2.

SAS recognizes three drawing areas: the data area, the graphics output area, and the procedure output area. The data area is the rectangle bound by the horizontal and vertical axes. The graphics output area is the total area available to the device on screen. Lastly, the procedure output area is the graphics output area minus the space allotted for footnotes, titles, and legends.

Either absolute values, which identify the exact location in the output, or relative values, which give an offset from the previous location, can be specified for the coordinate system. In this paper, only absolute coordinate system values will be discussed. Variables XSYS and YSYS can take on the absolute values of 1 through 6, as seen in Table 2.

Frame of Reference	Units	Value	Range
Data Area	Percents	1	0% to 100% of axis area
Data Area	Data Values	2	Minimum axis value to maximum axis value
Graphics Output area	Percents	3	0% to 100% of graphics output area
Graphics Output Area	Cells	4	0 to edge of graphics output area
Procedure Output Area	Percents	5	0% to 100% of procedure output area
Procedure Output Area	Cells	6	0 to edge of procedure output area

Table 2. Absolute Coordinate System Values

ANNOTATE MACROS

Annotate macros are a useful option for programmers who have previously been intimated by the cumbersome assignment statements necessary to build an Annotate data set.

Annotate macros essentially provide short cuts for creating observations in the Annotate dataset. The macros generate assignment statements that the programmer would otherwise create manually. Within the Annotate macros, you are given control of the basic variables associated with a particular function.

The SAS code below shows how to manually create several observations in an Annotate data set. This involves defining the function variable and all modifications as well as employing an OUTPUT statement. A RETAIN statement is used to define the values of XSYS and YSYS and retain their values for all observations until they are redefined.

```
DATA anno;
  LENGTH function $10 text $20;
  RETAIN xsys ysys '1';

  function='label'; x=50; y=50; text='Look What I Can Do!'; color='maroon';
    angle=0; rotate=0; position='5'; style='centx'; size=5; output;
  function='move'; x=15; y=35; output;
  function='draw'; x=15; y=60; color='maroon'; line=1; size=2; output;
  function='draw'; x=85; y=60; color='maroon'; line=1; size=2; output;
  function='draw'; x=85; y=35; color='maroon'; line=1; size=2; output;
  function='draw'; x=15; y=35; color='maroon'; line=1; size=2; output;
RUN;
```

The same data set above (ANNO) containing several observations can be produced in a more straightforward manner. The following DATA step utilizes Annotate macros instead of the assignment statements in the above code.

```
%ANNOMAC()
DATA anno;
  %DCLANNO
  %SYSTEM(1, 1)

  %LABEL(50, 50, 'Look What I Can Do!', maroon, 0, 0, 5, centx, 5)
  %LINE(15, 35, 85, 35, maroon, 1, 2)
  %LINE(15, 60, 85, 60, maroon, 1, 2)
  %LINE(15, 35, 15, 60, maroon, 1, 2)
  %LINE(85, 35, 85, 60, maroon, 1, 2)
RUN;
```

Both methods produce the same Annotate data set and the same graphical output, as seen below in Figure 1.
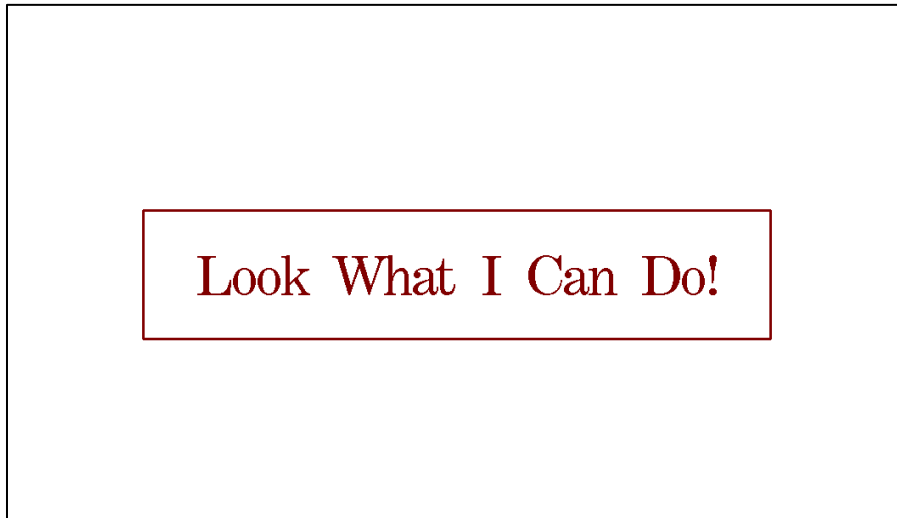

Figure 1. A Simple Example

Next, we will discuss the two types of Annotate macros: those that prepare and manage the environment and those that define functions.

PREPARATION MACROS

There are three macros that prepare the graphics environment for annotations: %ANNOMAC, %DCLANNO, and %SYSTEM. Table 3 describes these macros.

Macro	Parameters	Description	Notes
%ANNOMAC	None	Compiles the ANNOTATE macros and prepares them for use	Must be submitted before using other ANNOTATE macros
%DCLANNO	None	Sets correct length and data type for the ANNOTATE variables	Length for the TEXT variable is not set
%SYSTEM	(XSYS, YSYS)	Defines coordinate system for annotations	Coordinates can be redefined for each observation in the ANNOTATE data set

Table 3. Annotate Preparation Macros

FUNCTION MACROS

There are many Annotate macros available, but two of the most commonly used, %LABEL and %LINE, along with %ARROW will be discussed in this paper. The %LABEL macro adds text, %LINE draws a line, and %ARROW draws an arrow. The arguments may be numeric values, variable names, or strings. It is good practice to specify all arguments, and although missing values may resolve to the default value, it is common to receive errors as a

consequence of blank arguments. The parameters of each macro are described in the three tables below. See the SAS/GRAPH Annotate Dictionary for a list of possible values for each parameter.

%LABEL (x, y, text, color, angle, rotate, size, style, position)

Parameter	Description	Allowed Values
X, Y	Identify the text location	Coordinate values, numeric constants, or numeric variables
TEXT	Label text	Name of a character variable name or a string in quotation marks
COLOR	Text color	Character string (no quotation marks)
ANGLE	Angle of the text string with respect to the horizontal	Number, numeric constant, or numeric variable
ROTATE	Angle of each character in text string	Number, numeric constant, or numeric variable
SIZE	Size (height) of text	Number, numeric constant, or numeric variable
STYLE	Font	Character string (no quotation marks)
POSITION	Placement /alignment of text string	Text string (no quotation marks)

Table 4. %LABEL Macro Parameters

%LINE (x1, y1, x2, y2, color, line, size)

Parameter	Description	Allowed Values
X1, Y1	Starting coordinates of the line	Number, numeric constant, or numeric variable
X2, Y2	Ending coordinates of the line	Number, numeric constant, or numeric variable
COLOR	Color of the line	Character string (no quotation marks)
LINE	Type of the line (either continuous or dotted)	Number, numeric constant, or numeric variable
SIZE	Width of the line	Number, numeric constant, or numeric variable

Table 5. %LINE Macro Parameters

%ARROW (x1, y1, x2, y2, color, line, size, angle, style)

Parameter	Description	Allowed Values
X1, Y1	Starting coordinates for the arrow	Coordinate numbers, numeric constants, or numeric variables
X2, Y2	Ending coordinates for the arrow	Coordinate numbers, numeric constants, or numeric variables
COLOR	Color of the arrow line	Character string (no quotation marks)
LINE	Length of arrowhead sides	Number, numeric constant, or numeric variable
SIZE	Width of arrow line	Number, numeric constant, or numeric variable
ANGLE	Angle of arrowhead tip	Number, numeric constant, or numeric variable
STYLE	Type of arrowhead	CLOSED, FILLED, or OPEN

Table 6. %ARROW Macro Parameters

EXAMPLE 1: DATA DRIVEN ANNOTATIONS WITH PROC GPLOT

The first example deals with values of two lab tests, ALT and AST, over the course of a study for a particular subject. The data also includes specific instances in time (weeks) where the subject had either a dose reduction or interruption. PROC GPLOT is implemented to plot the lab results over time, and data driven annotations note where the subject had reductions or interruptions in a dose.

Also addressed in this example are situations where predetermined annotations must be included. The partial data sets for ALT and AST lab data and dose reduction/interruption data used for plotting and annotations, respectively, are shown in Table 7 and Table 8.

TEST	RESULT	TIMEWKS
ALT	17	0.00000
ALT	105	1.71429
ALT	84	2.00000
ALT	57	3.14286
ALT	93	4.14286

Table 7. LABS Data Set Used in PROC GPLOT

GROUP	TIMEWKS
DOSE INTERRUPTION	1.1429
DOSE INTERRUPTION	1.8571
DOSE INTERRUPTION	2.5714
DOSE INTERRUPTION	4.2857
DOSE REDUCTION	4.7143

Table 8. DOSE Data Set Used For Annotation

The following code prepares the environment by setting graphics options and specifying text options and title statements.

```
GOPTIONS RESET=ALL DEVICE=PNG TARGET=PNG FTITLE='Arial' FTEXT='Arial' HTITLE=1.5
    HTEXT=1;
FILENAME file2 "C:\SESUG\GPLOT.png";

OPTIONS ORIENTATION=LANDSCAPE NONUMBER NODATE CENTER;
TITLE color=black h=1.5 justify=center 'ALT and AST Laboratory Results Over Time';
TITLE2 color=black h=1.2 justify=center '(In Relation to Dose Reductions and
    Interruptions)';
TITLE3 color=black h=1.2 justify=center 'Subject 000001';
FOOTNOTE;
```

The SAS code below creates a data set (ANNOLEGEND) for data driven annotations.

In ANNOLEGEND, %SYSTEM(2, 2) allows for the lines and text for the dose reductions and interruptions legend to be drawn in the data area based on the actual X and Y axis values. See Table 2 for a refresher of coordinate system values, Table 4 for to reference the %LABEL macro parameters, and Table 5 for a description of the %LINE macro parameters.

```
%ANNOMAC()
DATA annolegend;
  %SYSTEM(2, 2)

  %LABEL(47, 485, 'Dose Interruption', black, 0, 0, 1, 'Arial', 6)
  %LINE(42, 485, 46, 485, black, 2, .75)

  %LABEL(47, 470, 'Dose Reduction', black, 0, 0, 1, 'Arial', 6)
  %LINE(42, 470, 46, 470, black, 8, .75)
RUN;
```

The next part of the annotation creates a data set (ANNOLINES) that includes a mix of data driven and predetermined annotations and is based on the DOSE data set in Table 8. ANNOLINES contains instructions to draw lines from the bottom to the top of the y-axis, with different line types based on whether a dose reduction or interruption occurred.

For this purpose, the value of YSYS must be changed to 1, which specifies a percentage of the data area. The system coordinates can be re-specified with %SYSTEM(2, 1) and then the Y1 value can be assigned to 0(%) and the Y2 value to 100(%). In this case, the reason we have a mix of data driven and predetermined annotations is that the

placement of lines on the x-axis is based on the variable TIMEWKS in the DOSE data set and the line itself simply spans from 0 to 100% of the data area.

See Table 2 for a refresher of coordinate system values and Table 5 for a reminder of the %LINE macro parameters.

```
DATA annolines;
   SET dose;
   %SYSTEM(2, 1)

   if group='DOSE INTERRUPTION' then do;
     %LINE(timewks, 0, timewks, 100, black, 2, 1)
   end;

   if group='DOSE REDUCTION' then do;
     %LINE(timewks, 0, timewks, 100, black, 8, 1)
   end;
RUN;
```

Since some annotations depend on a previously created data set while others do not, separate data sets were created and are set together in the code below for use as one data set in PROC GPLOT's ANNOTATE= option.

```
DATA anno;
   SET annolines
       annolegend;
RUN;
```

A sample of observations from the ANNO data set is given below in Table 9.

FUNCTION	XSYS	YSYS	X	Y	TEXT	SIZE	STYLE	COLOR	LINE
MOVE	2	1	1.1429	0		.		black	.
DRAW	2	1	1.1429	100		1.00		black	2
MOVE	2	1	1.8571	0		.		black	.
DRAW	2	1	1.8571	100		1.00		black	2
MOVE	2	1	2.5714	0		.		black	.
DRAW	2	1	2.5714	100		1.00		black	2
MOVE	2	1	4.2857	0		.		black	.
DRAW	2	1	4.2857	100		1.00		black	2
MOVE	2	1	6.5714	0		.		black	.
DRAW	2	1	6.5714	100		1.00		black	2
LABEL	2	2	47.0000	485	Dose Interruption	1.00	'Arial'	black	.
MOVE	2	2	42.0000	485	Dose Interruption	1.00	'Arial'	black	.
DRAW	2	2	46.0000	485	Dose Interruption	0.75	'Arial'	black	2
LABEL	2	2	47.0000	470	Dose Reduction	1.00	'Arial'	black	2
MOVE	2	2	42.0000	470	Dose Reduction	1.00	'Arial'	black	2
DRAW	2	2	46.0000	470	Dose Reduction	0.75	'Arial'	black	8

Table 9. Annotation Data Set ANNO Used in PROC GPLOT

Next, global statements for the axes, symbol values, and the lab test legend options are defined.

```
AXIS1 value=(h=1) order=(0 to 55 by 5) minor=(n=4) label=(h=1.25 "Time in Weeks")
     offset=(1);
AXIS2 value=(h=1) order=(0 to 500 by 50) minor=(n=1) label=(angle=90 h=1.25 'Lab
     Result (U/L)') offset=(1) origin=(, 20 pct);
```

```
SYMBOL1 v=none i=join l=1 c=black w=2;
SYMBOL2 v=none i=join l=4 c=black w=2;

LEGEND1 position=(bottom center outside) across=2 value=(j=c) shape=line(10)
label=(height=1 "Lab Test");
GOPTIONS GSFNAME=file2;
```

Now, all preparations are in place to plot the LABS data set (shown partially in Table 7) with PROC GPLOT. The ANNOTATE=anno option incorporates the instructions from the ANNO data set seen partially in Table 9. The LEGEND=legend1 option (see LEGEND1 statement above) places the lab test legend for ALT and AST line types below the graph, and the HAXIS= and VAXIS= options define which axis statements from the code above to assign to the X and Y axes.

```
PROC GPLOT DATA=labs ANNOTATE=anno;
    PLOT result*timewks=test / legend=legend1
                               haxis=axis1
                               vaxis=axis2;
RUN;
QUIT;
```

Below, Figure 2 shows the subject's lab values plotted over study time in weeks, displays a default and annotated legend, as well as incorporates vertical lines representing dose reductions and interruptions that occur over the course of the study.
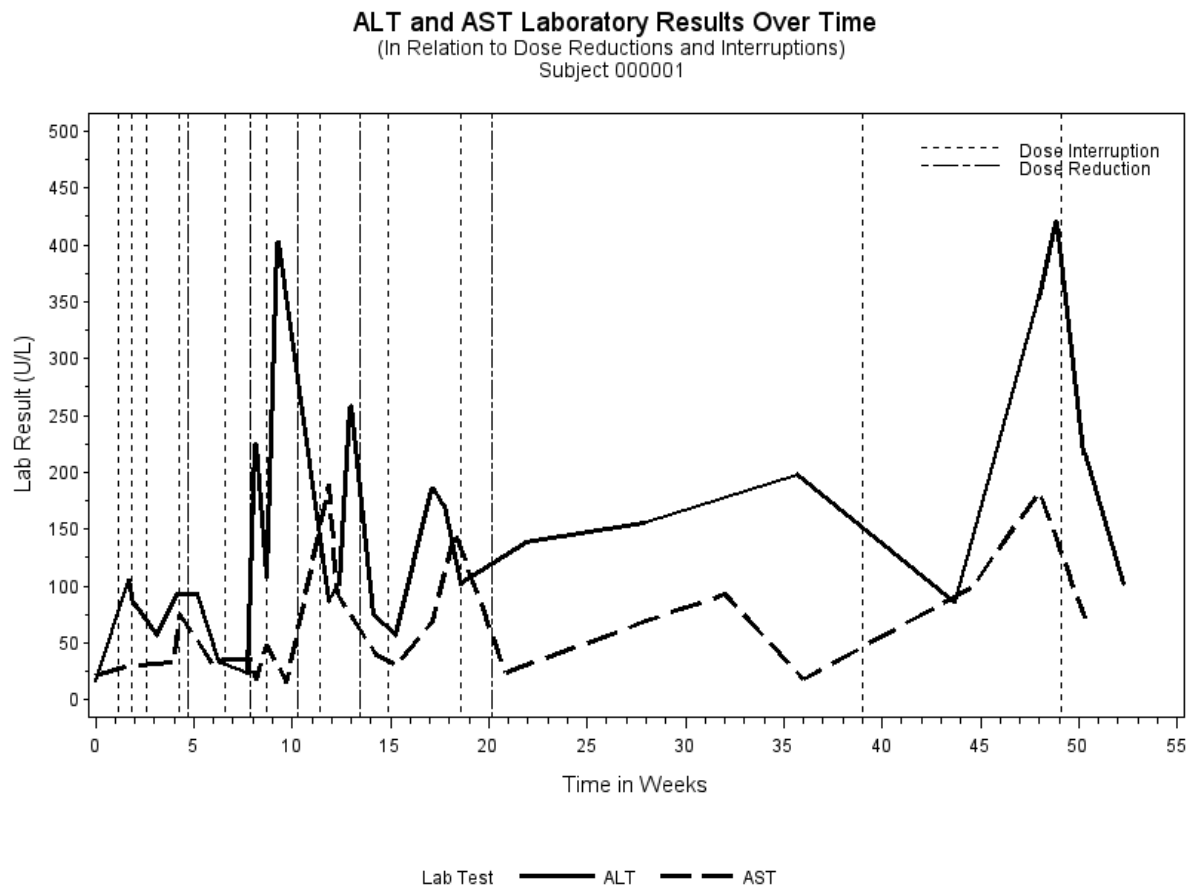

Figure 2. Plot of Lab Values over Time

EXAMPLE 2: ANNOTATIONS WITH PROC GCHART

The GCHART PROCEDURE offers limited control of the group axis (x-axis) values. Annotate allows modifications to the labels on the midpoint axis even without the midpoint annotate variable (see the SAS/GRAPH Annotate Dictionary), by using the discrete option (if the chart variable is numeric). Here TREATMENT is a numeric variable, so

we can specify numeric x-values in the annotation macros while XSYS is assigned a value of 2. Normally, the programmer has little control over the GAXIS while using PROC GCHART, but the methods below offer some options for customization.

The data used in this example deals with a contact lens comfort score given by patients at both a baseline and final study visit. The change in score between the visits is calculated. A change of 16 is the pre-specified threshold indicating no change in comfort. A change greater than 16 indicates the subject's contact lens comfort worsened, and a change in score less than 16 indicates the comfort between the visits improved. The first few observations of the comfort scores data set, SCORES, are displayed in Table 10.

TREATMENT	SUBJID	CHANGE
1	00001	12.5
1	00002	15.0
1	00003	5.0
1	00004	22.5
1	00005	17.5
1	00006	25.0
1	00007	22.5
1	00008	15.0
1	00009	15.0
1	00010	25.0
2	00011	17.5
2	00012	5.0

Table 10. SCORES Data Set Used in PROC GCHART

The MEANS PROCEDURE is used to calculate the total number of subjects (N) in each treatment group from the data set SCORES, and using The GLM PROCEDURE a one-way ANOVA model of change (from baseline) on treatment group is implemented and generates p-values (PVAL) for the three comparisons of each treatment group to the control group. From the output of these procedures a data set, FORANNO (Table 11), is created for use in the creation of the chart annotation data set.

TREATMENT	N	PVAL	NLABEL	PLABEL	XLABEL
1	10		N=10		Control
2	11	0.020	N=11	p=0.020	TREATMENT 1
3	4	0.410	N=4	p=0.410	TREATMENT 2
4	7	0.820	N=7	p=0.820	TREATMENT 3

Table 11. FORANNO Data Set for Annotation Preparation

The annotation code below creates the CHARTANNO1 data set, based on FORANNO, for data driven annotations. We need to employ the values of the group axis and the y-axis to place NLABEL and PLABEL on top of the bars, and treatment labels for the group axis under each bar. It is important to note that there is an option in the PROC GCHART VBAR statement (OUTSIDE=) to place a statistic on top of each bar. However, for the purposes of including the p-value on top of the bars as well, the frequency text (N=) is annotated. Note also that the group axis labels can be defined in an axis statement, but the advantage of annotating them is that it is possible, if needed, to change the angle or position of the label or break up lengthier text that may be truncated.

In CHARTANNO1, the text is set at a length of $20 so that the 'TREATMENT 3' label isn't truncated due to the length of the first label, 'CONTROL.' The WHEN = 'A' option tells the Annotate facility to draw the annotations AFTER the graph elements are drawn, so as to not cover up any specified annotations. By calling the first %SYSTEM macro, XSYS and YSYS are each assigned a value of 2. As we learned in the coordinates system section, XSYS=2 and

YSYS=2 indicate that the annotations will be drawn within the rectangle bound by both axes, and the actual data values will be used. NHEIGHT and PHEIGHT serve as values for the y location for NLABEL and PLABEL. The TREATMENT variable serves as the x location value, so the NLABEL and PLABEL will be written over each bar at a group values 1, 2, 3, and 4, at a height of 27 and 26, respectively. Note NHEIGHT and PHEIGHT are specified based on the maximum value of the y-axis, which is 30.

To draw the group axis labels, the YSYS coordinate must be changed by calling the %SYSTEM macro again. We assign the YSYS value to '3' (% of the graphics output area) and define Y=2, meaning that the labels will be placed at 2% of the graphics output area, just under the group axis.

```
%ANNOMAC()
DATA chartanno1;
  SET foranno;
  LENGTH TEXT $20;
  WHEN='A';

  NHEIGHT = 27;
  PHEIGHT = 26;

  %SYSTEM(2, 2)
  %LABEL(TREATMENT, NHEIGHT, NLABEL, black, 0, 0, 1.5, 'Arial', 5)
  %LABEL(TREATMENT, PHEIGHT, PLABEL, black, 0, 0, 1.5, 'Arial', 5)
  %SYSTEM(2, 3)
  %LABEL(TREATMENT, 2, XLABEL, black, 0, 0, 1.5, 'Arial', 5)
RUN;
```

The CHARTANNO2 data set (code follows below) contains instructions for drawing a dotted line across the graph to indicate where the threshold is set (Y=16) for improvement or worsening of contact lens comfort. Therefore, XSYS is assigned to '1' to specify that the horizontal line should span from 0% to 100% of the data area.

Now we will annotate arrows to the right of the threshold line, outside of the data area, to indicate the directions of improvement or worsening of comfort. The %SYSTEM macro is redefined to assign XSYS=5 but retains the YSYS value of 2. Therefore, the x values in the Annotate macros will specify percents of the procedure output area. This involves trial and error, and it is concluded that positioning the labels and arrows at 93% of the procedure output area is appropriate. See Figure 3 for the results of changing these coordinate system values.

```
DATA chartanno2;
  LENGTH TEXT $20;
  WHEN='A';

  %SYSTEM(1, 2)
  %LINE(0, 16, 100, 16, black, 3, 2)

  %SYSTEM(5, 2);
  %LABEL(93, 16, "Unchanged", black, 0, 0, 1.45, 'Arial', 5)
  %LABEL(93, 20, "Worse", black, 0, 0, 1.45, 'Arial', 5)
  %LABEL(93, 12, "Better", black, 0, 0, 1.45, 'Arial', 5)

  %ARROW (93, 16.5, 93, 19, black, 1, 1, 90, open)
  %ARROW (93, 15, 93, 12.5, black, 1, 1, 90, open)
RUN;
```

Since some of the annotations in this example are data driven while others are predetermined, separate data sets were necessary and are set together below for use as one data set in the ANNOTATE= option in PROC GCHART. If the predetermined modifications were specified along with the data driven ones using the FORANNO data set in CHARTANNO1, there would be duplicate observations for the text and lines, one for each record in the original data set. The annotations would appear bold and/or distorted on the graph as they would be drawn over top of each other multiple times.

```
DATA chartanno;
  SET chartanno1
      chartanno2;
RUN;
```

The following data set, CHARTANNO, is now ready to be implemented in PROC GCHART.

FUNCTION	XSYS	YSYS	X	Y	TEXT	WHEN	SIZE	STYLE	COLOR	LINE
LABEL	2	2	1	27.0	N=10	A	1.50	'Arial'	black	.
LABEL	2	2	1	26.0		A	1.50	'Arial'	black	.
LABEL	2	3	1	2.0	Control	A	1.50	'Arial'	black	.
LABEL	2	2	2	27.0	N=11	A	1.50	'Arial'	black	.
LABEL	2	2	2	26.0	p=0.020	A	1.50	'Arial'	black	.
LABEL	2	3	2	2.0	TREATMENT 1	A	1.50	'Arial'	black	.
LABEL	2	2	3	27.0	N=4	A	1.50	'Arial'	black	.
LABEL	2	2	3	26.0	p=0.410	A	1.50	'Arial'	black	.
LABEL	2	3	3	2.0	TREATMENT 2	A	1.50	'Arial'	black	.
LABEL	2	2	4	27.0	N=7	A	1.50	'Arial'	black	.
LABEL	2	2	4	26.0	p=0.820	A	1.50	'Arial'	black	.
LABEL	2	3	4	2.0	TREATMENT 3	A	1.50	'Arial'	black	.
MOVE	1	2	0	16.0		A	.		black	.
DRAW	1	2	100	16.0		A	2.00		black	3
LABEL	5	2	93	16.0	Unchanged	A	1.45	'Arial'	black	3
LABEL	5	2	93	20.0	Worse	A	1.45	'Arial'	black	3
LABEL	5	2	93	12.0	Better	A	1.45	'Arial'	black	3
MOVE	5	2	93	16.5	Better	A	1.45	'Arial'	black	3
ARROW	5	2	93	19.0	Better	A	1.00	open	black	1
MOVE	5	2	93	15.0	Better	A	1.00	open	black	1
ARROW	5	2	93	12.5	Better	A	1.00	open	black	1

Table 12. Final Data Set (CHARTANNO) for Annotations

The following code prepares the graph environment, defines the title and axis statements, and specifies other output options.

```
GOPTIONS RESET=ALL DEVICE=PNG TARGET=PNG HSIZE=28cm VSIZE=16cm FTITLE='Arial'
  FTEXT='Arial' HTITLE=1.75 HTEXT=1.5;
FILENAME file1 "C:\SESUG\barchart.png";

OPTIONS ORIENTATION=LANDSCAPE NONUMBER NODATE CENTER;
TITLE COLOR=BLACK JUSTIFY=CENTER 'CHANGE in Contact Lens Comfort Score';
TITLE2 COLOR=BLACK JUSTIFY=CENTER '(From Baseline to Final Visit)';
FOOTNOTE;

AXIS1 label=(a=90 "Mean Score Change (with 95% Confidence Intervals)") order=(0 to
  30 by 5) minor=none value=(font='Arial' h=1.5);
AXIS2 label=none value=none;

GOPTIONS GSFNAME=file1;
```

Below, the PROC GCHART code produces Figure 3. The discrete treatment variable from the data set SCORES (Table 10) is used as the chart variable, and mean of the CHANGE variable by treatment group is displayed on the y-axis (SUMVAR=CHANGE and TYPE=MEAN options). The ERRORBAR=BOTH option automatically draws 95% confidence bands on the bars. The other options specify the horizontal reference lines, the width and color of the bars, and which axis statements (created above) correlate with the y-axis (RAXIS) and the group axis (MAXIS).

```
PROC GCHART DATA = scores ANNOTATE = chartanno;
    VBAR TREATMENT / DISCRETE SUMVAR = CHANGE
                               TYPE = MEAN
                               SPACE = 16
                               WIDTH = 12
                               ERRORBAR = BOTH
                               CLM = 95
                               RAXIS = AXIS1
                               MAXIS = AXIS2
                               AUTOREF
                               CAUTOREF = VLIGB
                               CLIPREF
                               COUTLINE = BLACK;
        PATTERN COLOR = 'Light Gray';
    RUN;
    QUIT;
```

Of course, the most important option pertaining to the paper focus is the ANNOTATE=chartanno option in the PROC GCHART statement. See below in Figure 3 that all annotation instructions listed in the CHARTANNO data set (Table 12) are followed and the graph modifications are displayed on the bar chart. The annotated elements include the labels above the bars for N's by group and p-values for each treatment compared with the control, the score threshold dotted line at Y=16, the group axis treatment labels, and the improvement/worsening arrows to the right of the graph within the procedure output area but outside the data area.
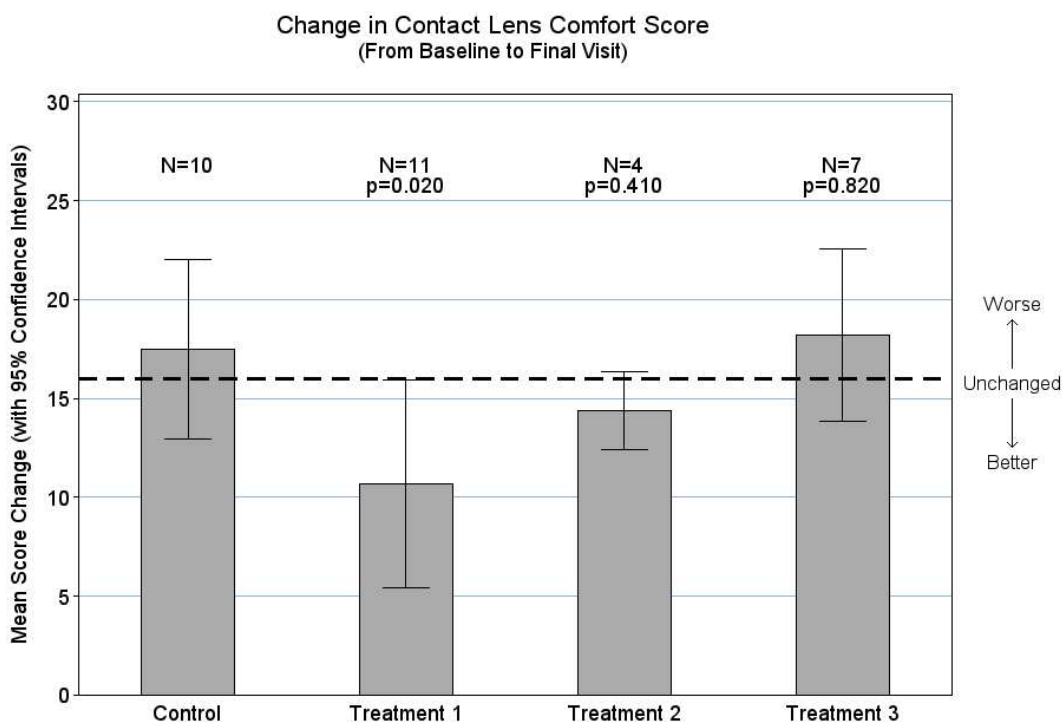

Figure 3. Bar Chart Displaying Mean Comfort Score Change by Treatment Group

CONCLUSION

The Annotate facility can be intimidating to even an experienced SAS/GRAPH programmer. However, with an understanding of the Annotate data set, coordinate systems, and the Annotate macros, great potential exists to overcome the limitations of default outputs available in SAS/GRAPH. The power of a DATA step paired with the SAS Macro Language offers the ability to produce a variety of both data driven and predetermined annotations that improve the clarity and aesthetics of graphical presentations.

REFERENCES

SAS Institute, Inc (2012) SAS/Graph® 9.2 Reference, 2nd Edition: Annotate Dictionary, Cary, NC: SAS Institute, Inc.

ACKNOWLEDGMENTS

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

Name: Sarah Mikol
Enterprise: Rho, Inc.
Address: 6330 Quadrangle Dr.
City, State ZIP: Chapel Hill, NC 27517
Work Phone: (919) 408-8000
Fax: (919) 408-0999
E-mail: sarah_mikol@rhoworld.com