Part 3: SEM

Quantitative Research Methods Seminar

Slides

Feel free to share these slides with anyone

This is version 1.1. For the most recent version of these slides, visit www.usabart.nl/QRMS

If you want to use these slides in your own lectures, use the above link for attribution

SEM

In this part I discuss the following:

- Why SEM?
- Marginal effects (for experiments)
- Modeling: theory
- Modeling: practice

What is SEM?

A Structural Equation Model (SEM) is a CFA where the factors are regressed on each other and on the experimental manipulations

(observed behaviors can also be incorporated)

The regressions are not estimated one-by-one, but at the same time

(and so is the CFA part of the model, actually)

Why SEM?

Benefits of Structural Equation Modeling

Why SEM?

Easy way to test for **mediation**

...without doing many separate tests

You can **keep factors** as factors

This ascertains normality, and leads to more statistical power in the regressions

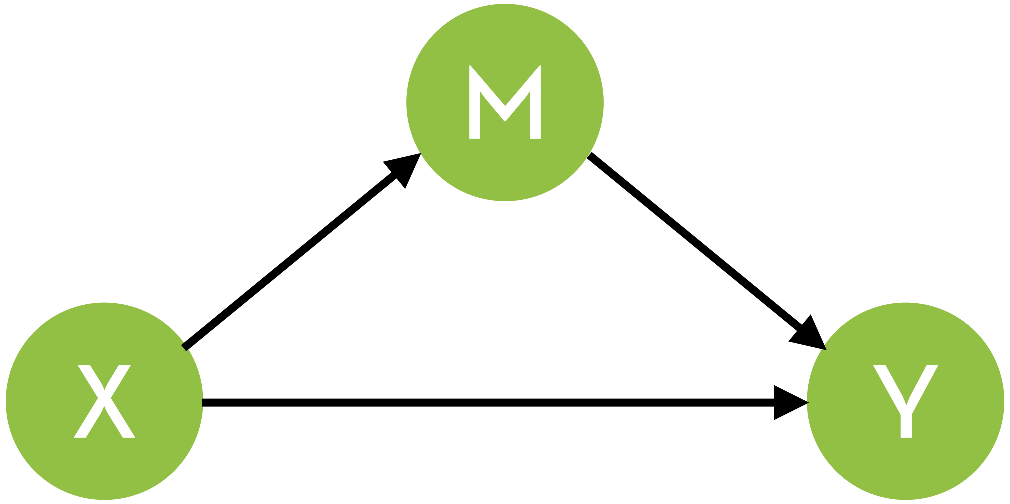
The model has several **overall fit indices**

You can judge the fit of an entire model, rather than just its parts

Mediation Analysis

X -> M -> Y

Does the system (X) influence usability (Y) via understandability (M)?



Types of mediation

Partial mediation

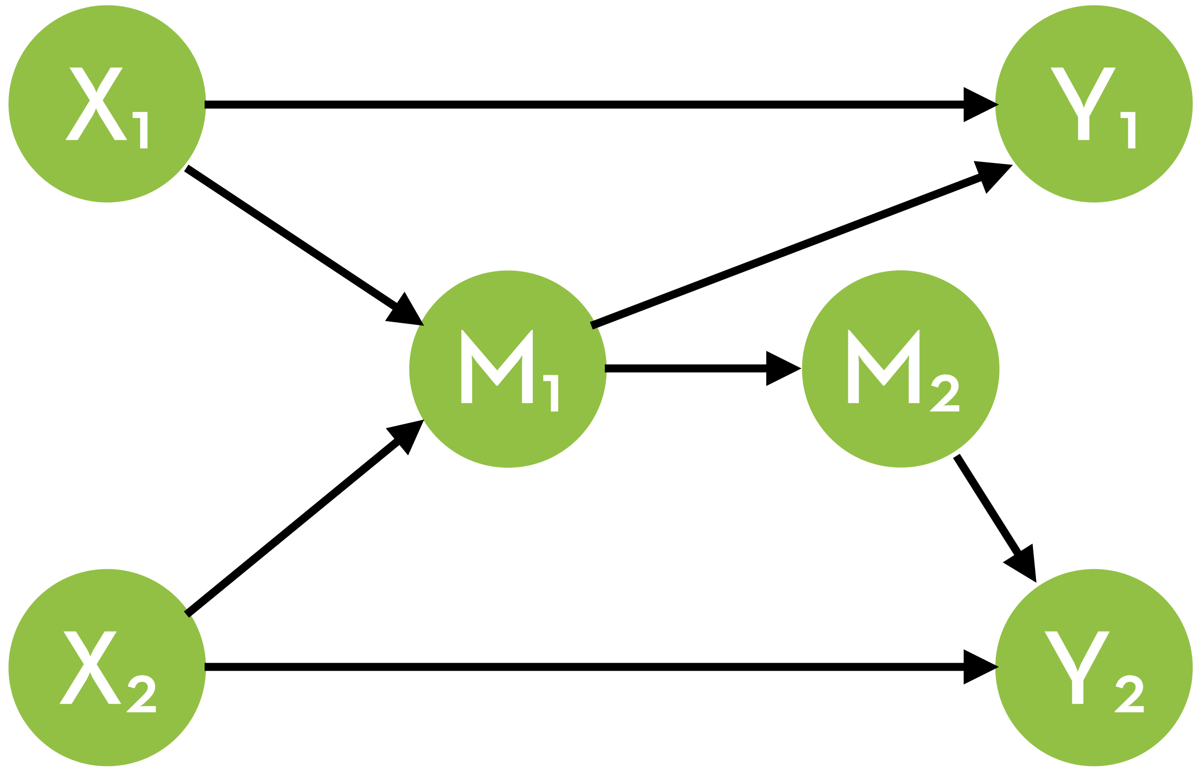
Full mediation

Negative mediation

Mediation Analysis

More complex models:

- What is the total effect of X1 on Y2?
- Is this effect significant?
- Is this effect fully or partially mediated by M1 and M2?



Keep the factors!

Let's say we have a factor F measuring trait Y, with AVE = 0.64

- On average, 64% of the item variance is communality, 36% is uniqueness

If we **sum the items** of the factor as S, this results in 36% error

- This is random noise that does not measure Y

Result: no regression with S as dependent can have an R-squared > 0.64!

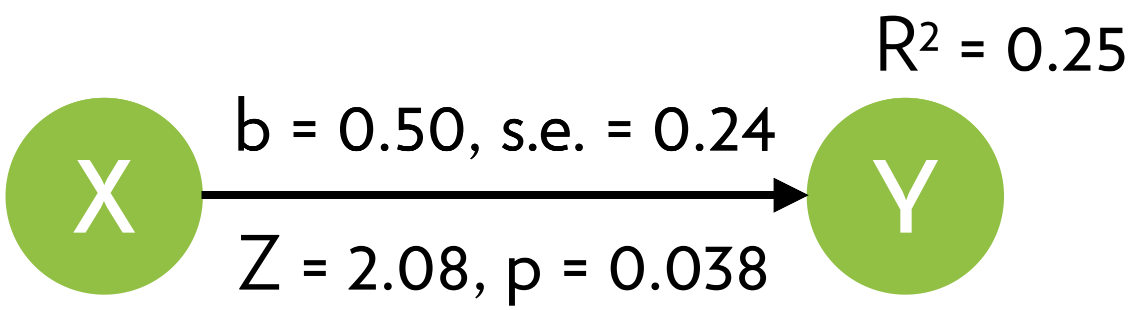
Keep the factors!

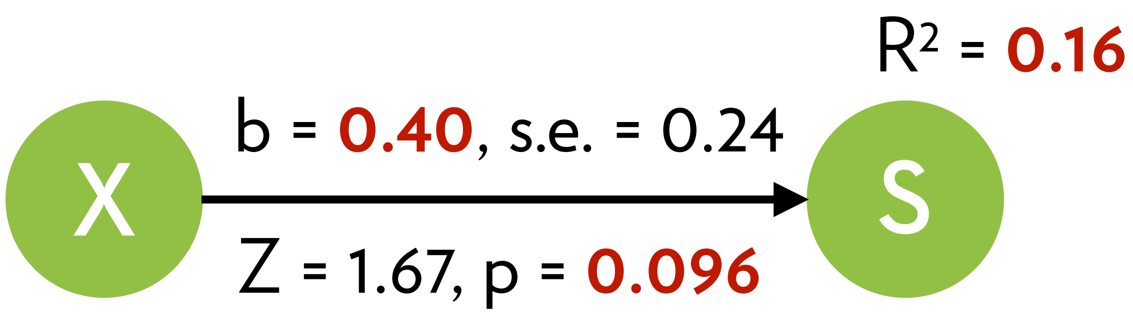
Any regression coefficient will be **attenuated** by the AVE of S!

Take for instance this X, which potentially explains 25% of the variance of Y...

...it only explains 16% of the variance of S!

...and the effect is non-significant!



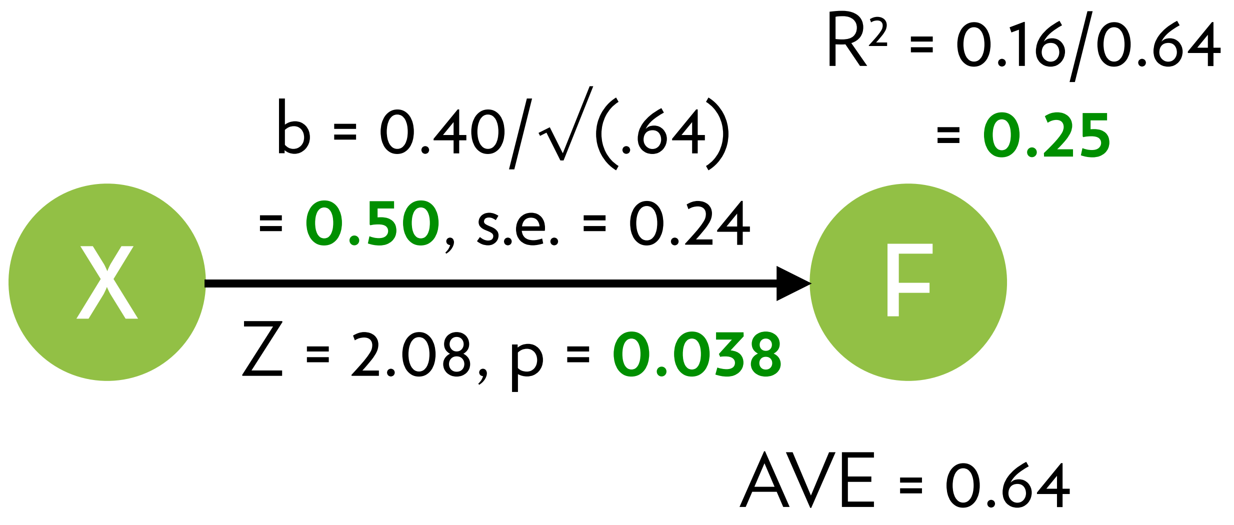


BREN:ICS INFORMATION AND COMPUTER SCIENCES

Keep the factors!

If we use F instead of S, we **know** that the AVE is 0.64

…so we can **compensate** for the incurred measurement error!



Estimates

In a SEM you can get the following estimates (all at once):

- Item loadings (see CFA slides; session 2)
- R^2 for every dependent variable (usually factors)
- Regression coefficients for all regressions (B, s.e., p-values)
- Total (mediated and non-mediated) effects

Plus, you can get omnibus tests for testing manipulations with > 2 conditions

- You have to run these one by one, though

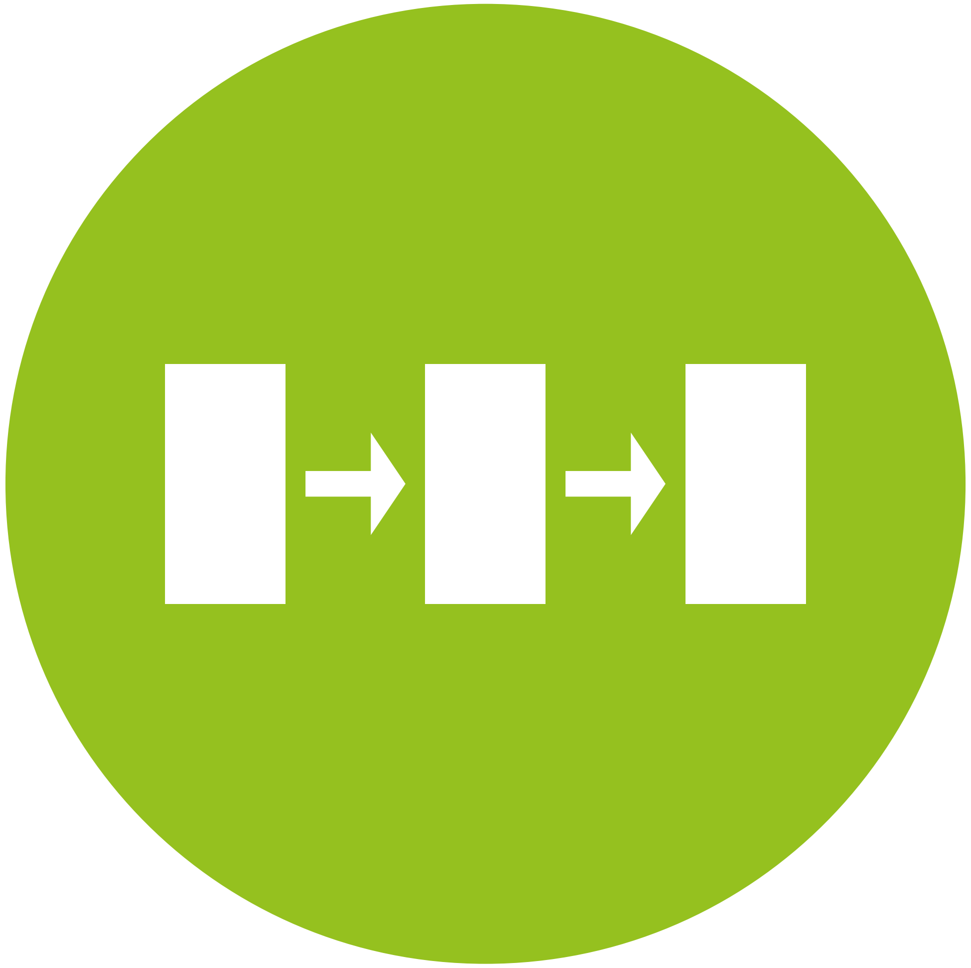
Fit statistics

Same fit statistics as in CFA! As a reminder:

- Item-fit: Loadings, communality, residuals
- Factor-fit: Average Variance Extracted
- Model-fit: Chi-square test, CFI, TLI, RMSEA

Also: modification indices for model improvement purposes

- Not just for items/factors, but also for regression coefficients!



Marginal effects

Getting an idea of the effect of experimental conditions

Marginal effects

First analysis: manipulations —> factors

MIMIC model (Multiple Indicators, Multiple Causes)

The SEM equivalent of a t-test / (factorial) ANOVA

Only for experiments (not for surveys)

Steps involved:

- Build your CFA (see session 2 slides)
- Create dummies for your experimental conditions
- Run regressions factor-by-factor

Create your CFA

Take the final CFA from last week

E.g., 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'
```

Don't run it yet! We are going to add extra lines to this model...

Create dummies

Main effects are already built for our dataset:

- Control conditions ("no control" is the baseline):

  `citem cfriend`

- Inspectability conditions ("list view" is the baseline):

  `cgraph`

What about the interaction effect?

- We need to create dummies for that too!

Create dummies

In MPlus, add:

```
DEFINE:
    cig = citem * cgraph;
    cfg = cfriend * cgraph;
```

In R, run:

```
twq$cig = twq$citem * twq$cgraph;
twq$cfg = twq$cfriend * twq$cgraph;
```

Run regressions

In MPlus (note the different notation for standardization!):

```
<...>

DEFINE:
 cig = citem * cgraph;
 cfg = cfriend * cgraph;

MODEL:
 satisf BY s1* s2-s7;
 quality BY q1* q2-q6;
 control BY c1* c2-c4;
 underst BY u2* u4-u5;
 satisf-underst@1;

 satisf ON citem cfriend cgraph cig cfg;
```

BREN:ICS INFORMATION AND COMPUTER SCIENCES

Run regressions

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 ~ citem+cfriend+cgraph+cig+cfg';

fit <-
sem(model,data=twq,ordered=names(twq[1:23]),std.lv=TRUE);

summary(fit);
```

Results

Note: effects are not significant (but that's okay for now)

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
SATISF ON				
CITEM	0.269	0.233	1.155	0.248
CFRIEND	0.197	0.223	0.883	0.377
CGRAPH	0.375	0.221	1.696	0.090
CIG	-0.131	0.320	-0.409	0.683
CFG	-0.048	0.309	-0.157	0.875

Interpretation

Citem: effect of item control vs. no control in the list view condition

Cfriend: effect of friend control vs. no control in the list view condition

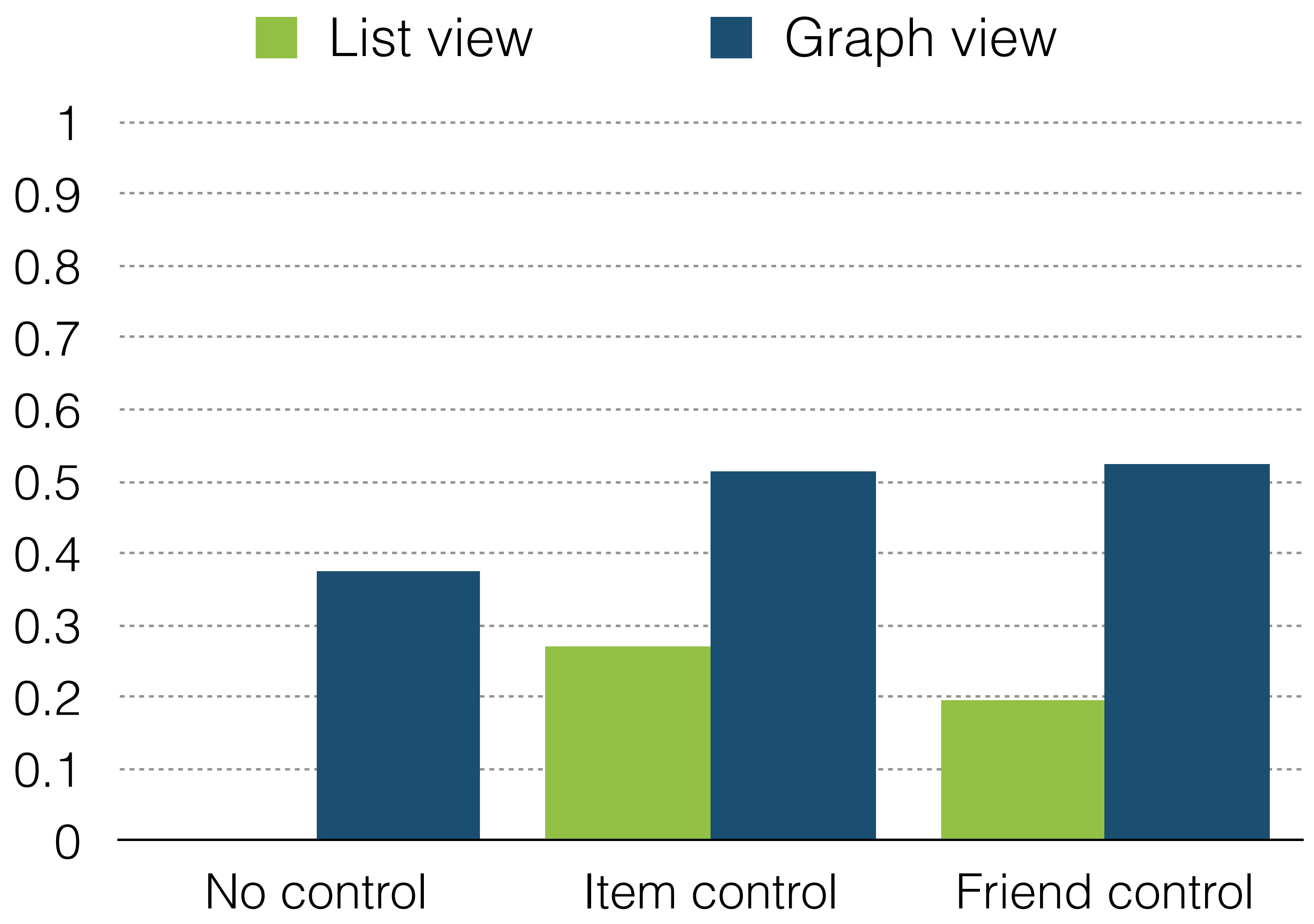
Cgraph: effect of graph view vs. list view in the "no control" condition

Cig: additional effect of item control in the graph view condition (or: additional effect of graph view in the item control condition)

Cfg: additional effect of friend control in the graph view condition (or: additional effect of graph view in the friend control condition)

Graph

Note: no control, list view is set to zero!



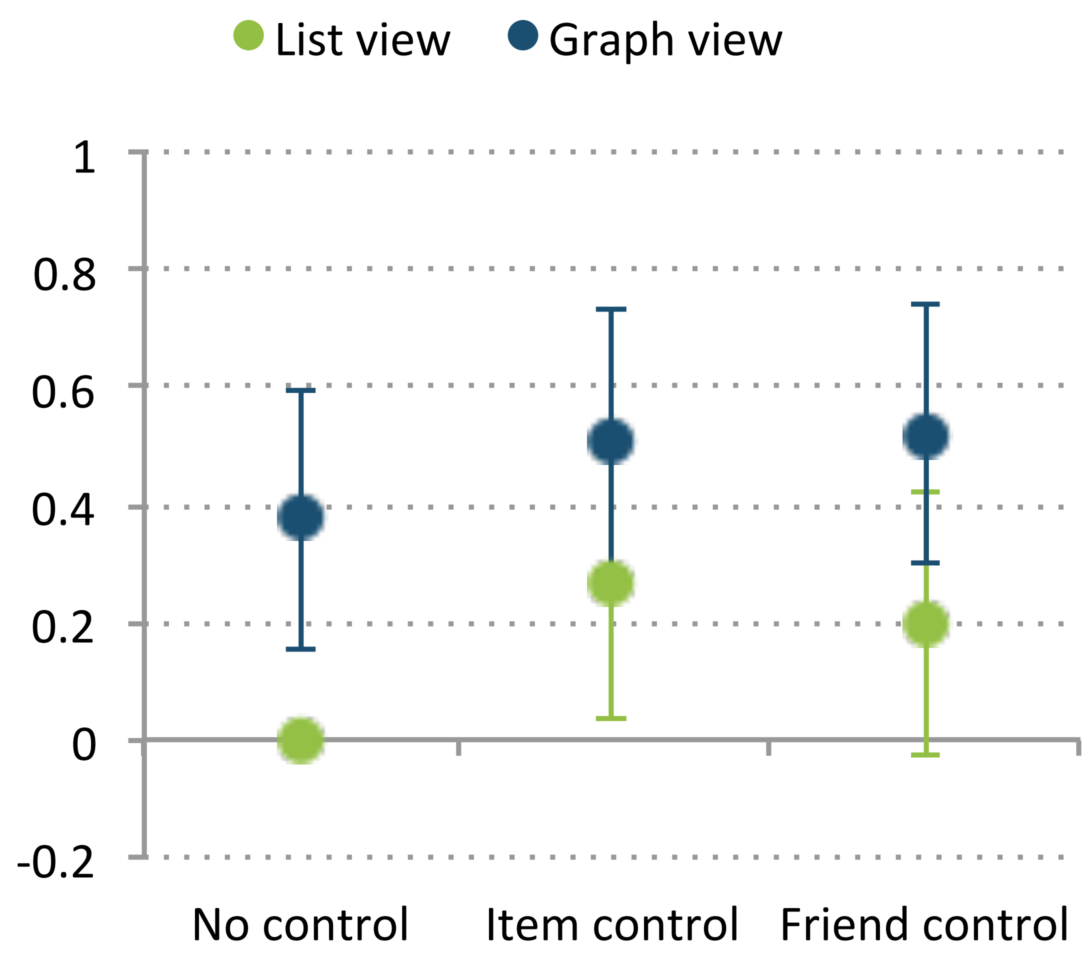
For a better graph

```
<...>

DEFINE:
 cil = citem * (1-cgraph);
 cfl = cfriend * (1-cgraph);
 cng = (1-citem) * (1-cfriend) * cgraph;
 cig = citem * cgraph;
 cfg = cfriend * cgraph;

MODEL:
 satisf BY s1* s2-s7;
 quality BY q1* q2-q6;
 control BY c1* c2-c4;
 underst BY u2* u4-u5;
 satisf-underst@1;

 satisf ON cil cfl cng cig cfg;
```

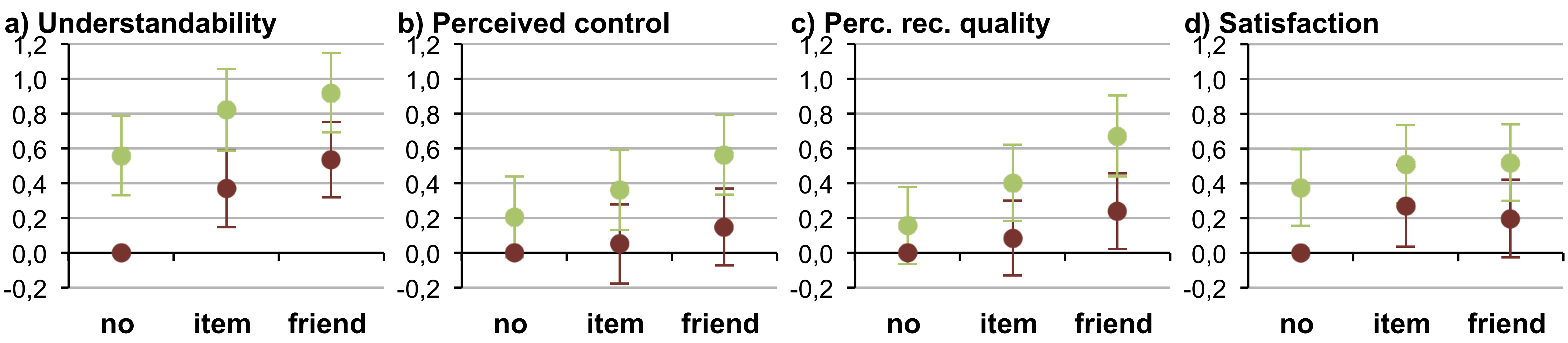
Better graph

Includes error bars (+/- 1 SE)

Easier to see that baseline is fixed to zero



Repeat



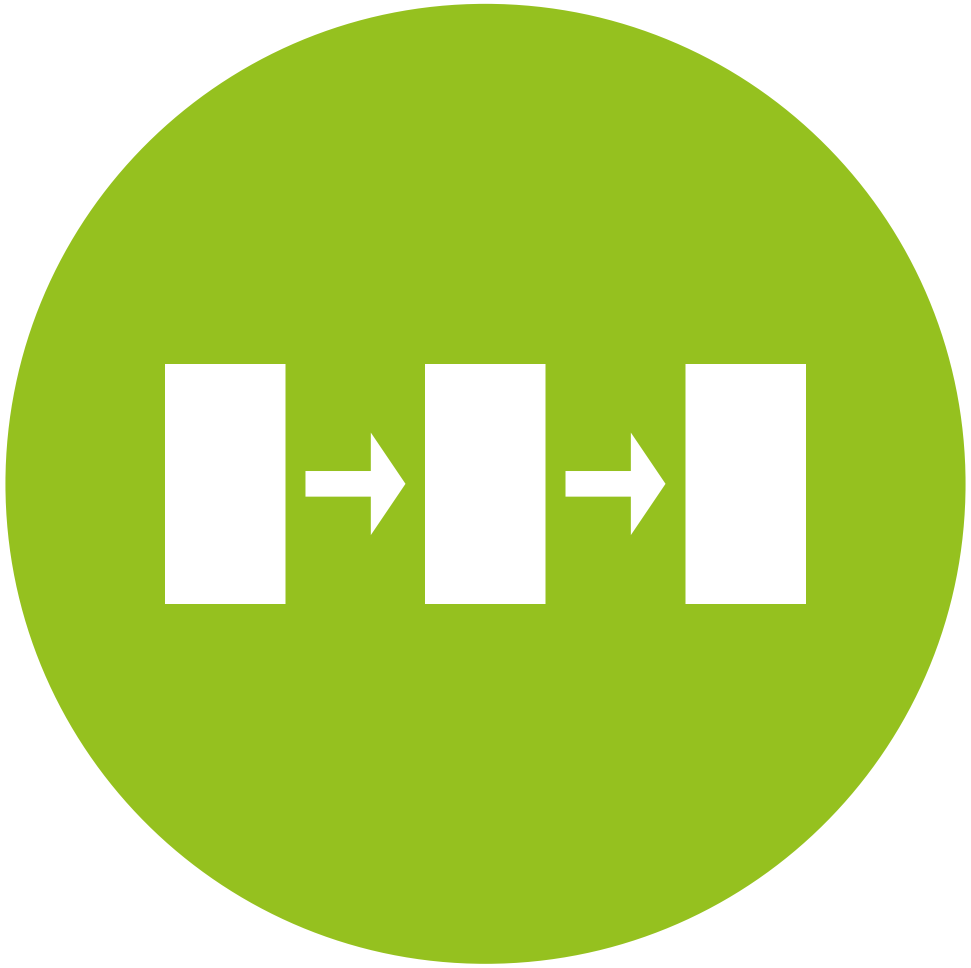
From: Knijnenburg et al. (2012): "Inspectability and Control in Social Recommenders", *RecSys'12*

Main finding

Main effects of inspectability and control conditions on understandability (no interaction effect)

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
UNDERST ON				
CITEM	0.365	0.229	1.598	0.110
CFRIEND	0.562	0.223	2.525	0.012
CGRAPH	0.596	0.232	2.566	0.010
CIG	-0.050	0.332	-0.151	0.880
CFG	-0.169	0.326	-0.519	0.604

BREN:ICS INFORMATION AND COMPUTER SCIENCES



Modeling: theory

Creating a research model

Modeling: theory

Do this **before** you do your study!

Motivate expected effects, based on:

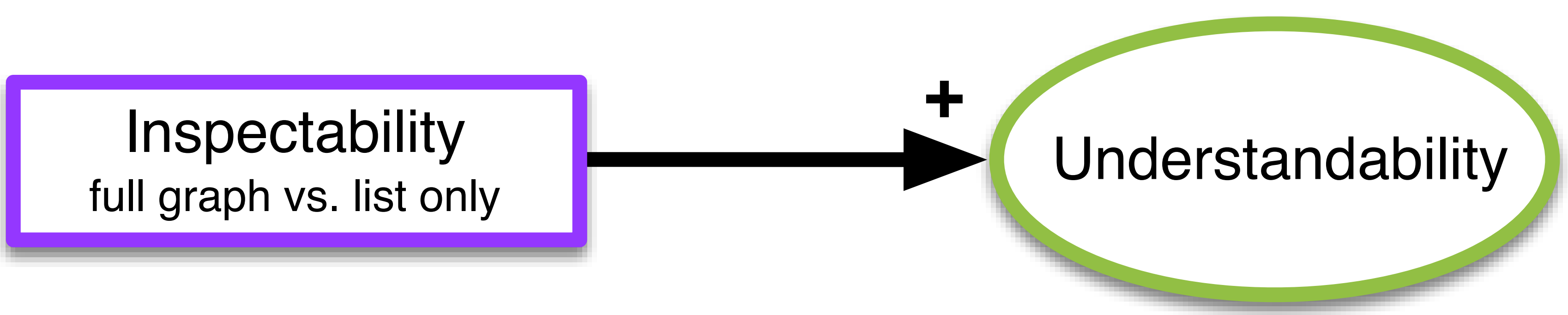
- previous work
- theory
- common sense

If in doubt, create alternate specifications!

BREN:ICS INFORMATION AND COMPUTER SCIENCES

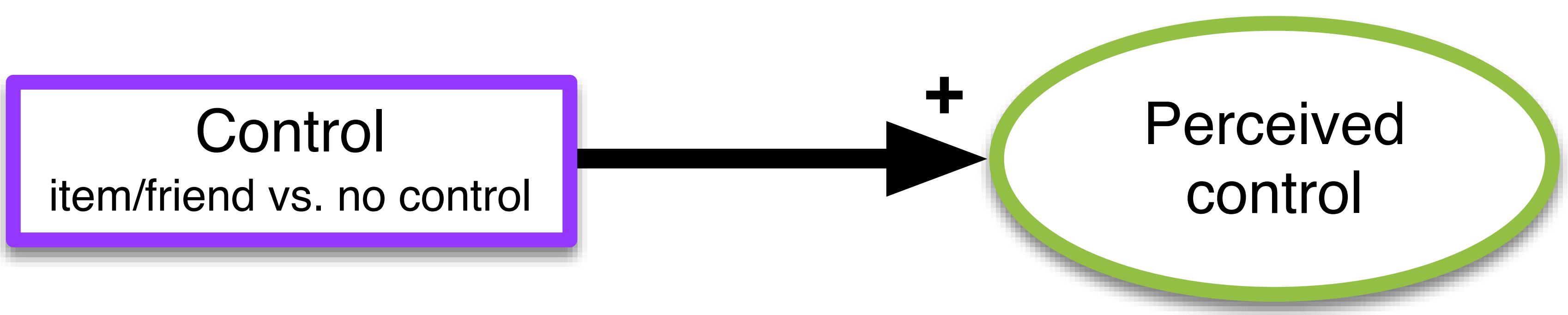
Inspectability

Herlocker argues that explanation provides transparency, "exposing the reasoning behind a recommendation".



Control

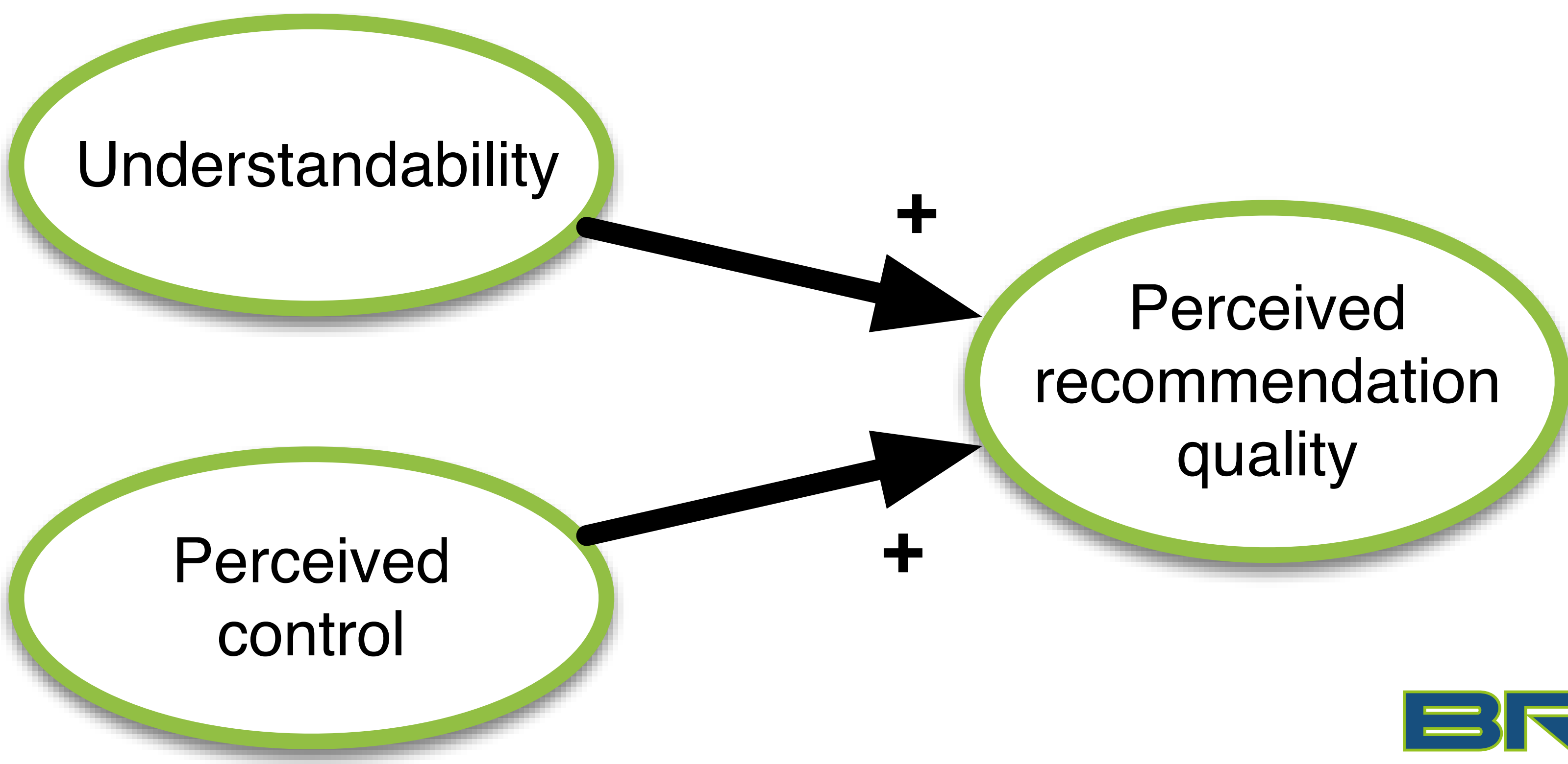
Multiple studies highlight the benefits of interactive interfaces that support control over the recommendation process.



Perceived quality

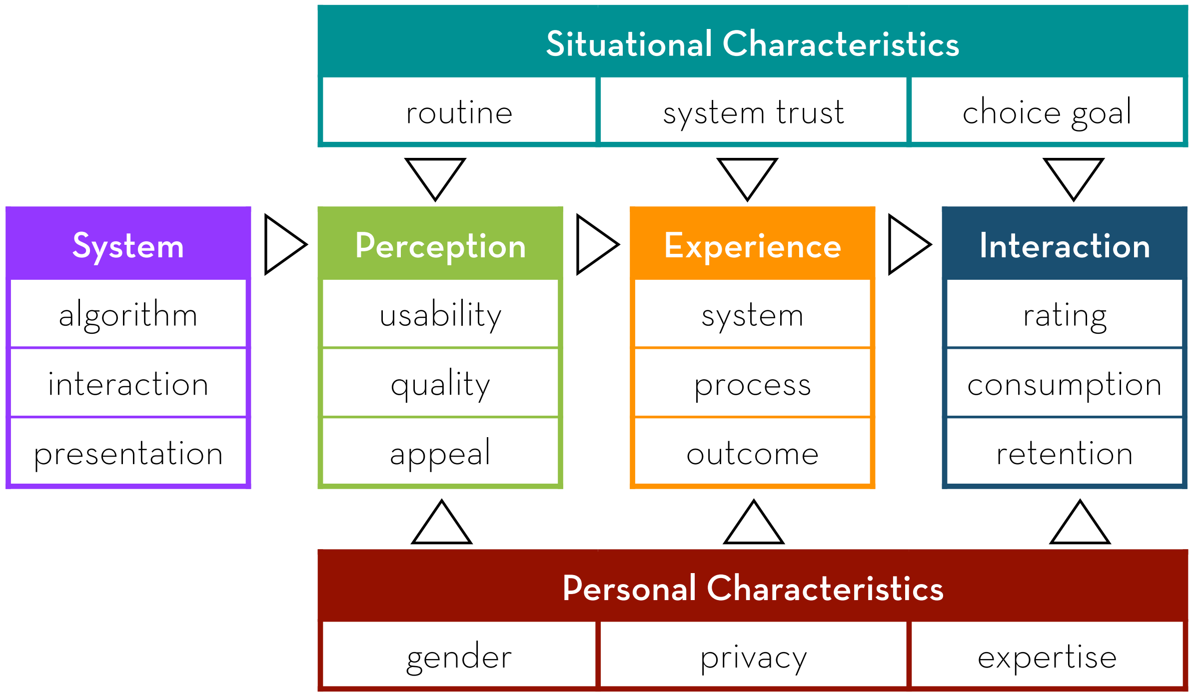
Tintarev and Masthoff show that explanations make it easier to judge the quality of recommendations.

McNee et al. found that study participants preferred user-controlled interfaces because these systems “best understood their tastes”.



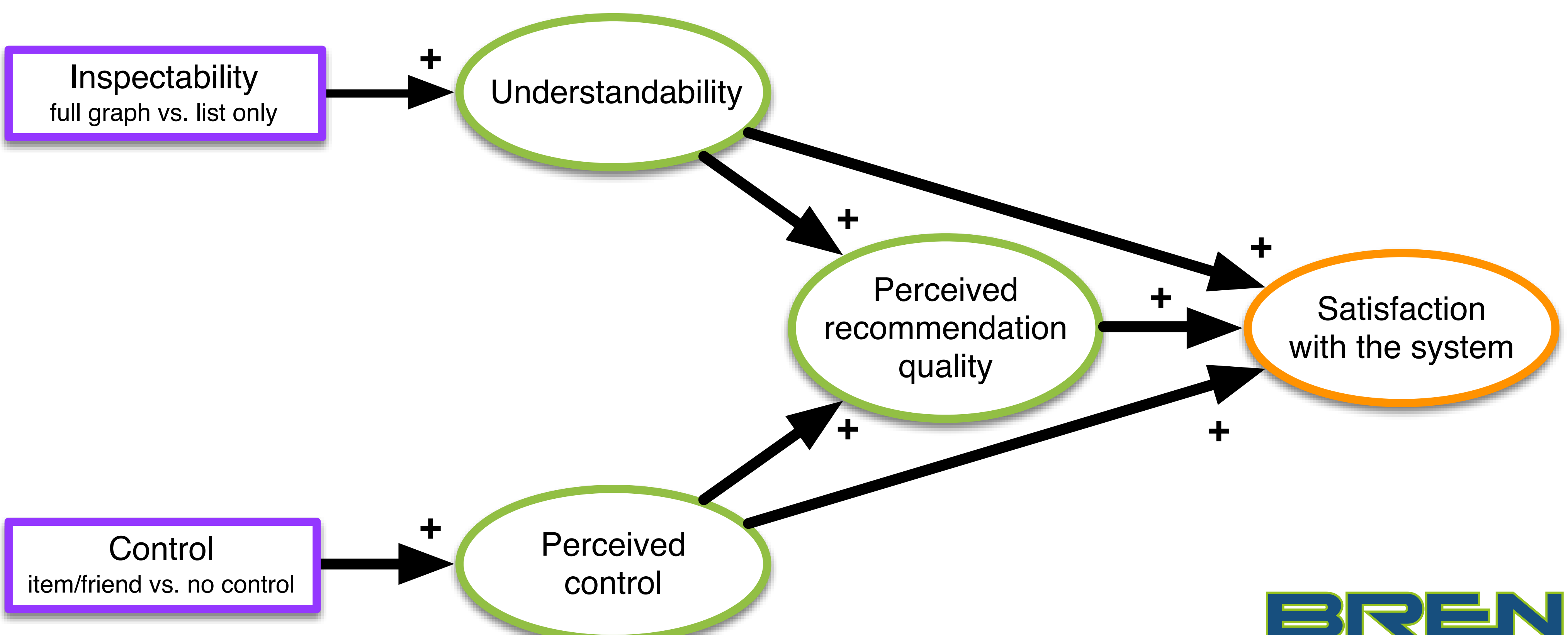
Satisfaction

Knijnenburg et al. developed a framework that describes how certain manipulations influence subjective system aspects (i.e. understandability, perceived control and recommendation quality), which in turn influence user experience (i.e. system satisfaction).



Satisfaction

Knijnenburg et al. developed a framework that describes how certain manipulations influence subjective system aspects (i.e. understandability, perceived control and recommendation quality), which in turn influence user experience (i.e. system satisfaction).



Modeling: practice

Testing your research model

Modeling: practice

Steps:

- Build and trim the core model
- Get model fit statistics
- Optional: expand the model
- Reporting

Model building

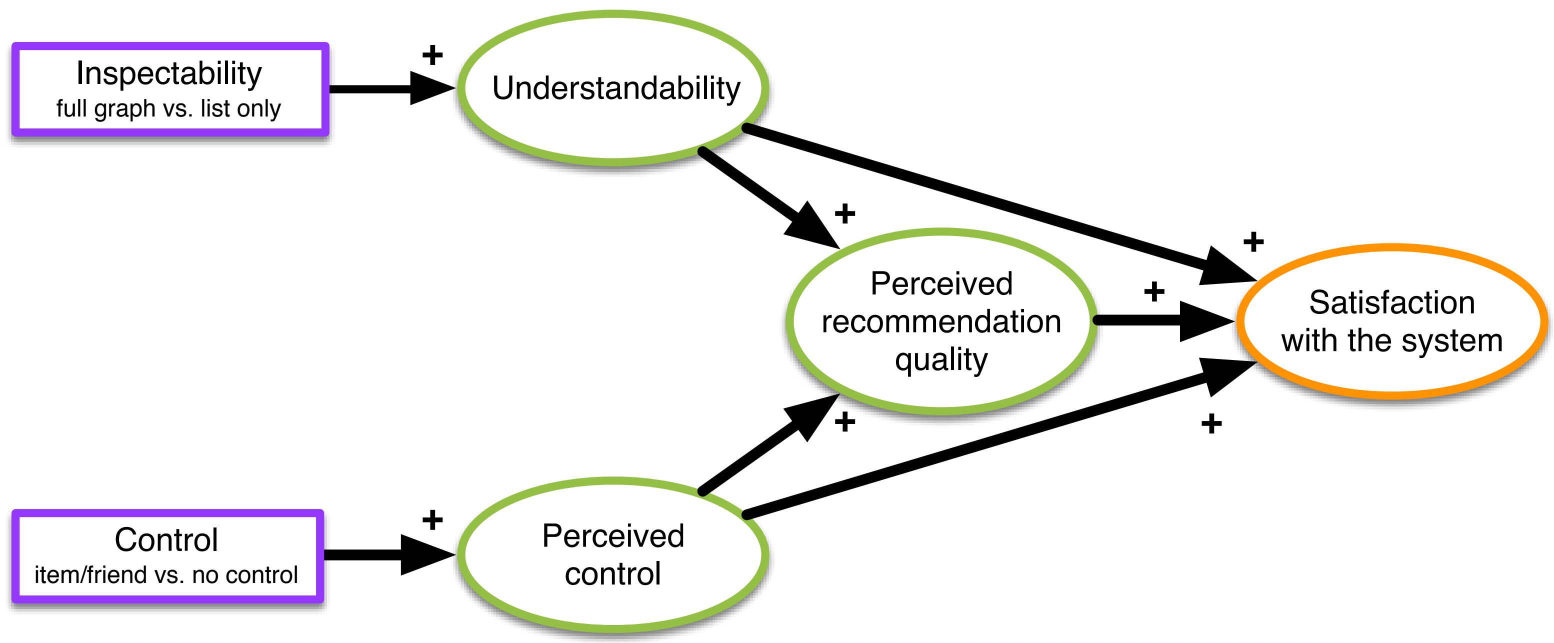
Steps:

- Determine the causal order and create a saturated model
- Trim the model
- Inspect modification indices
- Try alternative specifications, pick the best alternative (optional)

Causal order

Find the causal order of your model

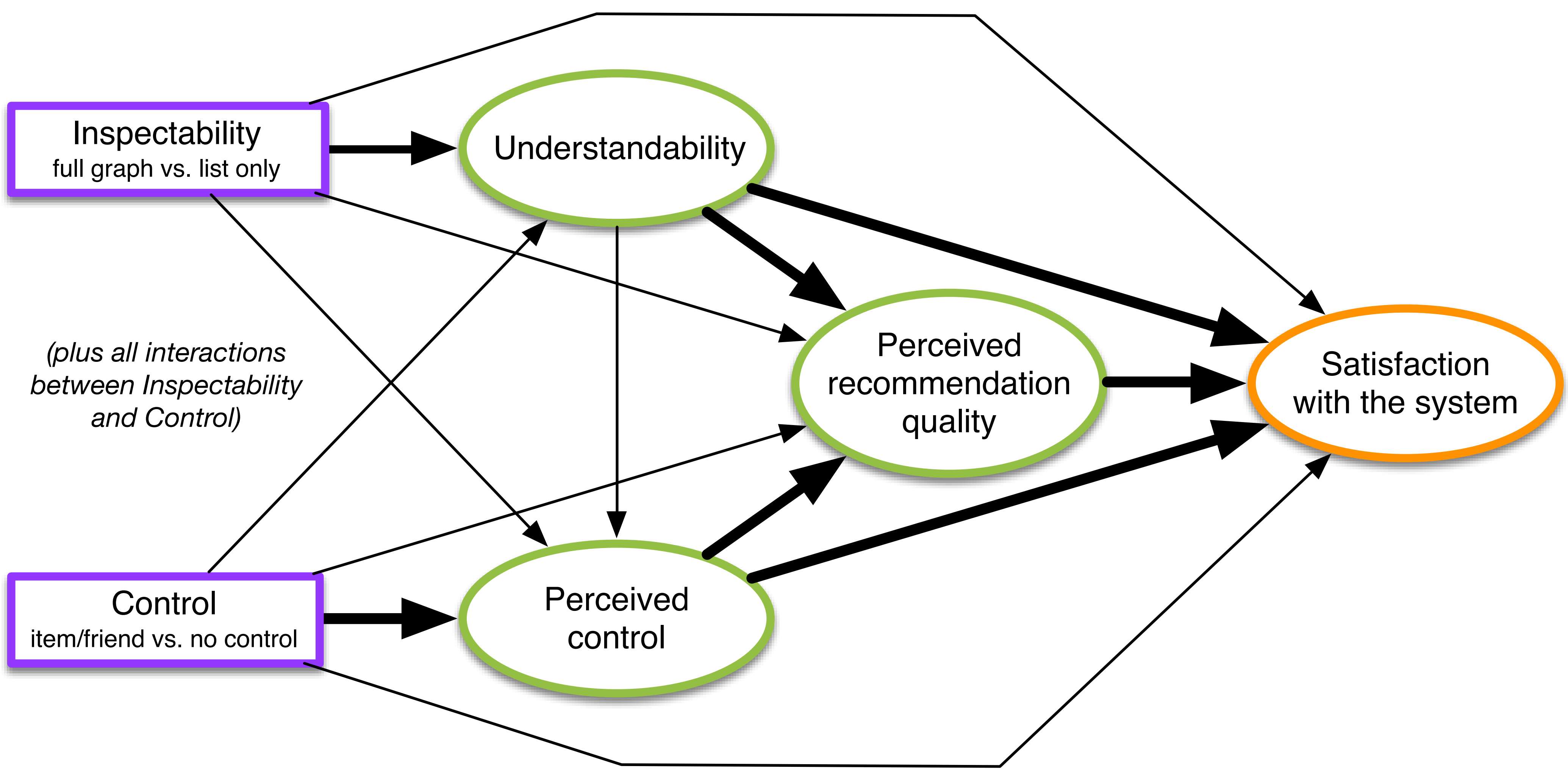
(fill the gaps where necessary)



conditions -> understandability -> perceived control -> perceived recommendation quality -> satisfaction

Saturated model

Fill in all forward-going arrows



BREN:ICS
INFORMATION AND COMPUTER SCIENCES

Run model

In MPlus:

```
MODEL:
 satisf BY s1* s2-s7;
 quality BY q1* q2-q6;
 control BY c1* c2-c4;
 underst BY u2* u4-u5;
 satisf-underst@1;

 satisf ON quality control underst citem cfriend cgraph cig cfg;
 quality ON control underst citem cfriend cgraph cig cfg;
 control ON underst citem cfriend cgraph cig cfg;
 underst ON citem cfriend cgraph cig cfg;
```

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[1:23]),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 only one is significant, keep the others as well)
- Interaction+main effects: never remove main effect before the interaction effect (if only the interaction is significant, keep the main effect regardless)

Results

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
SATISF ON				
QUALITY	0.438	0.076	5.744	0.000
CONTROL	-0.832	0.108	-7.711	0.000
UNDERST	0.105	0.078	1.354	0.176
QUALITY ON				
CONTROL	-0.757	0.085	-8.877	0.000
UNDERST	0.057	0.076	0.754	0.451
CONTROL ON				
UNDERST	-0.322	0.069	-4.685	0.000
SATISF ON				
CITEM	0.313	0.263	1.190	0.234
CFRIEND	0.004	0.256	0.014	0.988
CGRAPH	0.297	0.228	1.302	0.193
CIG	-0.389	0.356	-1.092	0.275
CFG	-0.391	0.356	-1.097	0.273

Results

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
QUALITY ON				
CITEM	0.041	0.203	0.203	0.839
CFRIEND	0.157	0.250	0.628	0.530
CGRAPH	0.000	0.235	-0.001	0.999
CIG	**0.105**	**0.316**	**0.333**	**0.739**
CFG	**0.182**	**0.373**	**0.488**	**0.625**
CONTROL ON				
CITEM	0.057	0.243	0.234	0.815
CFRIEND	0.024	0.221	0.109	0.913
CGRAPH	-0.024	0.240	-0.100	0.921
CIG	**-0.132**	**0.343**	**-0.384**	**0.701**
CFG	**-0.273**	**0.330**	**-0.828**	**0.408**
UNDERST ON				
CITEM	0.365	0.229	1.596	0.110
CFRIEND	0.562	0.223	2.522	0.012
CGRAPH	0.596	0.232	2.568	0.010
CIG	**-0.050**	**0.332**	**-0.149**	**0.881**
CFG	**-0.169**	**0.326**	**-0.518**	**0.604**

BREN:ICS INFORMATION AND COMPUTER SCIENCES

Trimming steps

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

with the latter, also remove the dummies from usevariables

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

BREN:ICS INFORMATION AND COMPUTER SCIENCES

Trimming steps

Remove citem and cfriend -> recommendation quality

Remove cgraph -> control

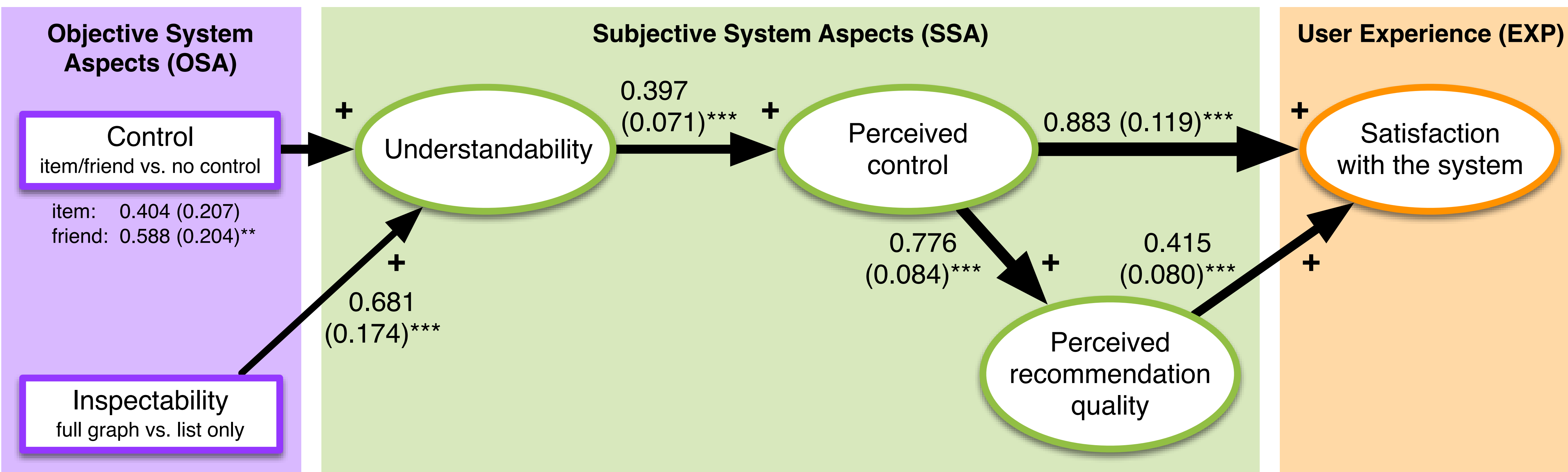
- Again: still mediated by understandability

Stop! All remaining effects are significant!

Trimmed model

	Estimate	S.E.	Est./S.E.	Two-Tailed P-Value
SATISF ON				
QUALITY	0.415	0.080	5.211	0.000
CONTROL	-0.883	0.119	-7.398	0.000
QUALITY ON				
CONTROL	-0.776	0.084	-9.198	0.000
CONTROL ON				
UNDERST	-0.397	0.071	-5.619	0.000
UNDERST ON				
CITEM	0.404	0.207	1.950	0.051
CFRIEND	0.588	0.204	2.878	0.004
CGRAPH	0.681	0.174	3.924	0.000

Trimmed model



BREN:ICS
INFORMATION AND COMPUTER SCIENCES

Modindices

```
ON/BY Statements

SATISF    ON UNDERST   /
UNDERST   BY SATISF          4.037      0.098       0.063        0.063
CONTROL   ON SATISF    /
SATISF    BY CONTROL         6.912      0.313       0.489        0.489
UNDERST   ON CONTROL   /
CONTROL   BY UNDERST        13.256      0.288       0.288        0.288

ON Statements

SATISF    ON CGRAPH          4.119      0.238       0.140        0.070
QUALITY   ON CFRIEND         6.691      0.301       0.230        0.108
QUALITY   ON CGRAPH          6.613      0.245       0.187        0.094
CONTROL   ON CGRAPH          9.164     -0.213      -0.196       -0.098
```

Some of these we removed earlier

For some of these we already have the alternate direction

Assess model fit

Item and factor fit should not have changed much

(please double-check!)

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 MPlus, change/add:

Under MODEL:

```
underst ON citem cfriend cgraph (p1-p3);
```

At the end:

```
MODEL TEST:
    p1=0;
    p2=0;
```

In R, change/add:

In model definition:

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

Then run:

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

Omnibus test

```
Wald Test of Parameter Constraints

          Value                                   8.516
          Degrees of Freedom                          2
          P-Value                                0.0142
```

Omnibus effect of control is significant

Total effects

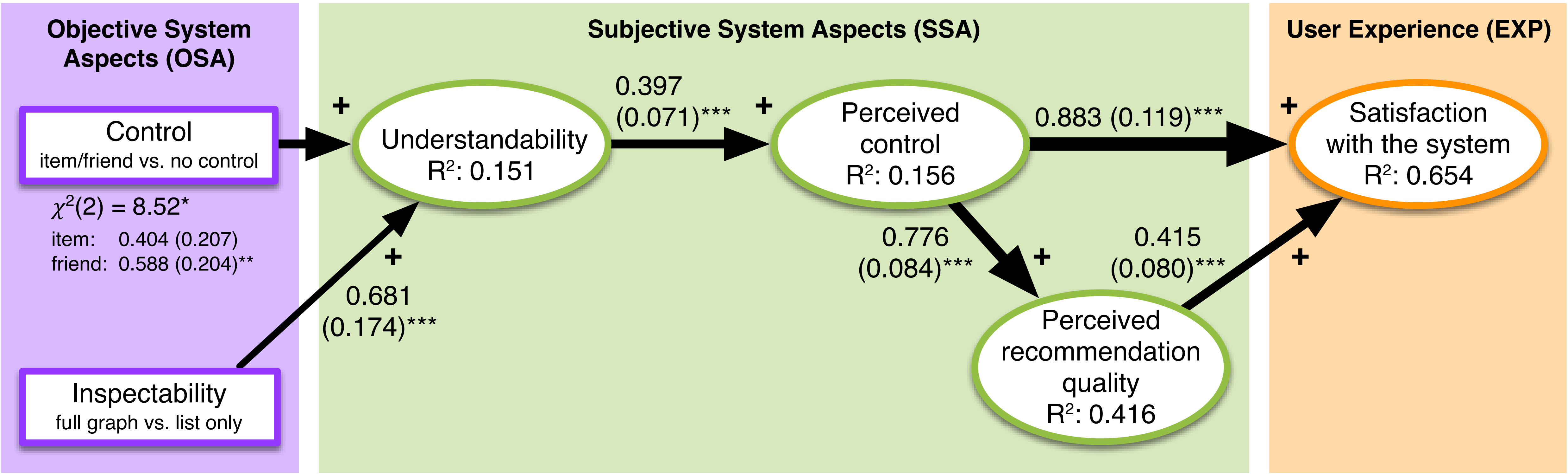
In MPlus:

```
MODEL INDIRECT:
 satisf IND citem;
 satisf IND cfriend;
 satisf IND cgraph;
 quality IND citem;
 quality IND cfriend;
 quality IND cgraph;
 control IND citem;
 control IND cfriend;
 control IND cgraph;
```

In R:

No automatic function for this; check out http://lavaan.ugent.be/tutorial/mediation.html

Final core model



BREN:ICS INFORMATION AND COMPUTER SCIENCES

Reporting

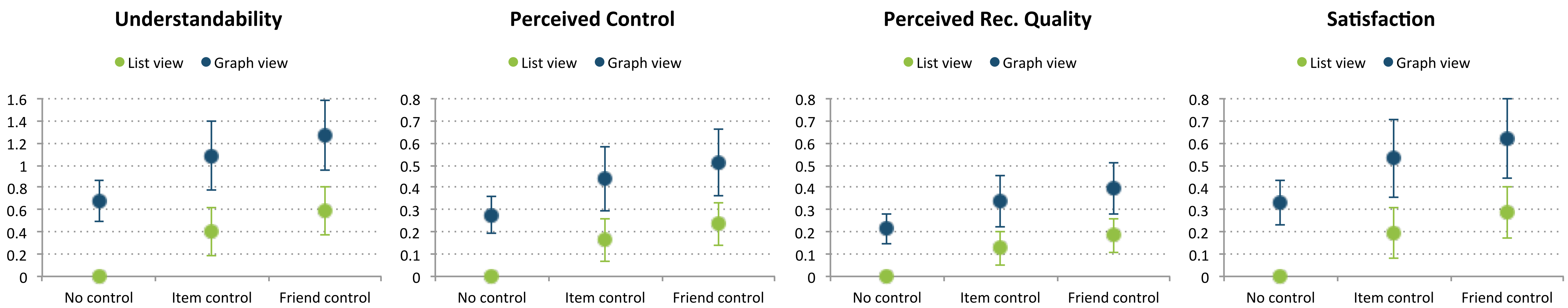
We subjected the 4 factors and the experimental conditions to structural equation modeling, which simultaneously fits the factor measurement model and the structural relations between factors and other variables. The model has a good* model fit: chi-square(223) = 306.685, p = .0002; RMSEA = 0.037, 90%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

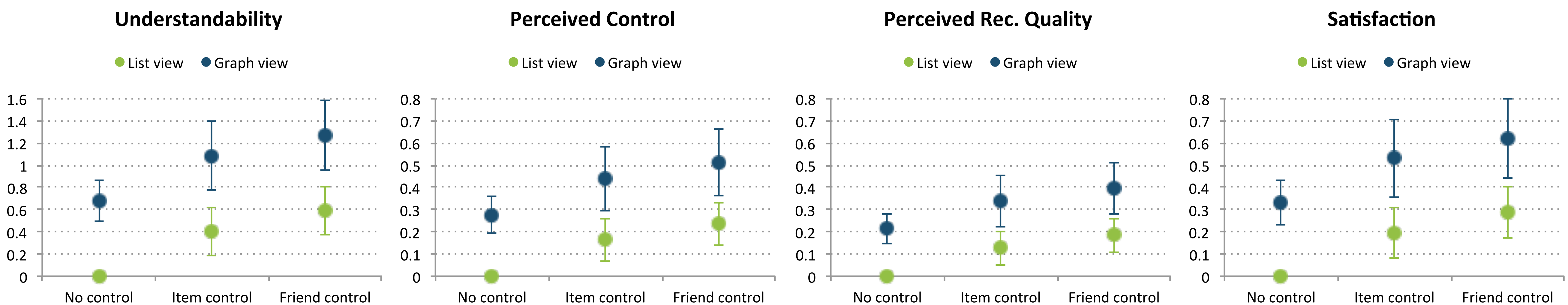
The model shows that the inspectability and control manipulations each have an independent positive effect on the understandability of the system: the full graph condition is more understandable than the list only condition, and the item control and friend control conditions are more understandable than 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.

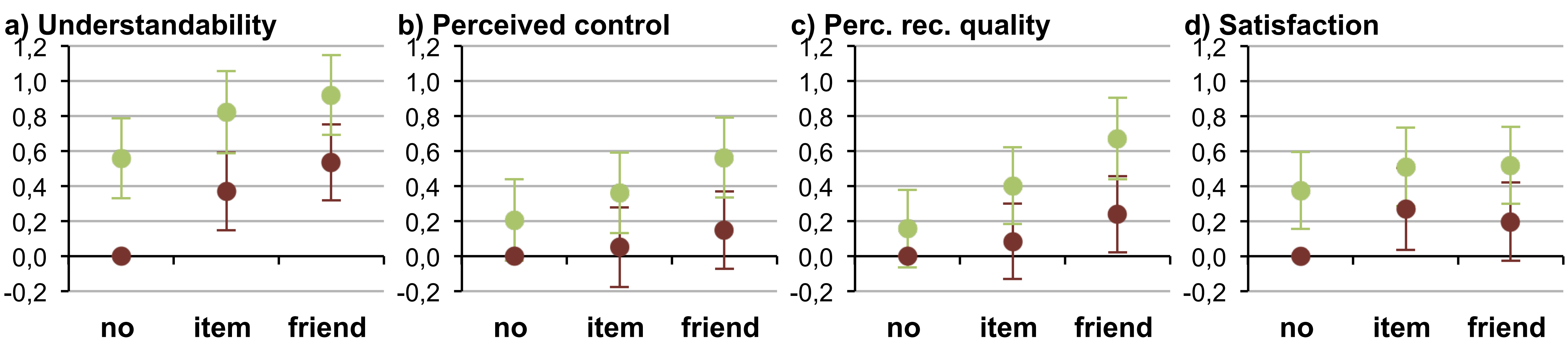
Total effect graphs



BREN:ICS INFORMATION AND COMPUTER SCIENCES

Why different?





BREN:ICS
INFORMATION AND COMPUTER SCIENCES

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"

Expand the model

Expanding the model by adding additional variables

- This is typically where behavior comes in

Redo model tests and additional stats

Expanded model

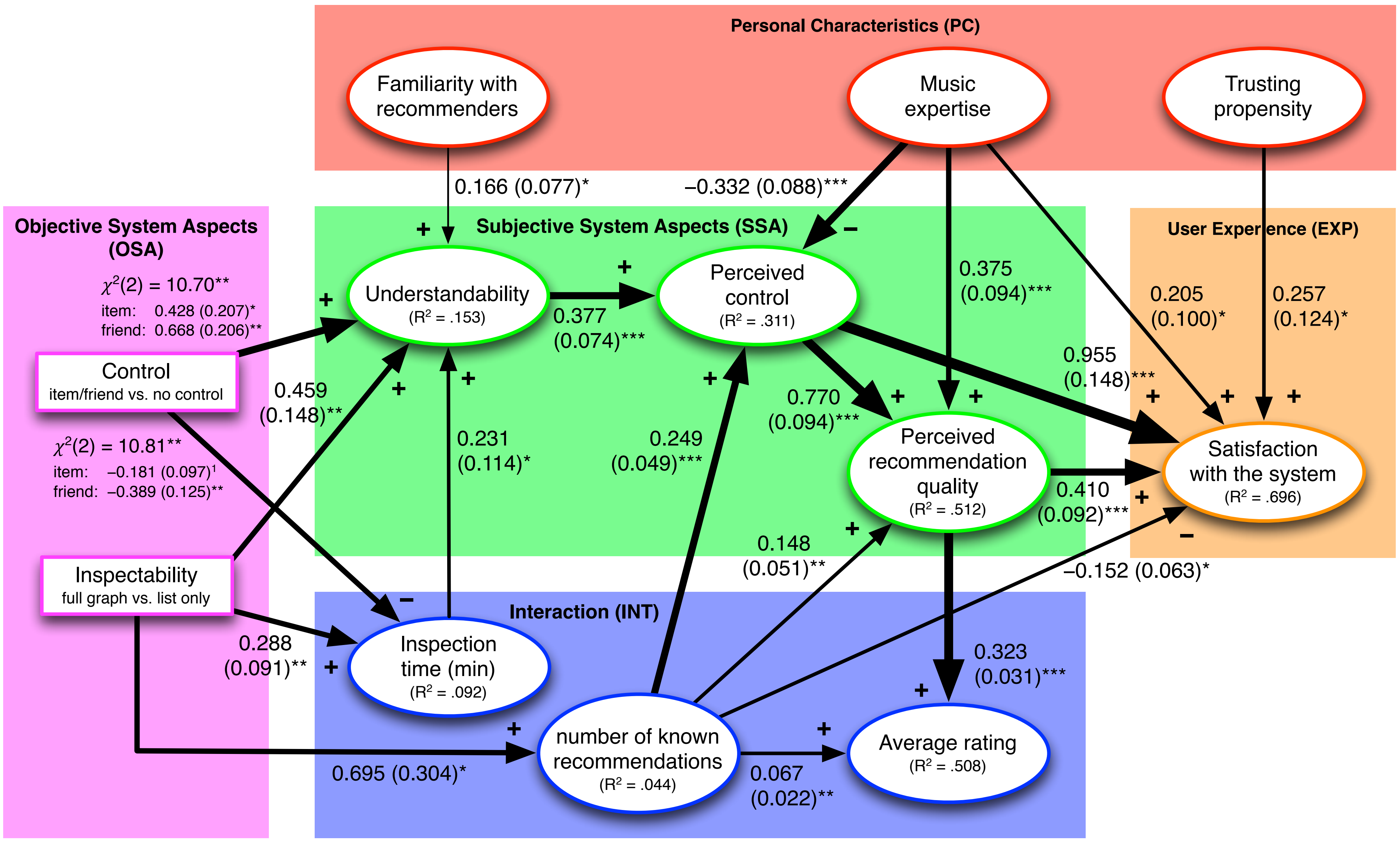


Figure 3. The structural equation model for the data of the experiment. Significance levels: *** $p < .001$, ** $p < .01$, 'ns' $p > .05$. R^2 is the proportion of variance explained by the model. Numbers on the arrows (and their thickness) represent the β coefficients (and standard error) of the effect.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

Learn more?

Learn it yourself:

- Rex Kline, “Principles and Practice of Structural Equation Modeling”, 3rd ed.
- MPlus: check the video tutorials at www.statmodel.com

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

T H A N K S !

George Bernard Shaw