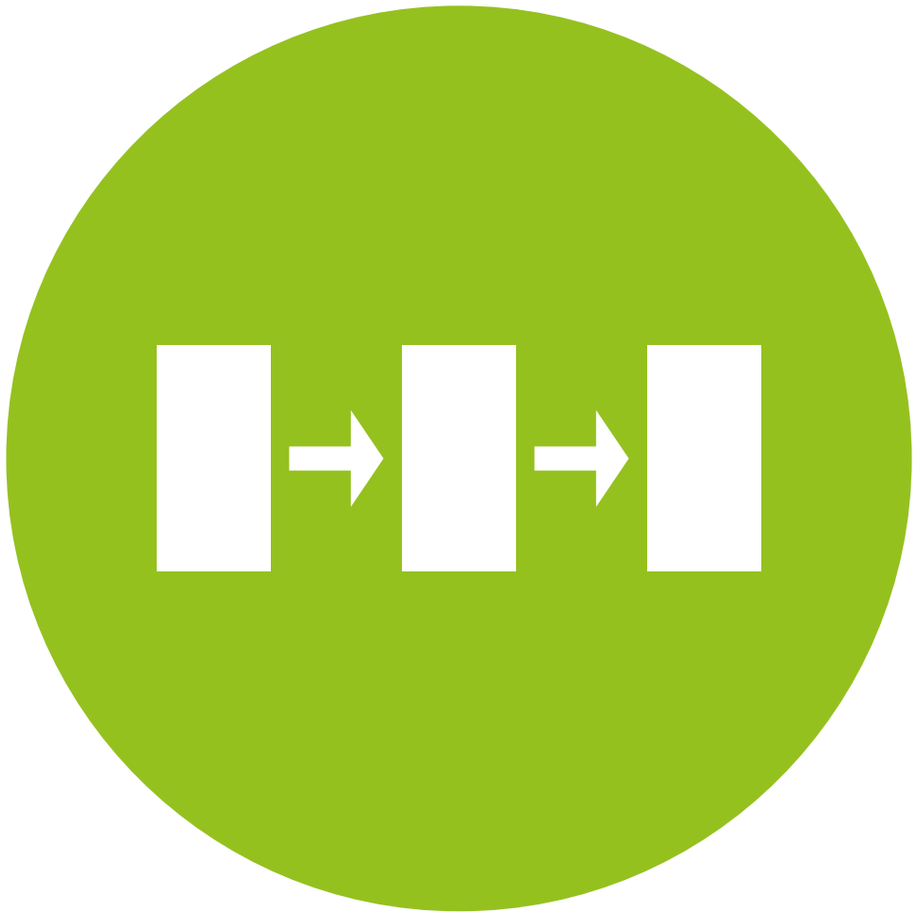


SEM part II

Structural Equation Modeling

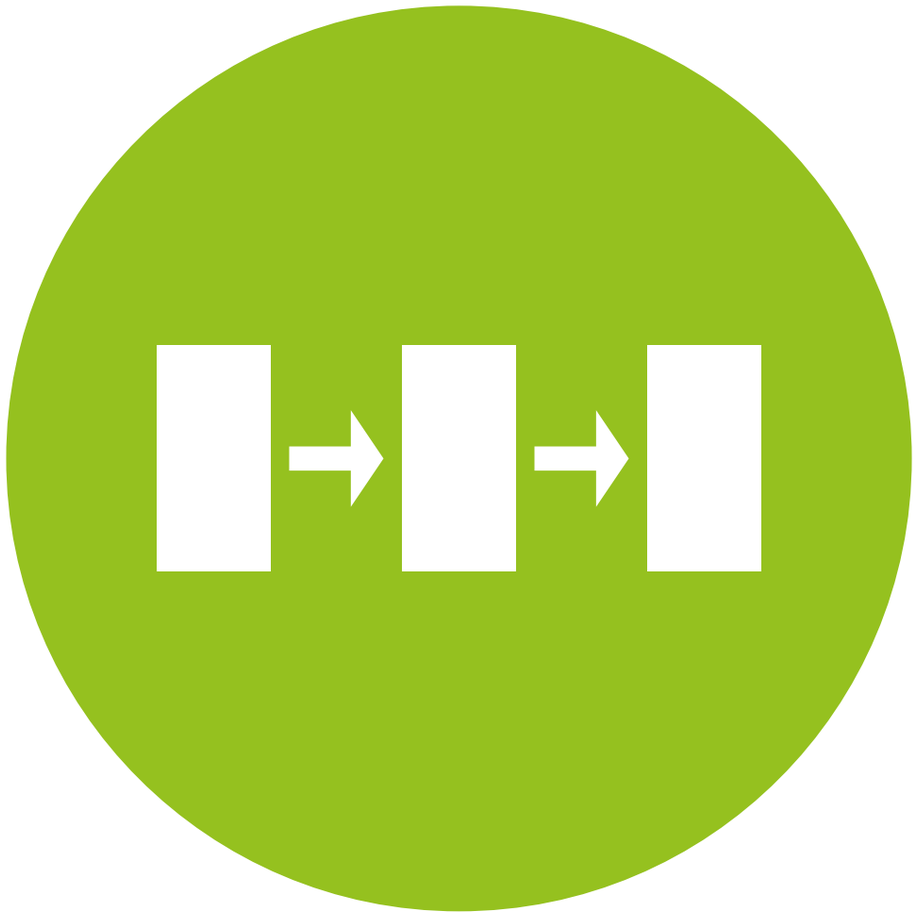
Intro

Today's goal:

Teach how to test theoretical models with SEM.

Outline:

- Running, trimming, and evaluating saturated models
- Expanding the model to include additional variables



Testing full models

The real power of SEM

Where we are now

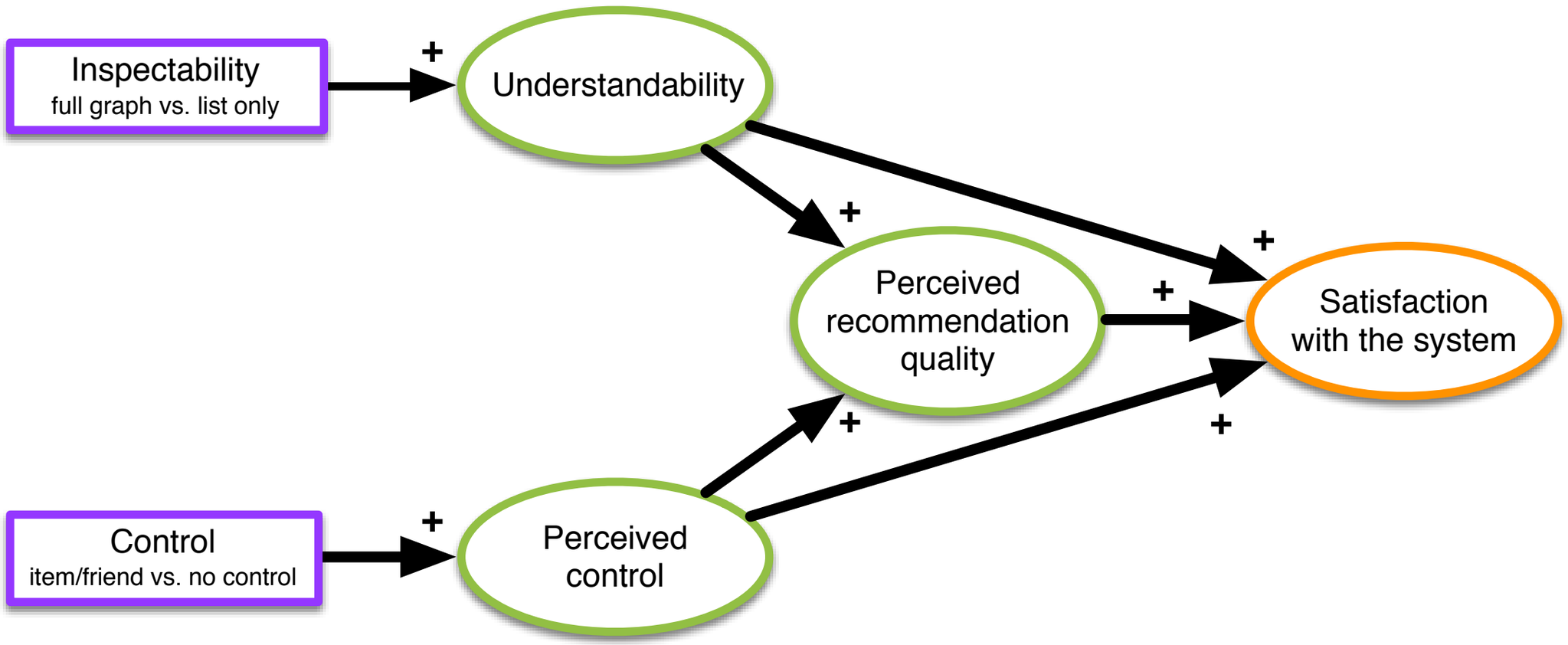
Step 1: Build your CFA ✓

Step 2: Analyze the marginal effects of the manipulations ✓

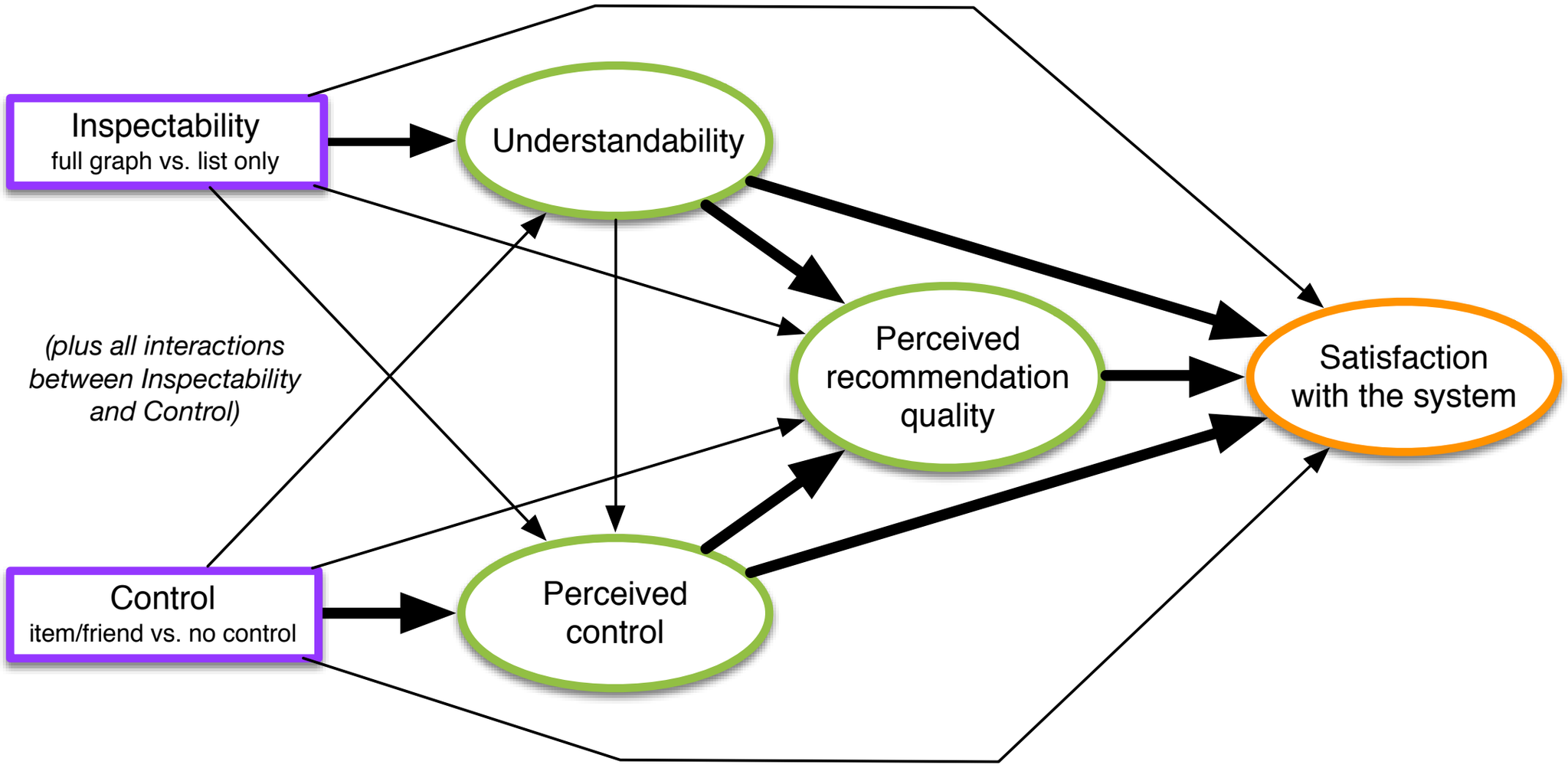
Step 3: Set up a model based on theory ✓

Step 4: Test and trim a saturated version of this model

Theoretical model



Saturated model



4. Test the model

Steps:

- Trim the model
- Get model fit statistics
- Optional: expand the model
- Reporting

Run model

In R:

```
model <- 'satisf =~ s1+s2+s3+s4+s5+s6+s7
  quality =~ q1+q2+q3+q4+q5+q6
  control =~ c1+c2+c3+c4
  underst =~ u2+u4+u5
  satisf ~ quality+control+underst+citem+cfriend+cgraph+cig+cfg
  quality ~ control+underst+citem+cfriend+cgraph+cig+cfg
  control ~ underst+citem+cfriend+cgraph+cig+cfg
  underst ~ citem+cfriend+cgraph+cig+cfg';

fit <- sem(model,data=twq,ordered=names(twq[9:31]),std.lv=TRUE);

summary(fit);
```

Trim model

Rules:

- Start with the least significant and least interesting effects (those that were added for saturation)
- Work iteratively
- Manipulations with >2 conditions: remove all dummies at once (if one is significant, keep the others as well)
- Interaction+main effects: never remove main effect before the interaction effect (if the interaction is significant, keep the main effect regardless)

Results

	Estimate	Std.err	Z-value	P(>|z|)
...(factors)...	...	...	...	...
Regressions:				
satisf ~				
quality	0.439	0.076	5.753	0.000
control	-0.838	0.107	-7.804	0.000
underst	0.090	0.073	1.229	0.219
citem	0.318	0.265	1.198	0.231
cfriend	0.014	0.257	0.054	0.957
cgraph	0.308	0.229	1.346	0.178
cig	-0.386	0.356	-1.082	0.279
cfg	-0.394	0.357	-1.103	0.270
quality ~				
control	-0.764	0.086	-8.899	0.000
underst	0.044	0.073	0.595	0.552
citem	0.046	0.204	0.226	0.821
cfriend	0.165	0.251	0.659	0.510
cgraph	0.009	0.236	0.038	0.970
cig	0.106	0.317	0.334	0.738
cfg	0.179	0.374	0.478	0.632

Results

control ~				
understanding	−0.308	0.066	−4.695	0.000
citem	0.053	0.240	0.220	0.826
cfriend	0.009	0.221	0.038	0.969
cgraph	−0.043	0.239	−0.181	0.857
cig	−0.148	0.341	−0.434	0.664
cfg	−0.273	0.331	−0.824	0.410
understanding ~				
citem	0.367	0.220	1.666	0.096
cfriend	0.534	0.217	2.465	0.014
cgraph	0.556	0.227	2.451	0.014
cig	−0.106	0.326	−0.324	0.746
cfg	−0.178	0.320	−0.555	0.579

Trimming steps

Remove interactions -> (1) understandability, (2) quality, (3) control, and (4) satisfaction

Remove cgraph -> (1) satisfaction, and (2) quality

Trimming steps

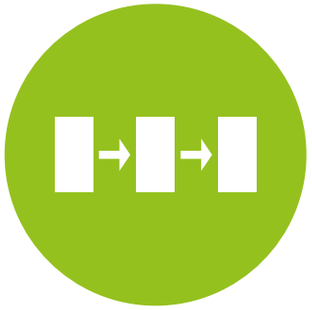
Remove citem and cfriend -> control

But wait… did we not hypothesize that effect?

Yes, but we still have citem+cfriend -> underst -> control!

In other words: the effect of item and friend control on perceived control is mediated by understandability!

Argument: “Controlling items/friends gives me a better understanding of how the system works, so in turn I feel more in control”



Trimming steps

Remove citem and cfriend -> satisfaction

Remove understandability -> recommendation quality

We hypothesized this effect, but it is still mediated by control.

Argument: “Understanding the recommendations gives me a feeling of control, which in turn makes me like the recommendations better.”

Remove understandability -> satisfaction

Same thing

Trimming steps

Remove citem and cfriend -> recommendation quality

Remove cgraph -> control

Again: still mediated by understandability

Stop! All remaining effects are significant!

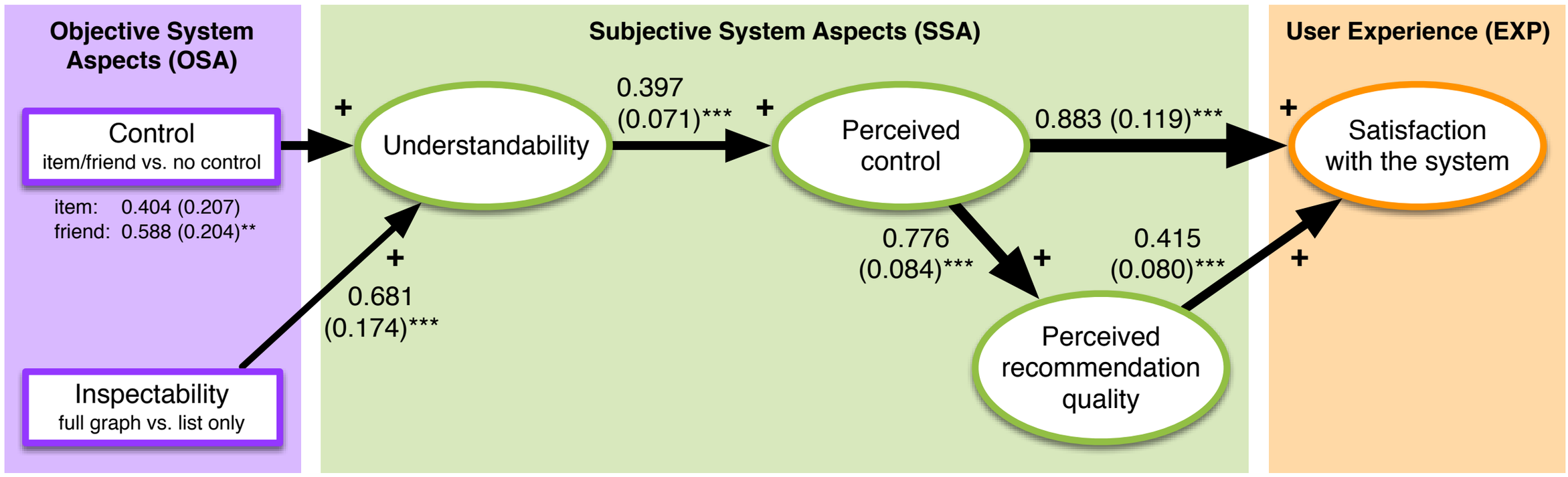
Trimmed model

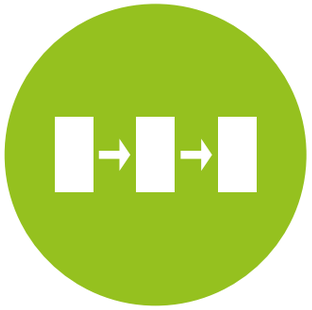
```
model <- 'satisf =~ s1+s2+s3+s4+s5+s6+s7
quality =~ q1+q2+q3+q4+q5+q6
control =~ c1+c2+c3+c4
underst =~ u2+u4+u5
satisf ~ quality+control
quality ~ control
control ~ underst
underst ~ citem+cfriend+cgraph'
```

Trimmed model

	Estimate	Std.err	Z-value	P(>|z|)
...(factors)...	...	...	...	...
Regressions:				
satisf ~				
quality	0.418	0.080	5.228	0.000
control	-0.887	0.120	-7.395	0.000
quality ~				
control	-0.779	0.084	-9.232	0.000
control ~				
underst	-0.371	0.067	-5.522	0.000
underst ~				
citem	0.382	0.200	1.915	0.056
cfriend	0.559	0.195	2.861	0.004
cgraph	0.628	0.166	3.786	0.000

Trimmed model





Modindices

Get modification indices

```
mods <- modindices(fit)
mods <- mods[mods$mi > 3.84 & !is.na(mods$mi),]
mods[order(-mods$mi),]
```

```
      lhs op     rhs     mi mi.scaled    epc sepc.lv sepc.all sepc.nox
1 control  ~  satisf 28.794    23.984  0.578   0.912    0.912    0.912
2 control ~~ underst 22.270    18.550  0.352   0.304    0.304    0.304
3      s3 ~~      s4 20.785    17.313  0.157   0.157    0.156    0.156
4 underst =~      q5 15.201    12.662 -0.150  -0.162   -0.161   -0.161
5      s2 ~~      s7 10.021     8.347  0.101   0.101    0.100    0.100
6  satisf =~      c3  8.796     7.327 -0.169  -0.286   -0.284   -0.284
7 underst =~      s6  8.049     6.705  0.109   0.117    0.117    0.117
```

No useful modification indices in the regression part of the model (only stuff we had left from the CFA)

Assess model fit

Item and factor fit should not have changed much

(please double-check the r-squares!)

Great model fit!

- Chi-Square value: 306.685, df: 223 (value/df = 1.38)
- CFI: 0.994, TLI: 0.993
- RMSEA: 0.037 (great), 90% CI: [0.026, 0.047]

Regression R^2

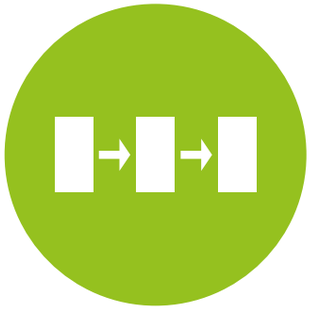
Satisfaction: 0.654

Perceived Recommendation Quality: 0.416

Perceived Control: 0.156

Understandability: 0.151

These are all quite okay



Omnibus test

In model definition:

```
underst ~ cgraph+p1*citem+p2*cfriend
```

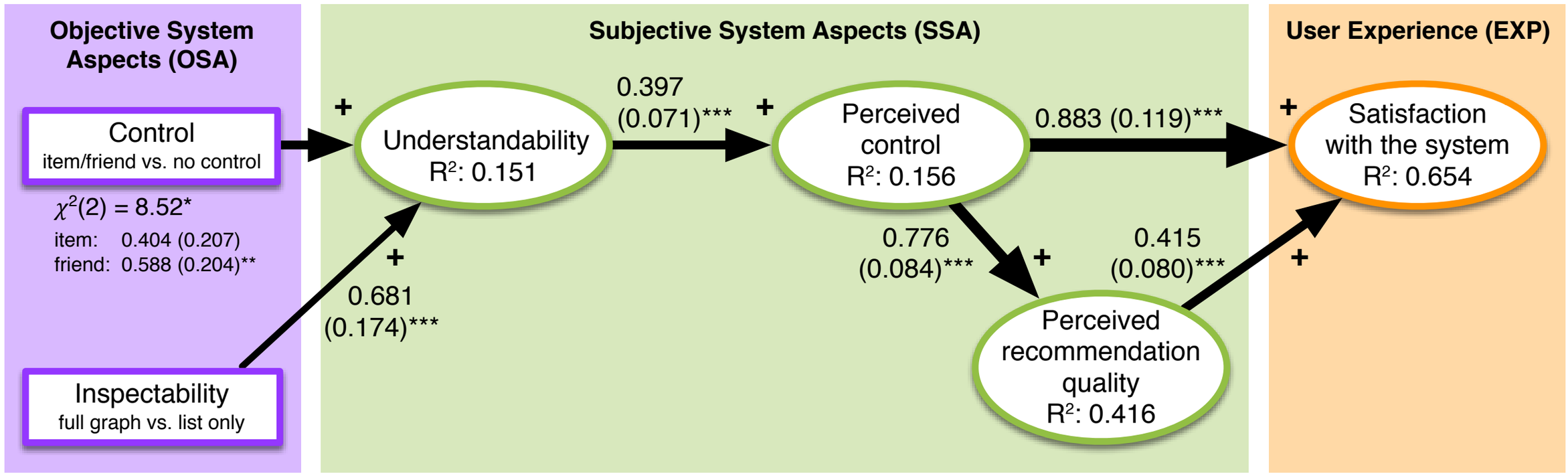
Then run:

```
lavTestWald(fit,'p1==0;p2==0');
```

Result: Omnibus effect of control is significant (this is a chi-square test)

```
$stat
[1] 8.386272

$df
[1] 2

$p.value
[1] 0.01509886
```

Final core model



Reporting

We subjected the 4 factors and the experimental conditions to structural equation modeling, which simultaneously fits the factor measurement model and the structural relations between factors and other variables. The model has a good* model fit: chi-square(223) = 306.685, p = .0002; RMSEA = 0.037, 90% CI: [0.026, 0.047], CFI = 0.994, TLI = 0.993.

* A model should not have a non-significant chi-square (p > .05), but this statistic is often regarded as too sensitive. Hu and Bentler propose cut-off values for other fit indices to be: CFI > .96, TLI > .95, and RMSEA < .05, with the upper bound of its 90% CI below 0.10.

Reporting

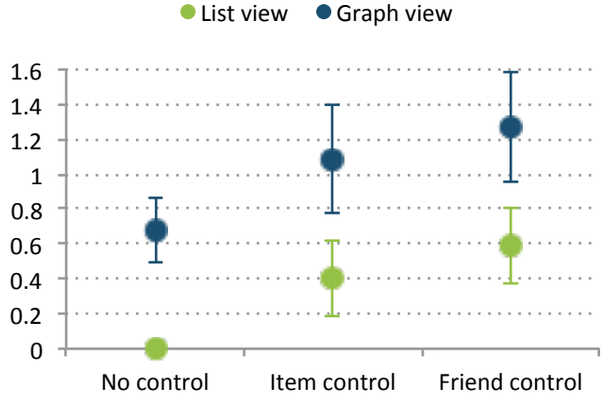
The model shows that the inspectability and control manipulations each have an independent positive effect on the understandability of the system: the full graph condition is more understandable than the list only condition, and the item control and friend control conditions are more understandable than the no control condition. Understandability is in turn related to users' perception of control, which is in turn related to the perceived quality of the recommendations. The perceived control and the perceived recommendation quality finally determine participants' satisfaction with the system.

Total effects

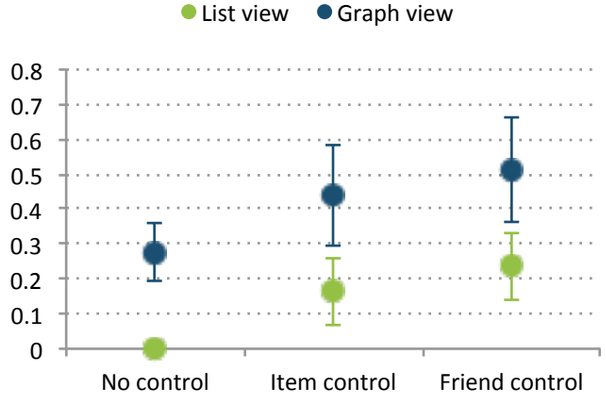
```
model <- 'satisf =~ s1+s2+s3+s4+s5+s6+s7
            quality =~ q1+q2+q3+q4+q5+q6
            control =~ c1+c2+c3+c4
            underst =~ u2+u4+u5
            satisf ~ pqs*quality+pcs*control
            quality ~ pcq*control
            control ~ puc*underst
            underst ~ piu*citem+pfu*cfriend+pgu*cgraph

item2sat := piu*puc*(pcs+pcq*pqs)
friend2sat := pfu*puc*(pcs+pcq*pqs)
graph2sat := pgu*puc*(pcs+pcq*pqs)

item2qual := piu*puc*pcq
friend2qual := pfu*puc*pcq
graph2qual := pgu*puc*pcq

item2ctrl := piu*puc
friend2ctrl := pfu*puc
graph2ctrl := pgu*puc’
```

Total effect graphs

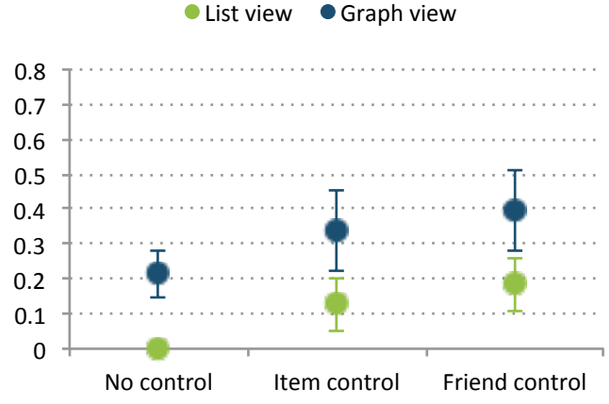
Understandability



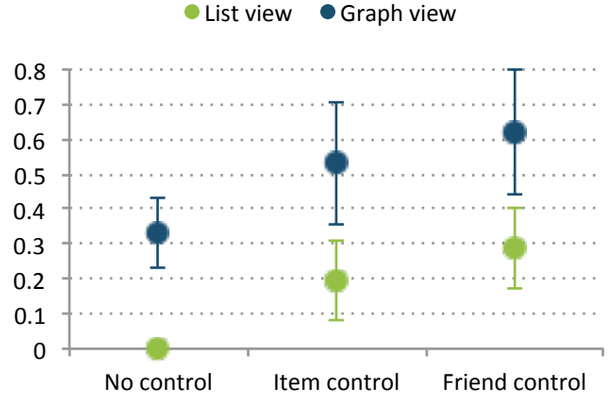
Perceived Control



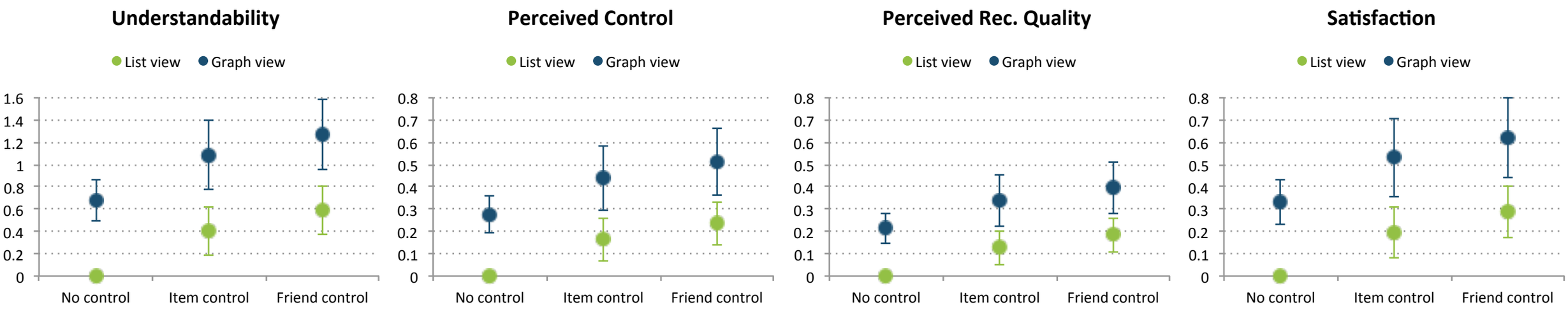
Perceived Rec. Quality

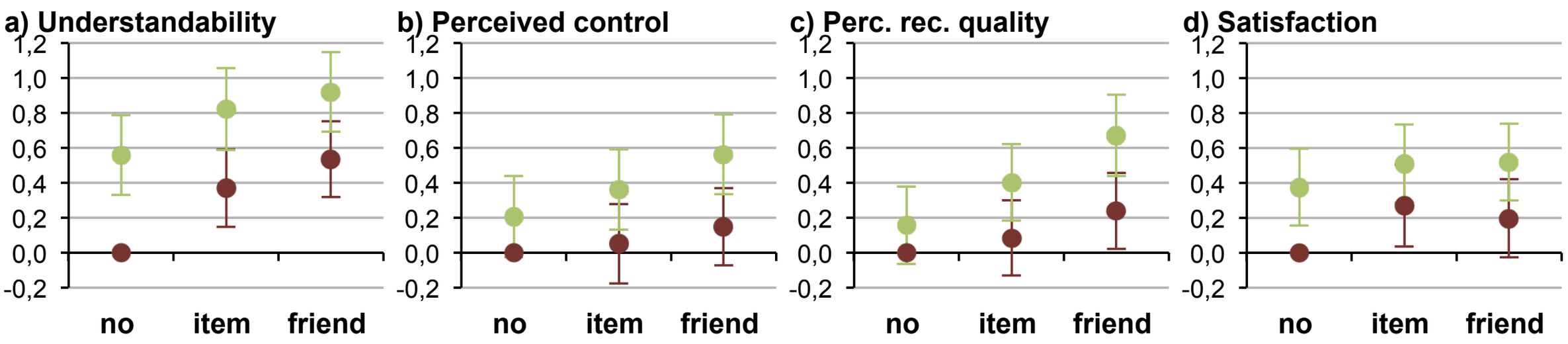


Satisfaction



Why different?





Why different?

Error bars are smaller because total effects are **mediated**

(mediation increases the accuracy of estimation)

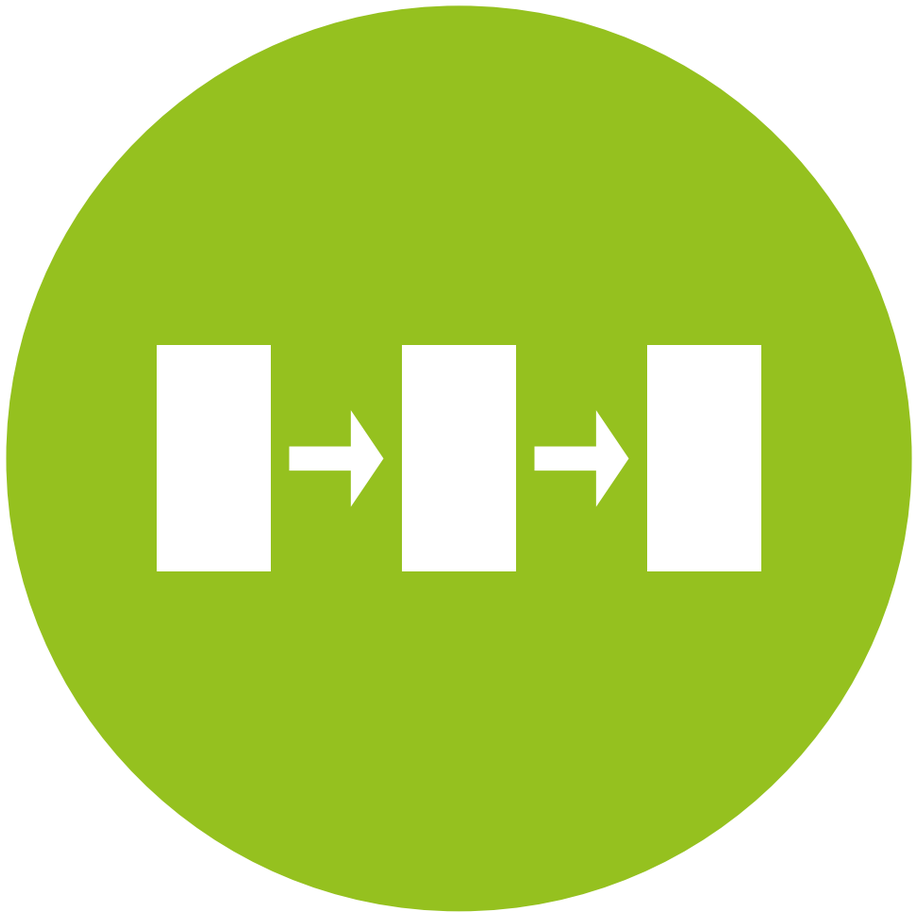
Values may be different because total effects are **modeled**

(there may be some model misspecification)

Which one should I use?

Marginal effect graphs are more “honest”

Use the p-values of the total effects (if needed); show the graphs of the marginal effects



Expanding the model

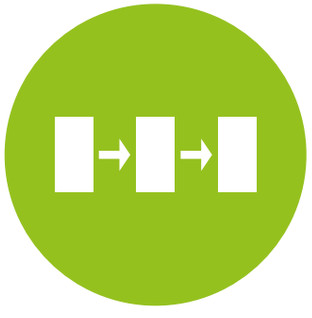
Add other variables

Expand the model

Expanding the model by adding additional variables

- Behavioral variables
- Personal and situational characteristics

Redo model tests and additional stats



Behaviors

We also measured:

- Recommendation inspection time (in seconds)
- Number of known recommendations (out of 10)
- Average rating (of the 10 recommendations)

To add these, we need to revisit the factor model!

Q: Why didn't we add these to begin with?

A: Because they are less important!

Behaviors

Inspection time

Turn into minutes (large numbers throw off the algorithm)

May be influenced by experimental conditions

May influence understandability, control, etc.

Behaviors

```
twq$mins <- twq$time/60;

model <- 'satisf =~ s1+s2+s3+s4+s5+s6+s7
          quality =~ q1+q2+q3+q4+q5+q6
          control =~ c1+c2+c3+c4
          underst =~ u2+u4+u5
          satisf ~ quality+control+mins
          quality ~ control+mins
          control ~ underst+mins
          underst ~ mins+citem+cfriend+cgraph
          mins ~ citem+cfriend+cgraph'
```

Behaviors

Regressions:				
satisf ~				
quality	0.419	0.080	5.214	0.000
control	−0.897	0.122	−7.369	0.000
mins	**0.258**	**0.102**	**2.534**	**0.011**
quality ~				
control	−0.780	0.084	−9.258	0.000
mins	**0.071**	**0.084**	**0.844**	**0.399**
control ~				
underst	−0.367	0.067	−5.447	0.000
mins	**0.129**	**0.090**	**1.434**	**0.152**
underst ~				
mins	0.245	0.088	2.793	0.005
citem	0.427	0.202	2.114	0.035
cfriend	0.660	0.201	3.288	0.001
cgraph	0.569	0.167	3.401	0.001
mins ~				
citem	−0.165	0.098	−1.685	0.092
cfriend	−0.393	0.125	−3.156	0.002
cgraph	0.286	0.091	3.128	0.002

Behaviors

Number of known recommendations (out of 10)

May be influenced by inspectability (remember the homework?)

May influence understandability, control, recommendation quality, satisfaction

Behaviors

```
model <- 'satisf =~ s1+s2+s3+s4+s5+s6+s7
          quality =~ q1+q2+q3+q4+q5+q6
control =~ c1+c2+c3+c4
underst =~ u2+u4+u5
satisf ~ quality+control+known
quality ~ control+known
control ~ underst+known
underst ~ known+mins+citem+cfriend+cgraph
mins ~ known+citem+cfriend+cgraph
known ~ citem+cfriend+cgraph'
```

Behaviors

Regressions:				
satisf ~				
quality	0.462	0.083	5.561	0.000
control	-0.884	0.119	-7.447	0.000
known	-0.142	0.051	-2.778	0.005
quality ~				
control	-0.680	0.082	-8.262	0.000
known	0.175	0.047	3.717	0.000
control ~				
underst	-0.361	0.070	-5.146	0.000
known	-0.228	0.048	-4.737	0.000
underst ~				
known	**0.031**	**0.045**	**0.683**	**0.495**
mins	0.219	0.091	2.404	0.016
citem	0.382	0.172	2.217	0.027
cfriend	0.541	0.178	3.046	0.002
cgraph	0.432	0.147	2.950	0.003
mins ~				
known	**0.013**	**0.039**	**0.325**	**0.745**
citem	-0.184	0.097	-1.898	0.058
cfriend	-0.396	0.128	-3.104	0.002
cgraph	0.279	0.095	2.933	0.003
known ~				
citem	**0.255**	**0.364**	**0.701**	**0.483**
cfriend	**0.602**	**0.355**	**1.696**	**0.090**
cgraph	0.750	0.303	2.477	0.013

Behaviors

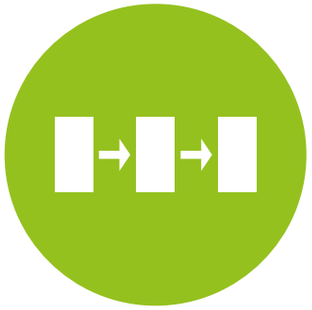
Average rating

Outcome of recommendation quality (and maybe number of known recommendations)

May influence satisfaction?

Behaviors

```
model <- 'satisf =~ s1+s2+s3+s4+s5+s6+s7
          quality =~ q1+q2+q3+q4+q5+q6
          control =~ c1+c2+c3+c4
          underst =~ u2+u4+u5
          satisf ~ rating+quality+control+known
          rating ~
quality+control+underst+mins+known+citem+cfriend+cgraph
          quality ~ control+known
          control ~ underst+known
          underst ~ mins+citem+cfriend+cgraph
          mins ~ citem+cfriend+cgraph
          known ~ cgraph'
```



Behaviors

Regressions:				
satisf ~				
rating	**-0.394**	**0.214**	**-1.841**	**0.066**
quality	0.605	0.134	4.532	0.000
control	-0.885	0.125	-7.081	0.000
known	-0.108	0.050	-2.156	0.031
rating ~				
quality	0.354	0.039	9.090	0.000
control	**-0.028**	**0.046**	**-0.614**	**0.539**
underst	**-0.050**	**0.041**	**-1.210**	**0.226**
mins	**-0.052**	**0.047**	**-1.109**	**0.267**
known	0.048	0.021	2.252	0.024
citem	**-0.183**	**0.089**	**-2.060**	**0.039**
cfriend	**0.054**	**0.101**	**0.533**	**0.594**
cgraph	**-0.054**	**0.077**	**-0.698**	**0.485**
...	**...**	**...**	**...**	**...**

Additional factors

We also measured:

- Music expertise (e1-e4)
- Trusting propensity (t1-t6)
- Familiarity with recommenders (f1-f6)

Additional factors

```
model <- 'satisf =~ s1+s2+s3+s4+s5+s6+s7
quality =~ q1+q2+q3+q4+q5+q6
control =~ c1+c2+c3+c4
underst =~ u2+u4+u5
expert =~ e1+e2+e3+e4
trust =~ t1+t2+t3+t4+t5+t6
fam =~ f1+f2+f3+f4+f5+f6’

fit <- sem(model,data=twq,ordered=names(twq[9:47]),std.lv=T)

summary(fit, rsquare=T)
```

Additional factors

Trimming:

- Remove t5 and t6 (low R^2)
- Remove t4 (low R^2)

Look at the modification indices:

There are a ton of things wrong with “familiarity”…

Additional factors

Items:

- I am familiar with online recommender systems.
- I have occasionally followed the advice of a recommender system.
- I have heard of systems similar to TasteWeights.
- I have never used anything like TasteWeights before.
- I regularly use systems similar to TasteWeights.
- Using TasteWeights was a completely new experience for me.

Additional factors

Might be two factors!

f1-f2: familiarity with recommender systems (fam)

f3-f6: familiarity with system like TasteWeights (famtw)

We eventually decided to only keep fam (remove famtw)

Model:

```
model <- 'satisf =~ s1+s2+s3+s4+s5+s6+s7
quality =~ q1+q2+q3+q4+q5+q6
control =~ c1+c2+c3+c4
underst =~ u2+u4+u5
expert =~ e1+e2+e3+e4
trust =~ t1+t2+t3
fam =~ f1+f2'
```

New model

```
model <- 'satisf =~ s1+s2+s3+s4+s5+s6+s7
quality =~ q1+q2+q3+q4+q5+q6
control =~ c1+c2+c3+c4
underst =~ u2+u4+u5
expert =~ e1+e2+e3+e4
trust =~ t1+t2+t3
fam =~ f1+f2

satisf ~ quality+control+known+expert+trust+fam
rating ~ quality+known+expert+trust+fam
quality ~ control+known+expert+trust+fam
control ~ underst+known+expert+trust+fam
underst ~ expert+trust+fam+mins+citem+cfriend+cgraph
mins ~ expert+trust+fam+citem+cfriend+cgraph
known ~ expert+trust+fam+cgraph'
```

New model

Regressions:				
satisf ~				
quality	0.425	0.085	4.983	0.000
control	-0.898	0.136	-6.615	0.000
known	-0.159	0.065	-2.426	0.015
expert	0.131	0.089	1.470	0.141
trust	0.140	0.078	1.794	0.073
fam	**0.132**	**0.098**	**1.339**	**0.180**
rating ~				
quality	0.324	0.031	10.540	0.000
known	0.071	0.023	3.146	0.002
expert	**0.006**	**0.034**	**0.163**	**0.870**
trust	**0.045**	**0.033**	**1.372**	**0.170**
fam	**-0.020**	**0.035**	**-0.573**	**0.567**
quality ~				
control	-0.719	0.087	-8.287	0.000
known	0.143	0.049	2.895	0.004
expert	0.212	0.081	2.633	0.008
trust	**0.008**	**0.080**	**0.102**	**0.919**
fam	**0.197**	**0.084**	**2.349**	**0.019**
control ~				
underst	-0.410	0.077	-5.315	0.000
known	-0.280	0.050	-5.554	0.000
expert	0.249	0.081	3.062	0.002
trust	**-0.110**	**0.075**	**-1.479**	**0.139**
fam	**0.281**	**0.082**	**3.420**	**0.001**

New model

understand ~				
expert	**0.072**	**0.079**	**0.908**	**0.364**
trust	**0.009**	**0.068**	**0.131**	**0.896**
fam	0.167	0.073	2.307	0.021
mins	0.211	0.110	1.911	0.056
citem	0.421	0.208	2.024	0.043
cfriend	0.656	0.205	3.195	0.001
cgraph	0.456	0.147	3.108	0.002
mins ~				
expert	**0.027**	**0.048**	**0.577**	**0.564**
trust	**0.073**	**0.051**	**1.416**	**0.157**
fam	**−0.009**	**0.038**	**−0.233**	**0.816**
citem	−0.181	0.097	−1.874	0.061
cfriend	−0.389	0.126	−3.093	0.002
cgraph	0.288	0.091	3.176	0.001
known ~				
expert	**0.166**	**0.108**	**1.541**	**0.123**
trust	**−0.078**	**0.104**	**−0.751**	**0.453**
fam	**0.056**	**0.109**	**0.512**	**0.609**
cgraph	0.667	0.298	2.236	0.025

After trimming...

```
Regressions:
  satisf ~
    quality          0.411    0.092     4.449    0.000
    control         -0.956    0.148    -6.456    0.000
    known           -0.153    0.064    -2.397    0.017
    expert           0.205    0.100     2.055    0.040
    trust            0.258    0.124     2.071    0.038
rating ~
  quality            0.323    0.031    10.445    0.000
  known              0.066    0.022     3.047    0.002
quality ~
  control           -0.770    0.094    -8.216    0.000
  known              0.148    0.051     2.893    0.004
  expert             0.374    0.094     3.987    0.000
control ~
  underst           -0.376    0.074    -5.087    0.000
  known             -0.249    0.050    -5.037    0.000
  expert             0.332    0.088     3.775    0.000
```

After trimming...

understood ~				
fam	0.165	0.076	2.166	0.030
mins	0.230	0.114	2.025	0.043
citem	0.425	0.207	2.059	0.039
cfriend	0.665	0.205	3.247	0.001
cgraph	0.457	0.147	3.105	0.002
mins ~				
citem	-0.181	0.097	-1.874	0.061
cfriend	-0.389	0.126	-3.093	0.002
cgraph	0.288	0.091	3.176	0.001
known ~				
cgraph	0.698	0.305	2.291	0.022

Expanded model

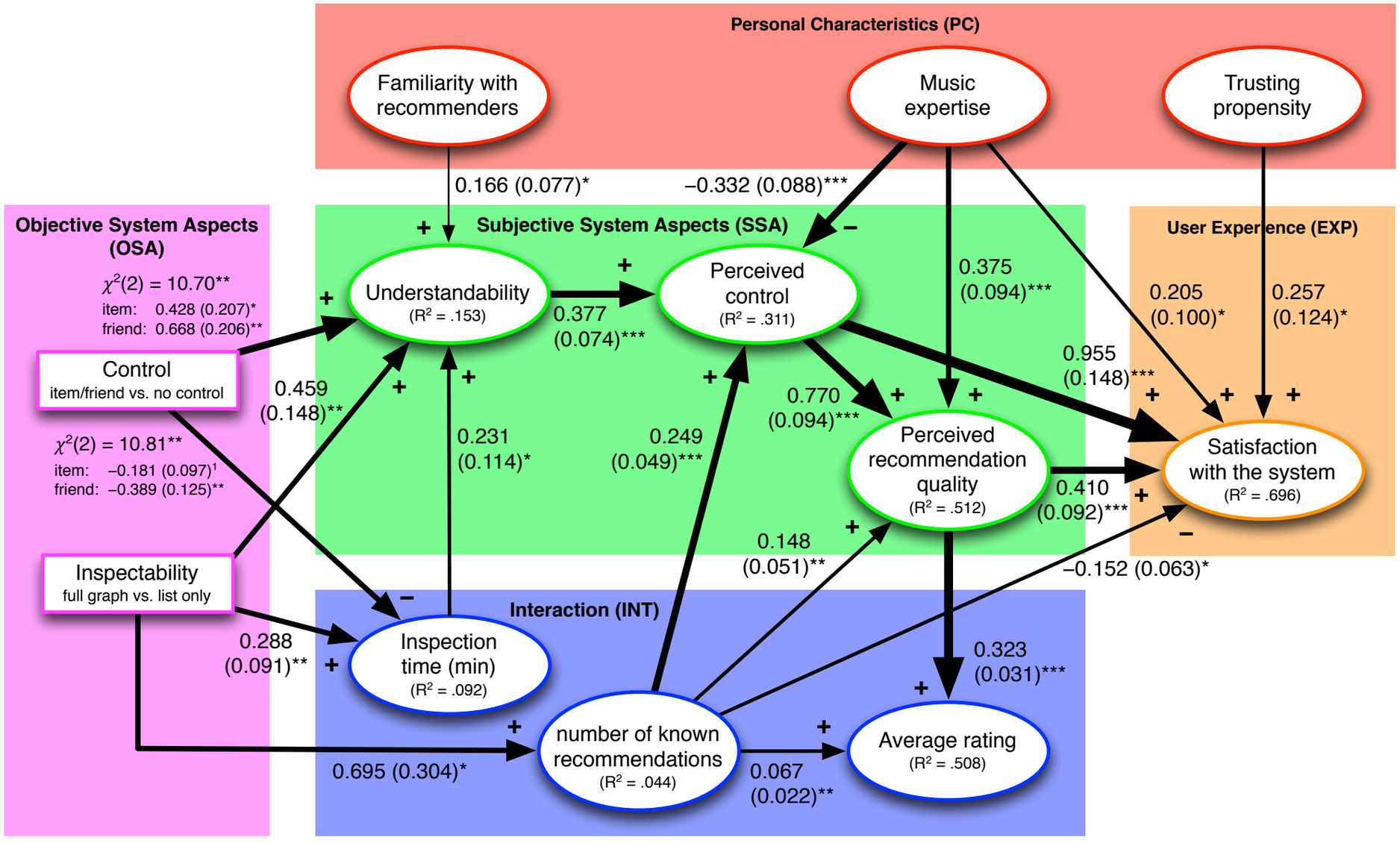


Figure 3. The structural equation model for the data of the experiment. Significance levels: *** $p < .001$, ** $p < .01$, 'ns' $p > .05$. R^2 is the proportion of variance explained by the model. Numbers on the arrows (and their thickness) represent the β coefficients (and standard error) of the effect. Factors are scaled to have an SD of 1.

“It is the mark of a truly intelligent person
to be moved by statistics.”

T H A N K S !

George Bernard Shaw