Growth Curve Modeling

Contents

Example Data Set: Kashy

Outcome:

- `ASATISF` Satisfaction

Predictor Variables:

- `TIME`: time in days (there are 14 days with 0 = study midpoint).

- `GENDER`: Gender effects coded (Women = -1 and Men = 1).

- `GenderS`: A string variable that labels women "Woman" and men "Man".

Moderators:

- `CAAvoid`, `CPAvoid`: Grand mean centered attachment avoidance (Actor and Partner—a mixed moderator)

First, read in the new dataset. It's already in the person-period pairwise structure.

```
library(dyadr)
library(dplyr)
library(nlme)
library(ggplot2)

kashy_ppp <- read.csv("/Users/randigarcia/Desktop/Data/kashy.csv", header=TRUE)

View(kashy_ppp)
```

Individual Growth Curve Modeling

First, for illustration purposes, we want to run a growth curve model for men only. We can use the Kashy data but select only men with this syntax:

```
kashy_men <- kashy_ppp %>%
  filter(GENDER == 1)
```
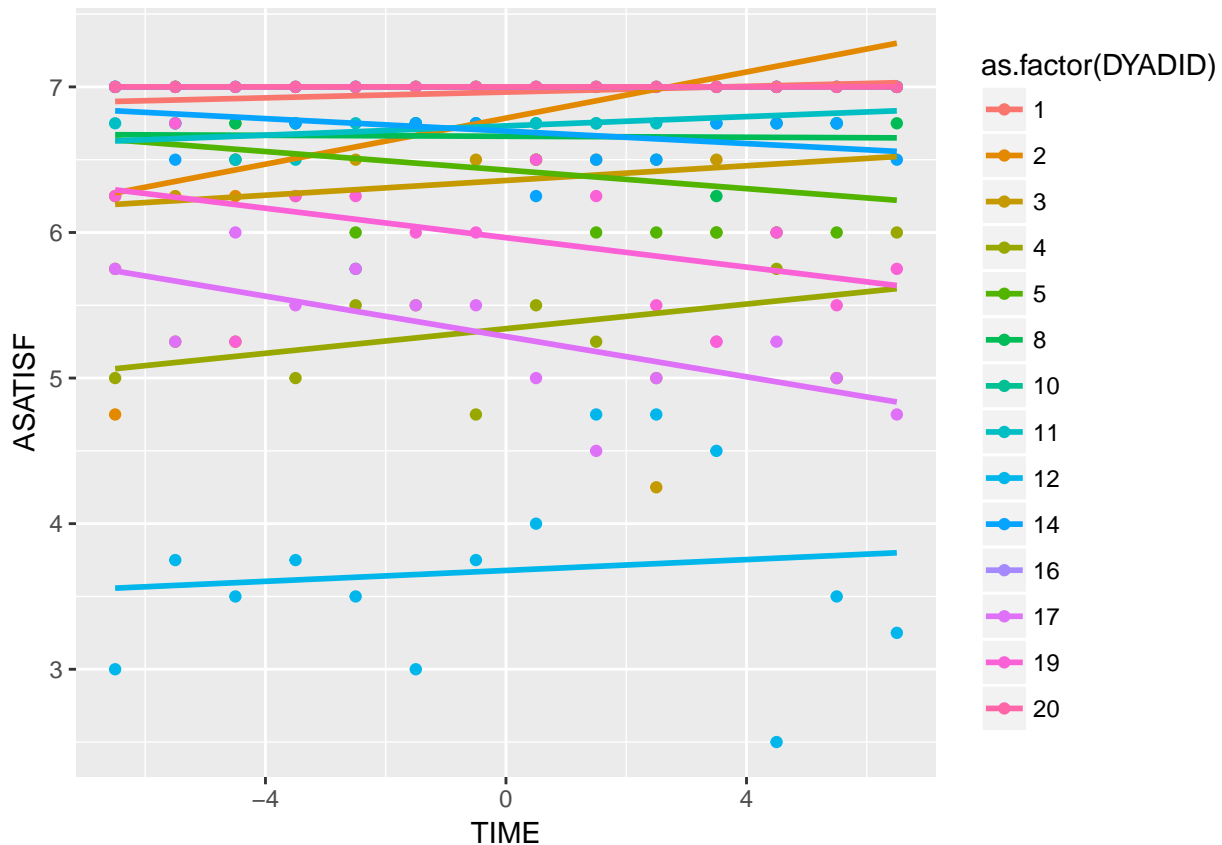
Spaghetti Plot

Spaghetti plot shows different slopes for different folks!

```
kashy_men_small <- kashy_men %>%
  filter(DYADID >= 1 & DYADID <= 20)

ggplot(kashy_men_small, aes(TIME, ASATISF,
                            group = as.factor(DYADID),
                            color = as.factor(DYADID))) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE)
```

Growth Curve for Men Only

Instead of `gls()`, we use the `lme()` in the `nlme` package, but now we have occasion nested within person and no need for correlated errors. Now we are using traditional MLM and will make use of the random option instead of the correlation option in the syntax below. The random option asks `lme()` to estimate variance in the intercepts of satisfaction at `TIME` = 0 (the study midpoint), and variance in the slopes (change in satisfaction overtime). The subject is `DYADID` because this variable is serving as our participant ID—recall that we are going to use men only and this is a heterosexual sample.

```
GC_men <- lme(ASATISF ~ TIME,
          data = kashy_men,
          random = ~ 1 + TIME|DYADID,
          na.action = na.omit)

summary(GC_men)

## Linear mixed-effects model fit by REML
##  Data: kashy_men
##        AIC       BIC    logLik
##   2913.754 2945.342 -1450.877
##
```

```
## Random effects:
##  Formula: ~1 + TIME | DYADID
##  Structure: General positive-definite, Log-Cholesky parametrization
##            StdDev     Corr
## (Intercept) 0.68969184 (Intr)
## TIME        0.04991018 -0.055
## Residual    0.57494853
##
## Fixed effects: ASATISF ~ TIME
##                 Value  Std.Error   DF t-value p-value
## (Intercept) 6.259438 0.06964341 1327 89.8784  0.0000
## TIME        0.019287 0.00621176 1327  3.1049  0.0019
##  Correlation:
##      (Intr)
## TIME -0.043
##
## Standardized Within-Group Residuals:
##        Min          Q1         Med          Q3         Max
## -6.6635944 -0.2279288  0.0955475  0.3916416  4.6840918
##
## Number of Observations: 1431
## Number of Groups: 103
```

We can use the `confintLME()` function in the `dyadr` package to get confidence intervals for the fixed effects.

```
coef(summary(GC_men))
```

```
##                   Value    Std.Error   DF    t-value      p-value
## (Intercept) 6.25943802 0.069643407 1327 89.878400 0.000000000
## TIME        0.01928692 0.006211763 1327  3.104902 0.001943707
```

```
confintLME(GC_men)
```

```
##                    2.5%      97.5%
## (Intercept) 6.122814832 6.39606120
## TIME        0.007100969 0.03147286
```

The final fixed equation is:

$$\widehat{Satisfaction} = 6.26 + .019(TIME)$$

The Intercept $= 6.26$, which is interpreted as the average level of satisfaction at TIME $= 0$ (the study midpoint). The coefficient for TIME $= .019$. That is, over time, satisfaction increases .019 units a day. The slope is small, although it is statistically significant—thus, there is some evidence of an average increase in satisfaction over time for men.

As for the random effects, the three important estimates are:

1. Standard deviation of Intercepts: `(Intercept)` = .670
2. Standard deviation of Slopes: `TIME` = .050
3. Correlation between Int/Slope: `Corr` = -.055

There is positive variance in the intercepts—some men were more satisfied than others at the midpoint. There is also positive variance in the slopes—some men are changing in satisfaction more than others. The slope-intercept covariance is in the negative direction. We would interpret this as, men with higher values at time 0 change more slowly than those with lower values.

Adding a Moderator

Next, we can include the men's attachment avoidance as a moderator of the level of satisfaction at the study midpoint (main effect of avoidance), and as a moderator of the change in satisfaction over time (TIME by avoidance interaction). This is the addition of two fixed effects—the random effects remain unchanged. The interpretation of the random effects are as before, but they are interpreted with the effect of avoidance "partialled out." `CAAvoid` is grand mean centered avoidance of the actor.

```
GC_men_mod <- lme(ASATISF ~ TIME + CAAvoid + TIME*CAAvoid,
               data = kashy_men,
               random = ~ 1 + TIME|DYADID,
               na.action = na.omit)

smallsummary(GC_men_mod)
```

```
## Random effects:
##  Formula: ~1 + TIME | DYADID
##  Structure: General positive-definite, Log-Cholesky parametrization
##              StdDev     Corr
## (Intercept) 0.67848447 (Intr)
## TIME        0.05026826 -0.032
## Residual    0.57764139
##
##                Value Std.Error   DF t-value p-value
## (Intercept)   6.2610    0.0690 1313 90.6848  0.0000
## TIME          0.0193    0.0063 1313  3.0616  0.0022
## CAAvoid      -0.1398    0.0688  100 -2.0321  0.0448
## TIME:CAAvoid  0.0045    0.0063 1313  0.7193  0.4721
##
##               2.5%    97.5%
## (Intercept)  6.1255  6.3964
## TIME         0.0069  0.0316
## CAAvoid     -0.2762 -0.0033
```

```
## TIME:CAAvoid -0.0078   0.0168
```

The final fixed effects growth equation with attachment avoidance as a moderator is:

$$\widehat{Satisfaction} = 6.26 + .019(TIME) - .140(Avoid) + .0045(TIME * Avoid)$$

The $intercept = 6.26$, the predicted satisfaction for men at the study midpoint who are at the mean on avoidance. There is a significant main effect of time, $b = 0.019$, that is, satisfaction increases by .019 each day for men who are at the mean on avoidance. There is a significant main effect of avoidance, $b = -.140$, such that men who are higher in avoidance are less satisfied at the study midpoint than men who are lower in avoidance. There is no significant Time X Avoidance interaction, $b = .0045$, but to interpret the direction of this interaction we could calculate the slope for men who are high and low in avoidance:

```
kashy_men %>%
  summarize(sd(CAAvoid, na.rm = TRUE))
```

```
##   sd(CAAvoid, na.rm = TRUE)
## 1                 1.002279
```

$Men's Avoidance SD = 1.002$

```
High_Avoid_Slope <- .019+.0045*1.002
Low_Avoid_Slope <- .019-.0045*1.002

High_Avoid_Slope
```

```
## [1] 0.023509
```

```
Low_Avoid_Slope
```

```
## [1] 0.014491
```

$High Avoidance (+1sd) Slope = .019 + .0045 * 1.002 = .0235$
$Low Avoidance (-1sd) Slope = .019 - .0045 * 1.002 = .0145$

Alternatively, we could re-center Avoidance to obtain these two values. Recall that the re-centering method will give you significance tests of the simple slopes and then you can use the `graphMod()` function to make a figure. We interpret this interaction as: men higher on avoidance become more satisfied over the course of the study than men lower on avoidance, but this difference is not significant.
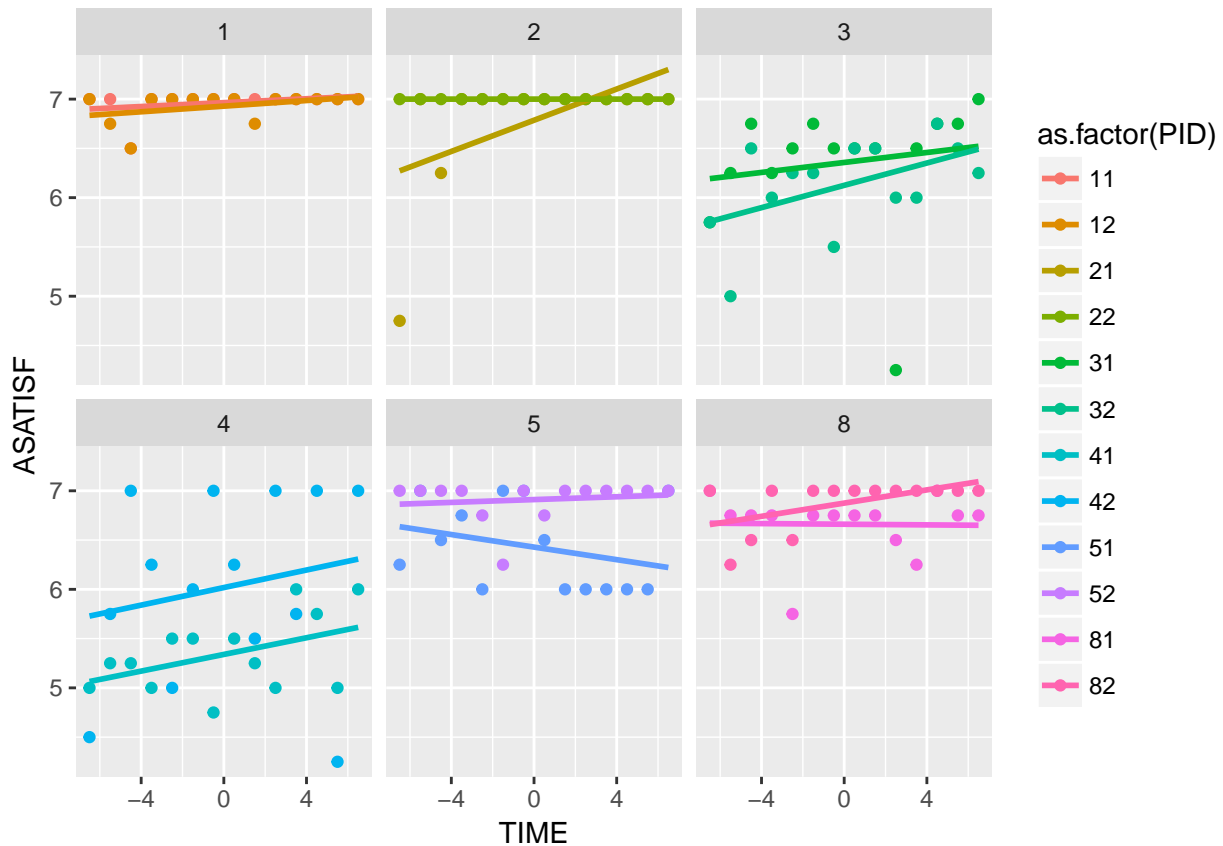
Dyadic Growth Curve Modeling

Dyadic Spaghetti Plot

Dyadic spaghetti plot shows different slopes for different folks (across different dyads)!

```
kashy_small <- kashy_ppp %>%
  filter(DYADID >= 1 & DYADID <= 8)

ggplot(kashy_small, aes(TIME, ASATISF,
                        group = as.factor(PID),
                        color = as.factor(PID))) +
  geom_point() +
  geom_smooth(method = "lm", se = FALSE) +
  facet_wrap(~as.factor(DYADID))
```


Dyadic Growth Curve for Distinguishable Dyads

Create obsid Index Variable

```
kashy_ppp <- kashy_ppp %>%
  mutate(obsid = Day+14*(DYADID-1))
```

Two-Intercept Approach

For dyadic growth curve modeling we are going to start with a two intercept model. Note how `GenderS` is included below as well as `-1`. This will give us separate intercepts for women and men. We again use the `lme()` procedure, but now we need a `random =` statement as well as a `correlation =` statement.

The `random =` statement estimates separate intercepts and slopes for men and women and all of the within-and-between person correlations (see slides for more description of all of these random effects).

The `correlation =` Statement specifies the variances and covariances between the residuals.

`weights = varIdent(form = ~1|GenderS)` asks R to estimate different residual variances for men and women (set to be the same at each time point). There is one correlation between the residuals (i.e., the time-specific correlation between satisfaction scores—will be `Rho` in the output).

```
dyadGC_di_two  <- lme(ASATISF ~ GenderS + GenderS:TIME  - 1,
                      data = kashy_ppp,
                      random = ~ GenderS + GenderS:TIME  - 1|DYADID,
                      correlation = corCompSymm(form = ~1|DYADID/obsid),
                      weights = varIdent(form = ~1|GenderS),
                      na.action = na.omit)

smallsummary(dyadGC_di_two)
```

```
## Random effects:
##  Formula: ~GenderS + GenderS:TIME - 1 | DYADID
##  Structure: General positive-definite, Log-Cholesky parametrization
##                    StdDev     Corr
## GenderSMan         0.68939425 GndrSM GndrSW GSM:TI
## GenderSWoman       0.55533836  0.811
## GenderSMan:TIME    0.04997165 -0.058  0.053
## GenderSWoman:TIME  0.04786064 -0.017  0.149  0.462
## Residual           0.57490536
##
##                   Value Std.Error   DF  t-value p-value
## GenderSMan        6.2591    0.0696 2760  89.9146   0.000
## GenderSWoman      6.3884    0.0572 2760 111.6052   0.000
## GenderSMan:TIME   0.0192    0.0062 2760   3.0876   0.002
## GenderSWoman:TIME 0.0099    0.0063 2760   1.5639   0.118

##                    2.5%   97.5%
## GenderSMan        6.1226 6.3956
## GenderSWoman      6.2762 6.5006
## GenderSMan:TIME   0.0070 0.0314
```

```
## GenderSWoman:TIME -0.0025 0.0222
```

We can also use the `lincomb()` function to test for differences between our estimates for men and women.

```
lincomb(summary(dyadGC_di_two), 1, 2)
```

```
## [1] -0.129261027  0.043308553  0.002838996
```

```
lincomb(summary(dyadGC_di_two), 3, 4)
```

```
## [1] 0.009329482 0.006577413 0.156070506
```

Fixed Effects:

- `GENDER`: Is the intercept different for men and women? Yes, p = .003.
  - Male Intercept = 6.259
  - Female Intercept = 6.388
- `TIME`: Do men and women have different slopes? No, men have steeper, more positive slopes than women, but not significantly so, p = .156
  - Male slope = .0191
  - Female slope = .0098

Interaction Approach

Finally, as a reminder, one should include gender in the model as a moderator. We've been referring to as this as the "interaction approach." The dyadic growth curve model for the gender moderation model is as follows.

```
dyadGC_di_int  <- lme(ASATISF ~ GENDER + TIME + GENDER*TIME,
                  data = kashy_ppp,
                  random = ~ GenderS + GenderS:TIME  - 1|DYADID,
                  correlation = corCompSymm(form = ~1|DYADID/obsid),
                  weights = varIdent(form = ~1|GenderS),
                  na.action = na.omit)
```

```
smallsummary(dyadGC_di_int)
```

```
## Random effects:
##  Formula: ~GenderS + GenderS:TIME - 1 | DYADID
##  Structure: General positive-definite, Log-Cholesky parametrization
##                    StdDev     Corr
## GenderSMan         0.68939513 GndrSM GndrSW GSM:TI
## GenderSWoman       0.55533714  0.811
## GenderSMan:TIME    0.04997161 -0.058  0.053
## GenderSWoman:TIME  0.04786057 -0.017  0.149  0.462
## Residual           0.57490548
##
```

```
##                 Value Std.Error    DF  t-value  p-value
## (Intercept)  6.3238     0.0599  2760 105.5093   0.0000
## GENDER      -0.0646     0.0217  2760  -2.9847   0.0029
## TIME         0.0145     0.0053  2760   2.7273   0.0064
## GENDER:TIME  0.0047     0.0033  2760   1.4184   0.1562

##                2.5%    97.5%
## (Intercept)  6.2063   6.4413
## GENDER      -0.1071  -0.0222
## TIME         0.0041   0.0250
## GENDER:TIME -0.0018   0.0111
```

Adding a Moderator

```
dyadGC_di_mod  <- lme(ASATISF ~ GENDER*TIME*CAAvoid + GENDER*TIME*CPAvoid,
                  data = kashy_ppp,
                  random = ~ GenderS + GenderS:TIME  - 1|DYADID,
                  correlation = corCompSymm(form = ~1|DYADID/obsid),
                  weights = varIdent(form = ~1|GenderS),
                  na.action = na.omit)

smallsummary(dyadGC_di_mod)
```

```
## Random effects:
##  Formula: ~GenderS + GenderS:TIME - 1 | DYADID
##  Structure: General positive-definite, Log-Cholesky parametrization
##                   StdDev      Corr
## GenderSMan        0.68146700 GndrSM GndrSW GSM:TI
## GenderSWoman      0.55815148  0.807
## GenderSMan:TIME   0.05028625 -0.032  0.075
## GenderSWoman:TIME 0.04736106  0.040  0.189  0.466
## Residual          0.57762304
##
##                   Value Std.Error    DF  t-value p-value
## (Intercept)      6.3236    0.0602  2725 105.0102  0.0000
## GENDER          -0.0643    0.0218  2725  -2.9520  0.0032
## TIME             0.0144    0.0054  2725   2.6798  0.0074
## CAAvoid         -0.0866    0.0480  2725  -1.8049  0.0712
## CPAvoid         -0.0456    0.0490  2725  -0.9310  0.3520
## GENDER:TIME      0.0052    0.0033  2725   1.5719  0.1161
## GENDER:CAAvoid  -0.0537    0.0469  2725  -1.1462  0.2518
## TIME:CAAvoid     0.0008    0.0048  2725   0.1766  0.8598
## GENDER:CPAvoid   0.0268    0.0479  2725   0.5602  0.5754
## TIME:CPAvoid     0.0097    0.0048  2725   2.0432  0.0411
```

```
## GENDER:TIME:CAAvoid  0.0038    0.0047 2725   0.8058  0.4204
## GENDER:TIME:CPAvoid -0.0023    0.0047 2725  -0.4934  0.6217

##                           2.5%    97.5%
## (Intercept)            6.2056   6.4417
## GENDER                -0.1071  -0.0216
## TIME                   0.0039   0.0250
## CAAvoid               -0.1806   0.0075
## CPAvoid               -0.1416   0.0504
## GENDER:TIME           -0.0013   0.0117
## GENDER:CAAvoid        -0.1456   0.0382
## TIME:CAAvoid          -0.0085   0.0102
## GENDER:CPAvoid        -0.0671   0.1207
## TIME:CPAvoid           0.0004   0.0191
## GENDER:TIME:CAAvoid   -0.0054   0.0130
## GENDER:TIME:CPAvoid   -0.0115   0.0069
```

We see that there is no statistically significant difference in the `TIME:AVOID` effects by gender, no three-way interactions. Further, gender does not moderate the actor and partner effects.

So we can run a model where gender is only included as a main effect—i.e., gender affects the intercept only.

```
dyadGC_di_mod2  <- lme(ASATISF ~ GENDER + TIME*CAAvoid + TIME*CPAvoid,
                    data = kashy_ppp,
                    random = ~ GenderS + GenderS:TIME  - 1|DYADID,
                    correlation = corCompSymm(form = ~1|DYADID/obsid),
                    weights = varIdent(form = ~1|GenderS),
                    na.action = na.omit)

coef(summary(dyadGC_di_mod2))
```

```
##                    Value    Std.Error   DF     t-value      p-value
## (Intercept)   6.324579898 0.059767886 2730 105.8190323 0.000000000
## GENDER       -0.065304795 0.021752164 2730  -3.0022206 0.002704476
## TIME          0.014854749 0.005345869 2730   2.7787340 0.005494459
## CAAvoid      -0.075081635 0.046485535 2730  -1.6151613 0.106391330
## CPAvoid      -0.043491130 0.045562331 2730  -0.9545414 0.339894185
## TIME:CAAvoid  0.001003538 0.004669892 2730   0.2148953 0.829865020
## TIME:CPAvoid  0.009312222 0.004677778 2730   1.9907362 0.046609438
```

From this model we can see that there is a statistically significant interaction of `TIME` and `CPAvoid`, which means that the change over time in satisfaction gets larger as one's partner's avoidance goes up.

The next thing we would want to do is test the simple slopes (of `TIME`) at high and low partner avoidance. We probably also want a figure.

Graphing an Interaction

First we will run the model again with:

1. high partner avoidance, and
2. low partner avoidance.

```
kashy_ppp <- kashy_ppp %>%
  mutate(High_CPAvoid = CPAvoid - sd(CPAvoid, na.rm = TRUE),
         Low_CPAvoid = CPAvoid + sd(CPAvoid, na.rm = TRUE))



dyadGC_di_mod2_High  <- lme(ASATISF ~ GENDER + TIME*CAAvoid + TIME*High_CPAvoid,
                     data = kashy_ppp,
                     random = ~ GenderS + GenderS:TIME  - 1|DYADID,
                     correlation = corCompSymm(form = ~1|DYADID/obsid),
                     weights = varIdent(form = ~1|GenderS),
                     na.action = na.omit)

dyadGC_di_mod2_Low  <- lme(ASATISF ~ GENDER + TIME*CAAvoid + TIME*Low_CPAvoid,
                     data = kashy_ppp,
                     random = ~ GenderS + GenderS:TIME  - 1|DYADID,
                     correlation = corCompSymm(form = ~1|DYADID/obsid),
                     weights = varIdent(form = ~1|GenderS),
                     na.action = na.omit)
```

The first model tests the simple time slope for high partner avoidance:

```
coef(summary(dyadGC_di_mod2_High))
```

```
##                         Value     Std.Error   DF   t-value       p-value
## (Intercept)         6.283267770 0.073964685 2730 84.949564 0.0000000000
## GENDER             -0.065304794 0.021752164 2730 -3.002221 0.0027044767
## TIME                0.023700407 0.006968113 2730  3.401266 0.0006803931
## CAAvoid            -0.075081639 0.046485502 2730 -1.615162 0.1063910683
## High_CPAvoid       -0.043491129 0.045562298 2730 -0.954542 0.3398938495
## TIME:CAAvoid        0.001003536 0.004669892 2730  0.214895 0.8298652604
## TIME:High_CPAvoid   0.009312224 0.004677778 2730  1.990737 0.0466093699
```

and the second model tests the simple time slope for low partner avoidance:

```
coef(summary(dyadGC_di_mod2_Low))
```

```
##                        Value    Std.Error   DF    t-value      p-value
## (Intercept)         6.365892051 0.073619636 2730 86.4700286 0.000000000
## GENDER             -0.065304794 0.021752164 2730 -3.0022206 0.002704476
## TIME                0.006009079 0.006934685 2730  0.8665252 0.386278380
## CAAvoid            -0.075081642 0.046485498 2730 -1.6151627 0.106391025
```

```
## Low_CPAvoid       -0.043491130 0.045562294 2730 -0.9545421 0.339893797
## TIME:CAAvoid       0.001003536 0.004669892 2730  0.2148949 0.829865294
## TIME:Low_CPAvoid   0.009312225 0.004677778 2730  1.9907368 0.046609366
```

Looking at the effect of `TIME` in these two models. We see that the change over time is statistically significant when partner's avoidance is high, $p = .0007$, but it is not significant when partner's avoidance is low, $p = .386$.

We can then use these two models in the `graphMod()` function to make an interaction figure. We will need to note that the intercept is in position **1**, and the slope of interest, `TIME`, is in position **3**.

```
plot <- graphMod(kashy_ppp,
                 kashy_ppp$TIME,
                 kashy_ppp$ASATISF,
                 kashy_ppp$CPAvoid,
                 dyadGC_di_mod2_High,
                 dyadGC_di_mod2_Low,
                 1, 3)

plot +
  geom_point(color = "white") + #to supress dots
  theme_classic() +
  ylab("Satisfaction") +
  xlab("Time (0 is study midpoint)") +
  scale_color_manual(values = c("tomato", "springgreen"),
                     name = "Partner's Avoidance") +
  ylim(5.5, 6.5)
```
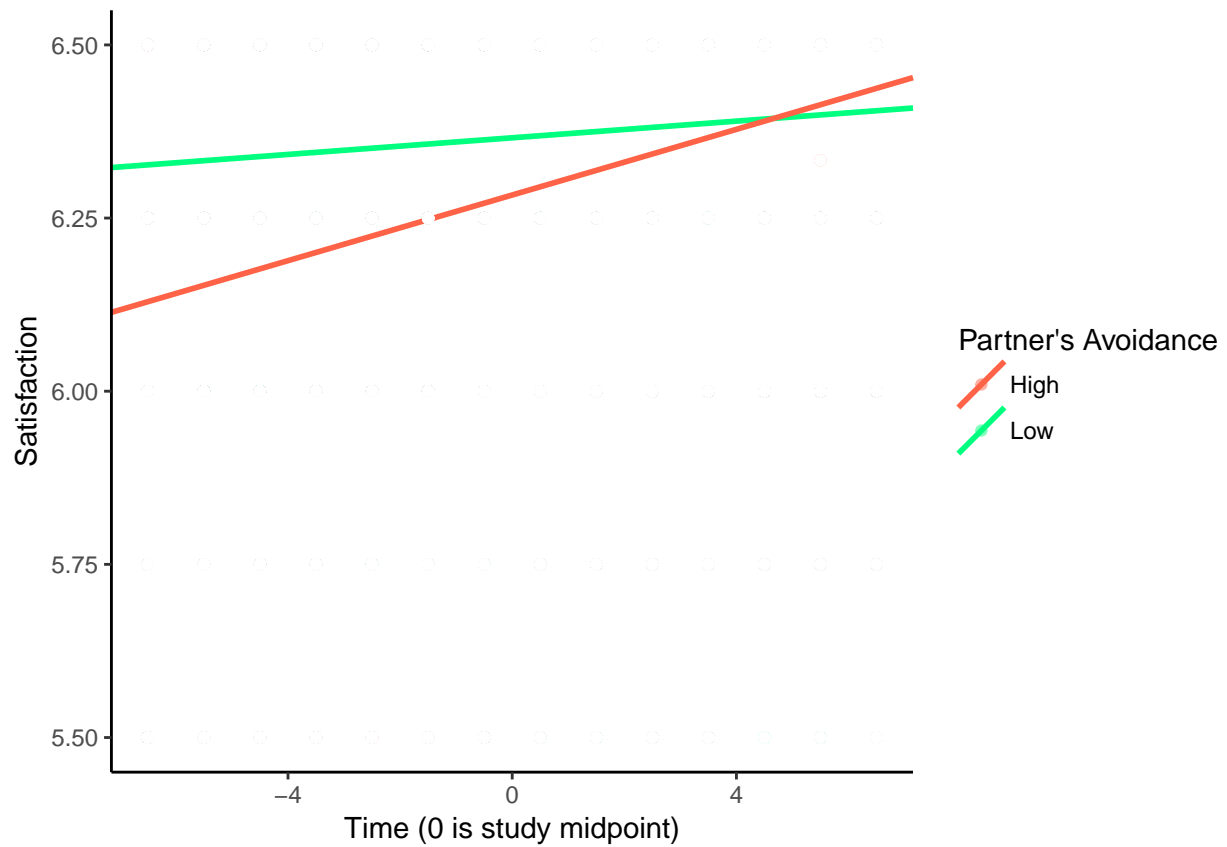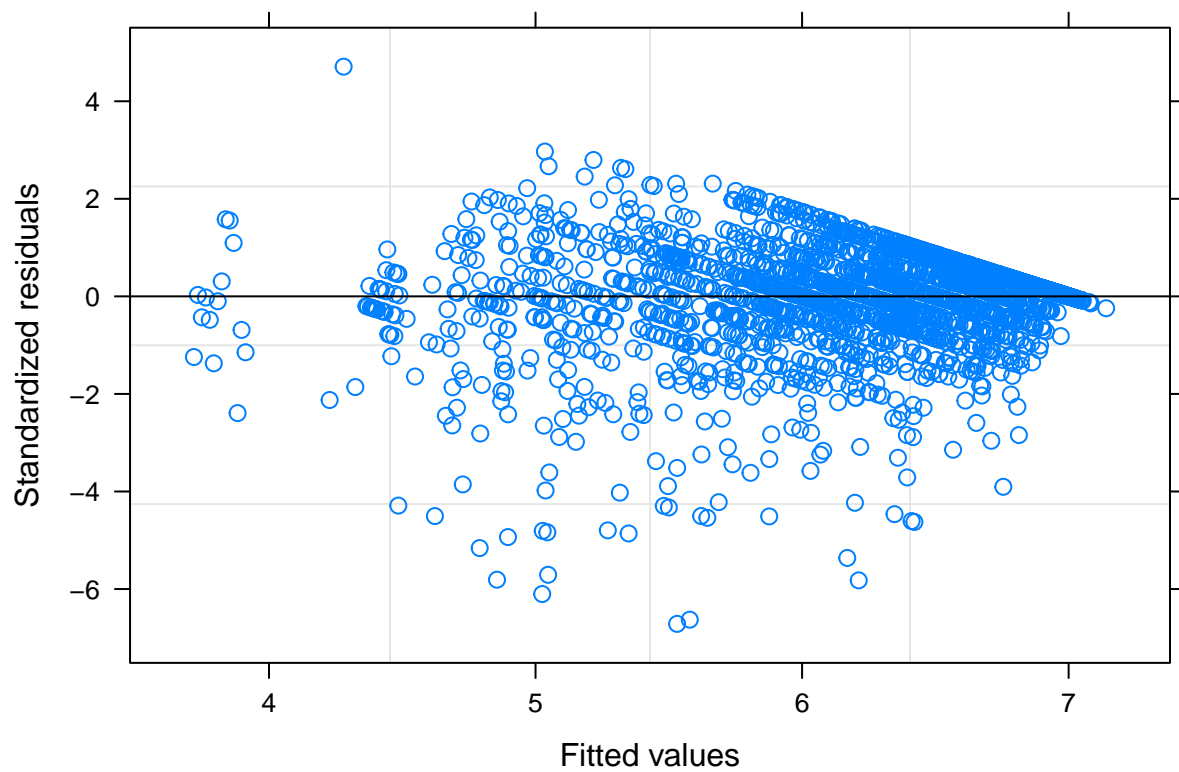
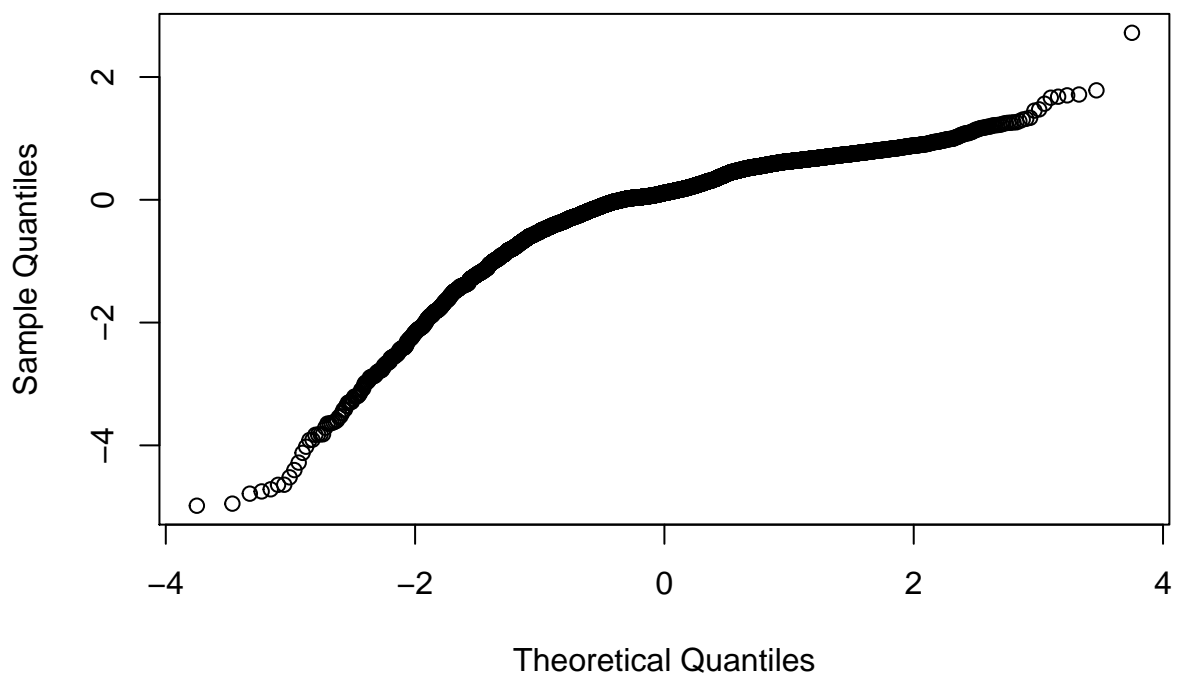This looks like a ceiling effect to me. Maybe satisfaction is skewed to the left?

Model Diagnostics

```
plot(dyadGC_di_mod2)
```

```
qqnorm(dyadGC_di_mod2$residuals)
```

Normal Q–Q Plot

It turns out, in this example, satisfaction is skewed to the left.

```r
qplot(x = ASATISF, data = kashy_ppp, bins = 20)
```