Statistically Speaking

Power Analysis and Sample Size Determination Using Simulation

Sean P. Lane, Ph.D.

THE ANALYSIS FACTOR

Presentation Outline

- General Approaches to Power Analysis
  - The Basic Idea
  - Existing Formulas & Software
  - Limitations
- Power Analysis Using Simulation
  - The Process & Different Components
  - Comparing Results to Closed-Form Solutions
- Mechanics of Generating and Analyzing Simulated Power Analyses
  - A Practical Example
- Where Do You Get the Parameter Estimates & What if You're Wrong?
  - Strategies for Guessing, Buffering, & Optimizing Power

General Approaches to Power Analysis

An Old, but Unappreciated Idea

"The problem is that, as practiced, current research hardly reflects much attention to power….Last year, in *Psychological Bulletin,* Sedlmeier and Gigerenzer (1989) published an article entitled "Do Studies of Statistical Power Have an Effect on the Power of Studies?". The answer was no."

"But I do not despair. I remember that W.S. Gosset, the fellow who worked in a brewery and appeared in print modestly as "Student," published the *t* test a decade before we entered World War I, and the test didn't get into the psychological statistics textbooks until after World War II.

These things take time. So, if you publish something that you think is really good, and a year or a decade or two go by and hardly anyone seems to have taken notice, remember the *t* test, and take heart."

- Jacob Cohen, *American Psychologist* (1990)

THE ANALYSIS FACTOR

Sixth Edition

Publication
Manual
of the American Psychological Association

SP SP Society for Personality and Social Psychology

Home » Publications

Home

Best Practices

Journal of Personality and Social Psychology
2015, Vol. 108, No. 2, 275–297

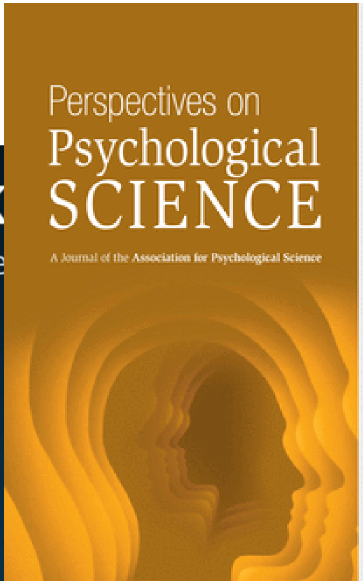
0022-3514/15/$12.00 http://dx.doi.org/10.1037/pspi0000007

Best Research Practices in Psychology: Illustrating Epistemological and Pragmatic Considerations With the Case of Relationship Science

Eli J. Finkel
Northwestern University

Paul W. Eastwick
University of Texas at Austin

Harry T. Reis
University of Rochester



VOLUME 9 NUMBER 3 MAY 2014

Open Science Framework

A scholarly commons to connect the entire research cycle

OPEN ACCESS Freely available online

PLOS | ONE

The N-Pact Factor: Evaluating the Quality of Empirical Journals with Respect to Sample Size and Statistical Power

R. Chris Fraley[1]*, Simine Vazire[2]*

1 Department of Psychology, University of Illinois at Urbana-Champaign, Champaign, Illinois, United States of America, 2 Department of Psychology, University of California Davis, Davis, California, United States of America

It's Time to Broaden the Replicability Conversation: Thoughts for and From Clinical Psychological Science

Perspectives on Psychological Science
2017, Vol. 12(5) 742–756

DOI: 10.1177/1745691617690042
www.psychologicalscience.org/PPS
SAGE

Jennifer L. Tackett[1], Scott O. Lilienfeld[2], Christopher J. Patrick[3], Sheri L. Johnson[4], Robert F. Krueger[5], Joshua D. Miller[6], Thomas F. Oltmanns[7], and Patrick E. Shrout[8]
[1]Northwestern University; [2]Emory University; [3]Florida State University; [4]University of California, Berkeley; [5]University of Minnesota; [6]University of Georgia; [7]Washington University in St. Louis; and [8]New York University

Psychonomic Society Guidelines on Statistical Issues

THE ANALYSIS FACTOR

What is Power?

- **Power:** The probability of rejecting a false null hypothesis
  - The inverse of the probability of a Type II Error (β)
  - $1 - \beta$

OR

  - The chance that your study will "work" given that your hypothesis is correct

Decision	H_0 True	H_0 False
Reject H_0	Type I Error $p = \alpha$	$\boxed{\text{Correct Decision } p = 1 - \beta}$ → **POWER**
Retain H_0	Correct Decision $p = 1 - \alpha$	Type II Error $p = \beta$

THE ANALYSIS FACTOR

Decreasing the Standard Error

$$Power = f(\alpha, \beta_1, N, \sigma_X^2, \sigma_e^2)$$

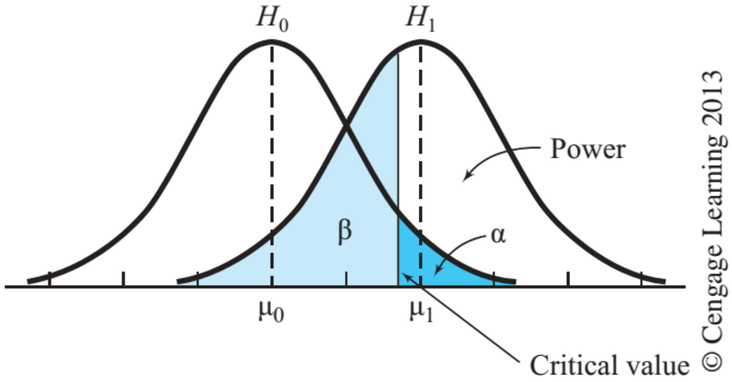


Figure 8.3 Effect on β of increasing $\mu_0 - \mu_1$

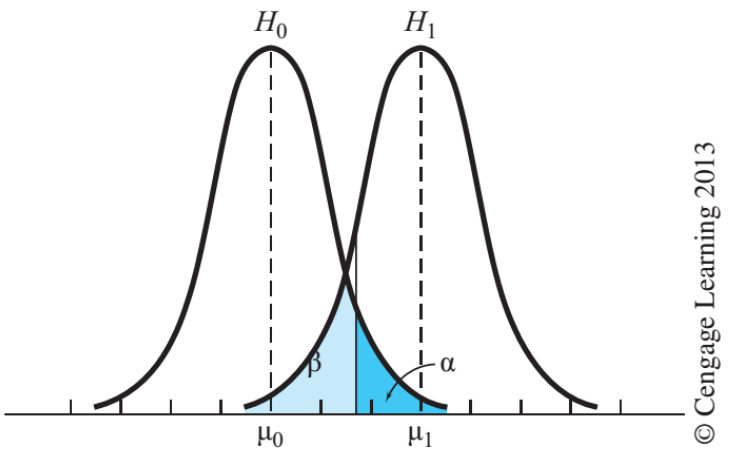


Figure 8.4 Effect on β of decrease in standard error of the mean

THE ANALYSIS FACTOR

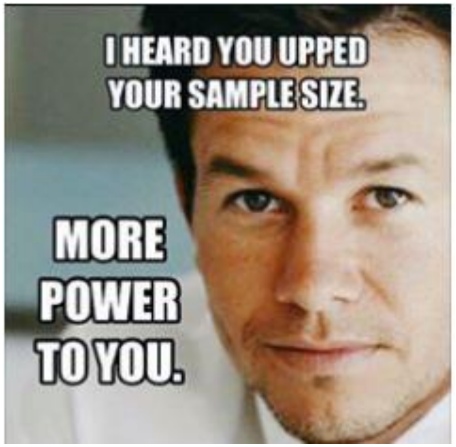
Effect Sizes vs. *p* Values

- **Effect Size:** Magnitude of effect (e.g., r, r^2, β, η^2, d)
  - Independent of sample size
  - Mean difference or association relative to variability
- ***p* Value:** Probability of obtaining an effect at least as large as what was actually observed if the null hypothesis is true
  - Highly dependent on sample size
  - Mean difference or association relative to the ratio of variability to sample size

GOAL: Identify minimum sample to have a high likelihood (e.g., 80%) of a low likelihood (e.g., 5%) of obtaining predicted effect if null is true

Why Should I Do a Power Analysis?

- Limited Resources
  - Studies are effortful, expensive, and time-consuming
  - Power analyses are free and less time consuming
- Efficiency
  - Increases chances of finding significant effects
  - Increases chances of replicating prior effects
  - Decreases null effects (file drawer)
  - Increase confidence in null effects



THE ANALYSIS FACTOR

Why Should I Do a Power Analysis?

- Grants require them
- Good for science
  - Protects from temptation to use researcher degrees of freedom
  - Increases precision of estimates – increasing study-to-study stability (replicability crisis)

THE ANALYSIS FACTOR

What Information Do I Need?

- Generally working within a regression framework
  - Although we can also power ANOVA designs using *M* and *SD*
- Effect sizes and additional sources of variability
  - Using standardized effects will often reduce the number of parameters that we need to determine
- Ex. Simple linear regression

$$Y_i = \beta_0 + \boxed{\beta_1} X_i + \varepsilon_i$$

Example: Calculating A Priori Power By Hand for Simple Linear Regression

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

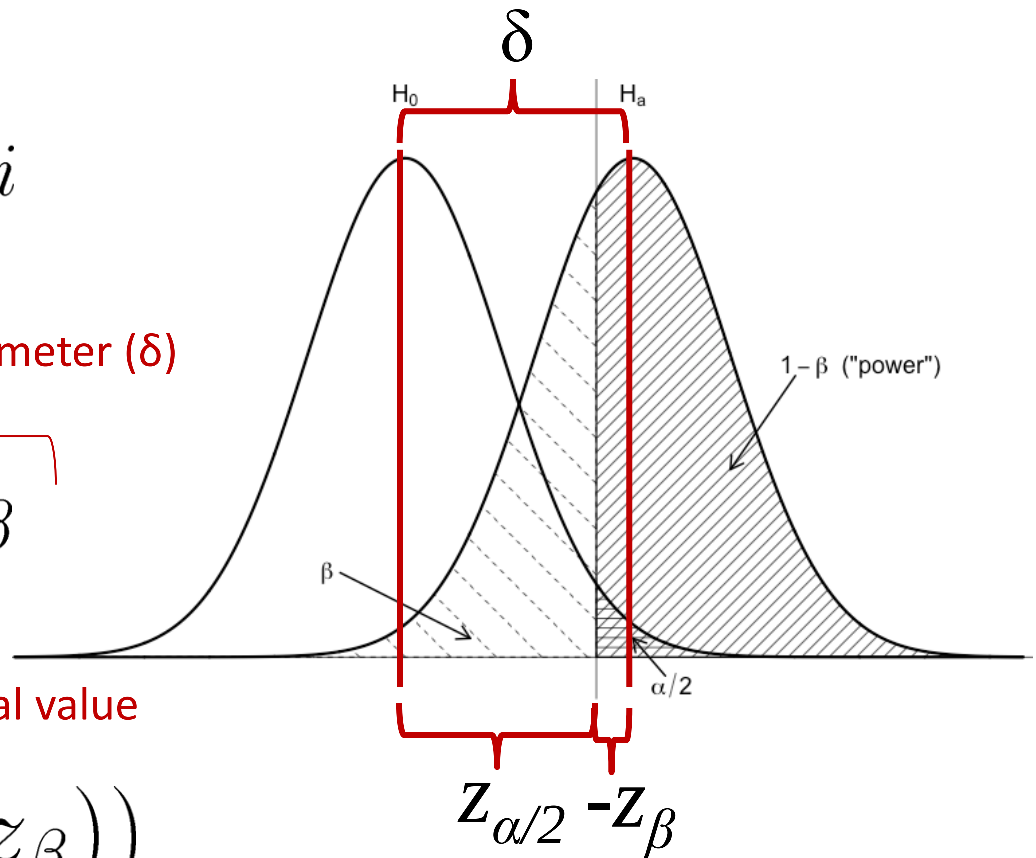
Standardized slope

1 – β²

$$\frac{\beta_1}{\sqrt{\frac{\sigma_\varepsilon^2}{N\sigma_X^2}}} = \overbrace{z_{\frac{\alpha}{2}} - z_\beta}^{\text{Noncentrality parameter }(\delta)}$$

Critical value　Critical value

$$Power = (1 - Pr(z_\beta))$$



Example: Calculating A Priori Power By Hand for Simple Linear Regression

Use $\beta = .15$

- Small effect
- Approximate association between height and weight in the population

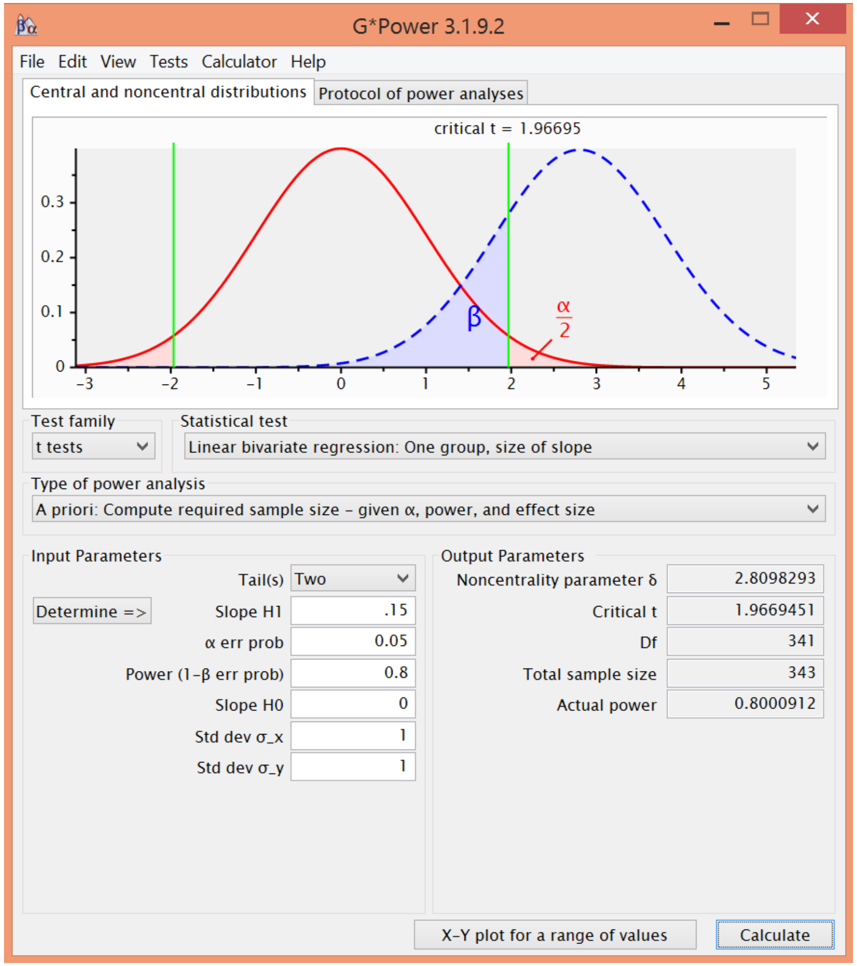
$$N = \frac{\sigma_\varepsilon^2 \left(z_{\frac{\alpha}{2}} - z_\beta\right)^2}{\beta^2 \sigma_X^2}$$

$$= \frac{(1-\beta^2)\left(z_{\frac{\alpha}{2}} - z_\beta\right)^2}{\beta^2 \sigma_X^2}$$

$$= \frac{(1-[.15]^2) \times (1.96 - [-.84])^2}{.15^2 \times 1^2} = 343$$

THE ANALYSIS FACTOR

Example: Calculating A Priori Power Using G*Power for Simple Linear Regression



t tests – Linear bivariate regression: One group, size of slope

Analysis:	A priori: Compute required sample size		
Input:	Tail(s)	=	Two
	Slope H1	=	0.1500000
	α err prob	=	0.05
	Power (1−β err prob)	=	0.8
	Slope H0	=	0
	Std dev σ_x	=	1
	Std dev σ_y	=	1
Output:	Noncentrality parameter δ	=	2.8098293
	Critical t	=	1.9669451
	Df	=	341
	Total sample size	=	343
	Actual power	=	0.8000912

Existing Power Software

Design	G*Power	Power & Precision	PS	PASS
t-tests	Yes	Yes	Yes	Yes
Correlation	Yes	Yes	Yes	Yes
ANOVA	Yes*	Yes*	Yes*	Yes*
Multiple regression	Yes*	Yes*	Yes*	Yes*
Multilevel regression	No	No	No	No
Path analysis	No	No	No	No
Latent variables	No	No	No	No

- Many models are not supported and even some basic models give limited or imperfect estimates

THE ANALYSIS FACTOR

Early Work

1990s: Closed form or iterative solutions for estimating standard errors and power for two-level models

- Approximation for estimating standard errors for two-level models (Snijders & Bosker, 1993; Liu & Liang, 1997)
- Accommodating CS, AR1, and TOEP residual covariance matrices (Hedeker, Gibbons, & Waternaux, 1999)
- Software for basic nested models [Optimal Design] (Raudenbush & Liu, 2000)

Early Work - Limitations

- Overly restrictive assumptions
- Approximations that had very limited generalizability
- Authors conducted analyses themselves and presented tables of sample sizes
  - Single papers often consisted of tables for only one type of model
  - Did not instruct researchers to conduct power analyses on their own
- Published in statistics and education journals

Early Work (Liu & Liang, 1997 example)

with $\mathbf{x}_{ij} = \mathbf{u}_{jl}$ and $\mathbf{z}_{ij} = \mathbf{v}_{jl}$. Similarly, Σ_1 reduces to

$$\Sigma_1 = m\mathrm{E}(\mathbf{P}^* V^{-1} \mathrm{cov}_{\mathrm{H}_1}(\mathbf{y}) V^{-1} \mathbf{P}^{*\prime}) = m \sum_{l=1}^{L} \pi_l \mathbf{P}_l^* V_l^{-1} \mathrm{cov}_{\mathrm{H}_1}(\mathbf{y}_l) V_l^{-1} \mathbf{P}_l^{*\prime}. \tag{10}$$

Let $\tilde{\xi} = \mathrm{E}[\mathbf{P}^* V^{-1}(\boldsymbol{\mu}^1 - \boldsymbol{\mu}^*)]$ and $\tilde{\Sigma}_1 = \mathrm{E}(\mathbf{P}^* V^{-1} \mathrm{cov}_{\mathrm{H}_1}(\mathbf{y}) V^{-1} \mathbf{P}^{*\prime})$, then the noncentral parameter ν in (4) can be re-expressed as

$$\nu = m\tilde{\xi}' \tilde{\Sigma}_1^{-1} \tilde{\xi},$$

and hence the sample size needed to achieve the nominal power is approximately

$$m = \nu / (\tilde{\xi}' \tilde{\Sigma}_1^{-1} \tilde{\xi}), \quad \leftarrow \text{OBVIOUSLY} \tag{11}$$

where ν is derived from a noncentral chi-square distribution and the given values of nominal power and significance level of the test.

Finally, equation (6), which leads to the solution of $\boldsymbol{\lambda}_0^*$, may now be expressed as

$$\sum_{l=1}^{L} \pi_l \left(\frac{\partial \tilde{\boldsymbol{\mu}}_l^*}{\partial \boldsymbol{\lambda}} \right)' V_l^{-1} (\tilde{\boldsymbol{\mu}}_l^1 - \tilde{\boldsymbol{\mu}}_l^*) = 0. \tag{12}$$

Note that $\tilde{\boldsymbol{\mu}}^1$, the expectation of $\mathbf{y}$, was computed given ψ_1 and $\boldsymbol{\lambda}_1$. This equation can be solved by the GEE method with the weights $\{\pi_l,\ l = 1, \ldots, L\}$.

In sample size calculations, all parameters in the models under the null and alternative hypotheses

Recent Work

2000s-: More flexible simulation approaches, additional software packages

Ahn, C., Heo, M., & Zhang, S. (2015). *Sample size calculations for clustered and longitudinal outcomes in clinical research.* Boca Raton, FL: CRC Press.

Bolger, N., & Laurenceau, J-P. (2013). *Intensive longitudinal methods: An Introduction to diary and experience sampling research*. New York: Guilford.

Cools, W., Van den Noortgate, W., & Onghena, P. (2008). ML-DEs: A program for designing efficient multilevel studies. *Behavior Research Methods, 40,* 236-249.

Gelman, A., & Hill, J. (2006). *Data analysis using regression and multilevel/hierarchical models*. Cambridge: Cambridge University Press.

Green, P., & MacLeod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*, 493-498.

Moerbeek, M., & Teerenstra, S. (2016). *Power analysis of trials with multilevel data.* Boca Raton, FL: CRC Press.

Zhang, Z., & Wang, L. (2009). Statistical power analysis for growth curve models using SAS. *Behavior Research Methods, 41,* 1083-1094.

THE ANALYSIS FACTOR

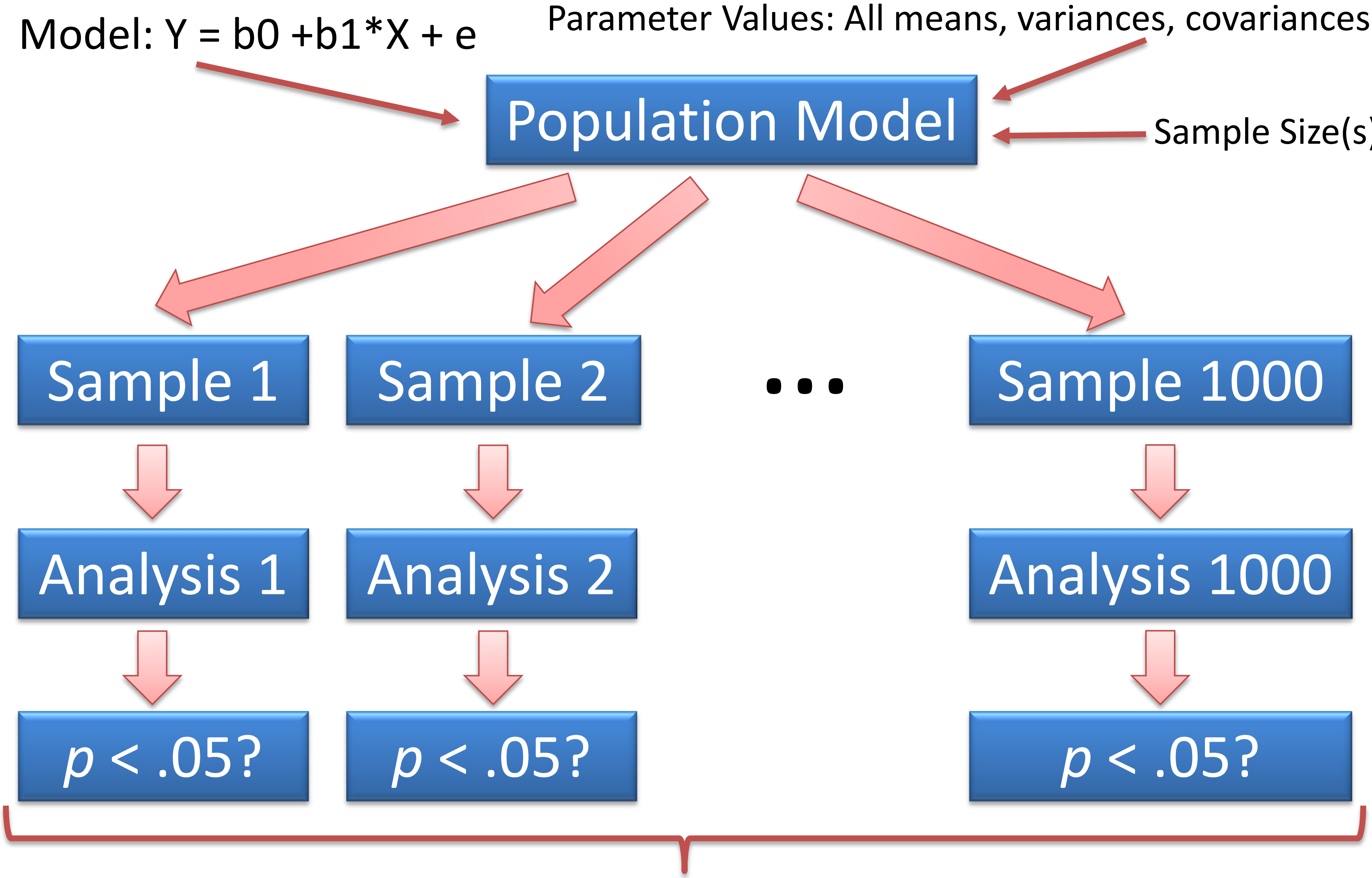
Recent Work - Limitations

Design	G*Power	Power & Precision	PS	PASS	RMASS	PinT	Optimal Design	ML-DEs	Simulation
t-tests	Yes	Yes	Yes	Yes	No	No	No	No	Yes
Correlation	Yes	Yes	Yes	Yes	No	No	No	No	Yes
ANOVA	Yes*	Yes*	Yes*	Yes*	No	No	No	No	Yes
Multiple regression	Yes*	Yes*	Yes*	Yes*	No	No	No	No	Yes
Multilevel regression	No	No	No	No	Yes*	Yes*	Yes*	Yes*	Yes
Path analysis	No	No	No	No	No	No	No	No	Yes
Latent variables	No	No	No	No	No	No	No	No	Yes

Power Analysis Using Simulation

What is Simulation?

1. Randomly generate data for a hypothetical study
   – Based on predefined model
2. Repeat 1000s of times to simulate 1000s of studies
3. Analyze each study
4. Record if hypothesized effect is significant
5. Aggregate results from all studies
6. % of significant results = Power



Power = % of significant effects

Simulate G*Power Example in SAS

```
DATA EXAMPLE1;
 RETAIN SUBNUM ;
 RETAIN SEED 20160129;
 RETAIN B0 1;
 RETAIN B1 .15;
 RETAIN SAMPLES 10000;
 DO SAMPLE=1 TO SAMPLES;
   DO SUBNUM=1 TO 343;
     X = RANNOR(SEED);
     Y = B0 + B1*X + SQRT(1-B1**2)*RANNOR(SEED);
     OUTPUT;
   END;
 END;
 DROP SEED b0 b1 SAMPLES;
 RUN;

 ODS GRAPHICS OFF; ODS SELECT NONE;

PROC reg DATA=EXAMPLE1;
 BY SAMPLE;
 MODEL Y = X / stb;
 ODS OUTPUT parameterestimates=fixedeffects;
 RUN;

 ODS SELECT ALL; ODS GRAPHICS ON; ODS LISTING;

data Xvar; set fixedeffects; if variable='X'; if probt le .05 & estimate gt 0 then sig=1; else sig=0; run;
proc means data=Xvar; var estimate standardizedest sig; run;
```

Generate the Data

```
DATA EXAMPLE1;              ← Declare the dataset we want to create
RETAIN SUBNUM ;             ← Create variable for subject id
RETAIN SEED 20160129;       ← Choose value for random number generator
RETAIN B0 1;                ← Intercept value
RETAIN B1 .15;              ← Slope value
RETAIN SAMPLES 10000;       ← Number of studies to simulate
DO SAMPLE=1 TO SAMPLES;     ← Loops to create data for studies and
  DO SUBNUM=1 TO 343;         subjects within studies
    X = RANNOR(SEED);       ← Random predictor value
    Y = B0 + B1*X + SQRT(1-B1**2)*RANNOR(SEED);   ← The model
    OUTPUT;                 ← Output generated data to dataset
  END;
END;
DROP SEED b0 b1 SAMPLES;
RUN;
```

Define parameters

Simulate studies

THE ANALYSIS FACTOR

Analyze the Data

```
ODS GRAPHICS OFF; ODS SELECT NONE;
```
← Turn off output

The analysis

```
PROC reg DATA=EXAMPLE1;
BY SAMPLE;
MODEL Y = X / stb;
ODS OUTPUT parameterestimates=fixedeffects;
RUN;
```

← Run a regression using generated data

← Analyze each sample separately

← The model

Save model parameters

```
ODS SELECT ALL; ODS GRAPHICS ON; ODS LISTING;
```

Turn output back on

Dataset of Estimates from Each Sample

```
data Xvar; set fixedeffects; if variable='X'; if probt le .05 & estimate gt 0 then sig=1; else sig=0; run;
```

VIEWTABLE: Work.Fixedeffects (Parameter Estimates)

	SAMPLE	Model	Dependent	Variable	DF	Estimate	StdErr	tValue	Probt	StandardizedEst
1	1	MODEL1	Y	Intercept	1	0.97384	0.05459	17.84	<.0001	0
2	1	MODEL1	Y	X	1	0.14345	0.05340	2.69	0.0076	0.14396
3	2	MODEL1	Y	Intercept	1	1.01207	0.05421	18.67	<.0001	0
4	2	MODEL1	Y	X	1	0.12749	0.05422	2.35	0.0193	0.12632
5	3	MODEL1	Y	Intercept	1	0.94712	0.05429	17.45	<.0001	0
6	3	MODEL1	Y	X	1	0.13906	0.05292	2.63	0.0090	0.14089
7	4	MODEL1	Y	Intercept	1	1.03225	0.05011	20.60	<.0001	0
8	4	MODEL1	Y	X	1	0.06480	0.04869	1.33	0.1841	0.07189

A Little Data Cleaning

```
data Xvar; set fixedeffects; if variable='X'; if probt le .05 & estimate gt 0 then sig=1; else sig=0; run;
```

p-value significant?

VIEWTABLE: Work.Xvar

	SAMPLE	Model	Dependent	Variable	DF	Estimate	StdErr	tValue	Probt	StandardizedEst	sig
1	1	MODEL1	Y	X	1	0.14345	0.05340	2.69	0.0076	0.14396	1
2	2	MODEL1	Y	X	1	0.12749	0.05422	2.35	0.0193	0.12632	1
3	3	MODEL1	Y	X	1	0.13906	0.05292	2.63	0.0090	0.14089	1
4	4	MODEL1	Y	X	1	0.06480	0.04869	1.33	0.1841	0.07189	0

Estimate Power

```
proc means data=Xvar; var estimate standardizedest sig; run;
```

The MEANS Procedure

Variable	Label	N	Mean	Std Dev	Minimum	Maximum
Estimate	Parameter Estimate	10000	0.1504423	0.0539107	-0.0443288	0.3694533
StandardizedEst	Standardized Estimate	10000	0.1504054	0.0533023	-0.0449166	0.3515363
sig		10000	0.8006000	0.3995693	0	1.0000000

"Same" as G*Power – (Sampling variability)

THE ANALYSIS FACTOR

Everything We are Doing Can Be Done in SPSS, R, Mplus, etc.

SPSS

Generate Data

```
INPUT PROGRAM.
LOOP #s=1 to 10000.
LOOP #i=1 to 343.

COMPUTE sample=#s.
COMPUTE id=#i.
COMPUTE x=RV.NORMAL(0,1).
COMPUTE y=1+.15*x+SQRT(1-.15**2)*RV.NORMAL(0,1).
END CASE.

END LOOP.
END LOOP.
END FILE.
END INPUT PROGRAM.
EXECUTE.
```

Analyze Data

```
SPLIT FILE SEPARATE BY sample.

DATASET DECLARE  estimates.
REGRESSION
  /MISSING LISTWISE
  /STATISTICS COEFF OUTS R ANOVA
  /CRITERIA=PIN(.05) POUT(.10)
  /NOORIGIN
  /DEPENDENT y
  /METHOD=ENTER x
  /OUTFILE=COVB(estimates).

DATASET ACTIVATE estimates.
FILTER OFF.
USE ALL.
SELECT IF (ROWTYPE_="EST"  |  ROWTYPE_="SIG").
EXECUTE.

SORT CASES BY sample ROWTYPE_.
CASESTOVARS
  /ID=sample
  /INDEX=ROWTYPE_
  /GROUPBY=VARIABLE.

COMPUTE X_SIG=0.
IF X.EST>0 & X.SIG<=.05 X_SIG=1.
EXECUTE.

DESCRIPTIVES VARIABLES=X.EST X_SIG
  /STATISTICS=MEAN STDDEV MIN MAX.
```

Everything We are Doing Can Be Done in SPSS, R, Mplus, etc.

SPSS

Power Results

Descriptives

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
x.EST	10000	-.03	.32	.1505	.05232
X_SIG	10000	.00	1.00	.8030	.39775
Valid N (listwise)	10000				

- It runs, but it took over an hour (SAS < 1 minute)
- Clunkier syntax
- Difficult to suppress output when using included procedures

THE ANALYSIS FACTOR

Everything We are Doing Can Be Done in SPSS, R, Mplus, etc.

Mplus

Generate & Analyze Data

```
MONTECARLO:
        NAMES ARE x y;
        NOBSERVATIONS = 343;
        SEED = 20160129;
        NREPS = 10000;
ANALYSIS:
        MODEL POPULATION:
        y ON x*.15;
        [x*0];
        [y*1];
        x*1;
        y*1;
MODEL:
        y ON x*.15;
        [x*0];
        [y*1];
        x*1;
        y*1;
OUTPUT: SAMPSTAT TECH9;
```

THE ANALYSIS FACTOR

Everything We are Doing Can Be Done in SPSS, R, Mplus, etc.

Mplus

Power Results

MODEL RESULTS

	Population	ESTIMATES Average	Std. Dev.	S. E. Average	M. S. E.	95% Cover	% Sig Coeff
Y ON							
X	0.150	0.1500	0.0542	0.0540	0.0029	0.950	0.793
Means							
X	0.000	-0.0005	0.0538	0.0539	0.0029	0.948	0.052
Intercepts							
Y	1.000	0.9997	0.0540	0.0539	0.0029	0.946	1.000
Variances							
X	1.000	0.9966	0.0756	0.0761	0.0057	0.945	1.000
Residual Variances							
Y	1.000	0.9945	0.0759	0.0759	0.0058	0.943	1.000

- Runs quickly
- Easily specify wide variety of models
- Can't loop multiple sample sizes without batch jobs (re: later)

THE ANALYSIS FACTOR

Everything We are Doing Can Be Done in SPSS, R, Mplus, etc.

R

Generate & Analyze Data

```r
ex1.sim <- function (J){
  person <- rep(1:J)
  x <- rnorm(J,0,1)
  b0 <- 1
  b1 <- .15
  y <- rnorm(J, b0+b1*x,(1-b1*b1))
  return(data.frame(person,y,x))
}

ex1.power <- function (J,n.sims=10000){
  signif <- rep(NA,n.sims)
  estim <- rep(NA,n.sims)
  pval <- rep(NA,n.sims)
  for (s in 1:n.sims){
    sim <- ex1.sim(J)
    lm.power <- lm(y ~ x, data=sim)
    est <- coef(summary(lm.power))["x","Estimate"]
    p <- coef(summary(lm.power))["x","Pr(>|t|)"]
    se <- coef(summary(lm.power))["x","Std. Error"]
    signif[s] <- est>0 & p<=.05
  }
  power <- mean(signif)
  return(power)
}

ex1.power(J=343,n.sims=10000)
```

Power Results

```
>
> ex1.power(J=343,n.sims=10000)
[1] 0.8061
>
```

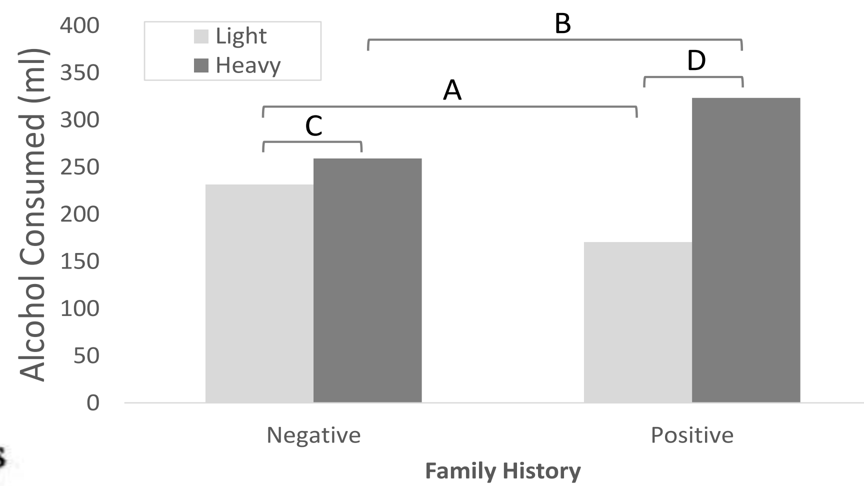
- Runs quickly
- Easily specify wide variety of models
- Can loop multiple sample sizes or effect sizes

THE ANALYSIS FACTOR

Mechanics of Generating and Analyzing a Simulated Power Analysis

2x2 Between-Subjects ANOVA Example (SAS)

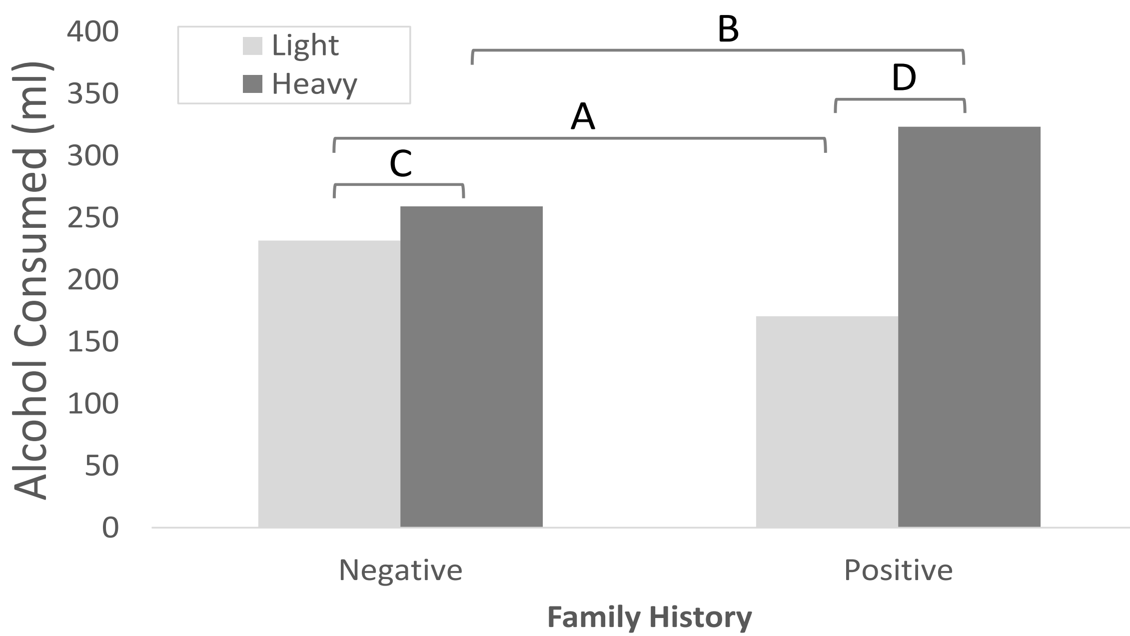
- Replicating Chipperfield & Vogel-Sprott (1988): *Family History of Problem Drinking among Young Male Social Drinkers: Modeling Effects on Alcohol Consumption*
  - Individuals model the drinking behavior of others
  - This effect is amplified among people who have a family history of alcoholism
  - Original study *N* = 50



(FH) × 2 (model) ANOVA of the milliliter scores. This analysis obtained a significant FH × Model interaction, $F(1, 46) = 4.10$, $p < 0.05$, and main effect of model, $F(1, 46) = 8.48$, $p < 0.01$. The FH groups did not differ in response to the model treatments, $F(1, 46) = 0.003$, $p > 0.99$. Figure 1, which plots the

2x2 Between-Subjects ANOVA Example (SAS)

```
DATA EXAMPLE1;                                                           /*1*/
 RETAIN SEED 20160129;                                                   /*2*/
 RETAIN B0 231.7; RETAIN B1 -61.1; RETAIN B2 27.5; RETAIN B3 125.4;      /*3*/
 RETAIN SAMPLES 10000;                                                   /*4*/
 ARRAY SIZE[3] (50 80 800);                                              /*5*/
 DO SSIZE=1 TO 3;                                                        /*6*/
   DO SAMPLE=1 TO SAMPLES;                                               /*7*/
     DO SUBNUM=1 TO SIZE[SSIZE];                                         /*8*/
       IF SUBNUM le SIZE[SSIZE]/2 THEN X1 = 0; ELSE X1=1;                /*9*/
       IF SUBNUM le SIZE[SSIZE]/4 THEN X2 = 0;                           /*10*/
       ELSE IF (SUBNUM gt SIZE[SSIZE]/2 & SUBNUM le SIZE[SSIZE]*(3/4))   /*11*/
                           THEN X2=0; ELSE X2=1;                         /*12*/
       Y = B0 + B1*X1 + B2*X2 + B3*X1*X2 + SQRT(10300)*RANNOR(SEED);     /*13*/
       OUTPUT;                                                           /*14*/
 END; END; END;                                                          /*15*/
 DROP SEED b0 b1 b2 b3 SAMPLES; RUN;                                     /*16*/
DATA EXAMPLE1; SET EXAMPLE1;                                             /*17*/
 if SSIZE=1 then SSIZE=size1; else if SSIZE=2 then SSIZE=size2;          /*18*/
 else SSIZE=size3;                                                       /*19*/
 DROP size1 size2 size3; RUN;                                            /*20*/
                                                                         /*21*/
 ODS GRAPHICS OFF; ODS SELECT NONE;                                      /*22*/
PROC MIXED data=EXAMPLE1 noclprint method=type3; by ssize sample;        /*23*/
 class x1 x2;                                                            /*24*/
 model y = x1 x2 x1*x2 / s;                                              /*25*/
 lsmeans x1*x2 / diff;                                                   /*26*/
 ods output tests3=anova diffs=contrasts; RUN;                           /*27*/
 ODS SELECT ALL; ODS GRAPHICS ON; ODS LISTING;                           /*28*/
                                                                         /*29*/
data anova; set anova; if probf le .05 then sig=1; else sig=0; RUN;      /*30*/
proc sort data=anova; by ssize sample; RUN;                              /*31*/
proc transpose data=anova out=anova2;                                    /*32*/
 by ssize sample; id effect; var sig; RUN;                               /*33*/
proc means data=anova2; by ssize; var x1 x2 x1_x2; RUN;                  /*34*/
                                                                         /*35*/
data contrasts; set contrasts;                                           /*36*/
 if x1=0 & x2=0 & _x1=0 & _x2=1 then simple='C';                         /*37*/
 if x1=0 & x2=0 & _x1=1 & _x2=0 then simple='A';                         /*38*/
 if x1=0 & x2=1 & _x1=1 & _x2=1 then simple='B';                         /*39*/
 if x1=1 & x2=0 & _x1=1 & _x2=1 then simple='D';                         /*40*/
 if missing(simple)=0;                                                   /*41*/
 if probt le .05 & estimate lt 0 & (simple~='A') then sig=1;             /*42*/
 else if probt le .05 & estimate gt 0 & (simple='A')                     /*43*/
 then sig=1; else sig=0; RUN;                                            /*44*/
                                                                         /*45*/
proc sort data=contrasts; by ssize sample simple; RUN;                   /*46*/
proc transpose data=contrasts out=contrasts2;                            /*47*/
 by ssize sample; id simple; var sig; RUN;                               /*48*/
proc means data=contrasts2; by ssize; var a b c d; RUN;                  /*49*/
```



Simulate the Data

Regression coefficients for main effects and interaction

```
DATA EXAMPLE1;
RETAIN SEED 20160129;
RETAIN B0 231.7; RETAIN B1 -61.1; RETAIN B2 27.5; RETAIN B3 125.4;
RETAIN SAMPLES 10000;
ARRAY SIZE[3] (50 80 800);
DO SSIZE=1 TO 3;
  DO SAMPLE=1 TO SAMPLES;
    DO SUBNUM=1 TO SIZE[SSIZE];
      IF SUBNUM le SIZE[SSIZE]/2 THEN X1 = 0; ELSE X1=1;
      IF SUBNUM le SIZE[SSIZE]/4 THEN X2 = 0;
      ELSE IF (SUBNUM gt SIZE[SSIZE]/2 & SUBNUM le SIZE[SSIZE]*(3/4))
                                  THEN X2=0; ELSE X2=1;
      Y = B0 + B1*X1 + B2*X2 + B3*X1*X2 + SQRT(10300)*RANNOR(SEED);
      OUTPUT;
END; END; END;
DROP SEED b0 b1 b2 b3 SAMPLES; RUN;
```

← **NEW** – Runs power analyses for 3 different sample sizes

← **NEW** – Tells loop to reference the array to determine sample size

← **NEW** – Assigns subjects to condition

Error variance

```
DATA EXAMPLE1; SET EXAMPLE1;
if SSIZE=1 then SSIZE=size1; else if SSIZE=2 then SSIZE=size2;
else SSIZE=size3;
DROP size1 size2 size3; RUN;
```

Clean up dataset so 'ssize' corresponds to sample size

THE ANALYSIS FACTOR

Analyze the Simulated Data

The analysis

```
ODS GRAPHICS OFF; ODS SELECT NONE;
PROC MIXED data=EXAMPLE1 noclprint method=type3; by ssize sample;
class x1 x2;
model y = x1 x2 x1*x2 / s;
lsmeans x1*x2 / diff;
ods output tests3=anova diffs=contrasts; RUN;
ODS SELECT ALL; ODS GRAPHICS ON; ODS LISTING;
```

← Save effects

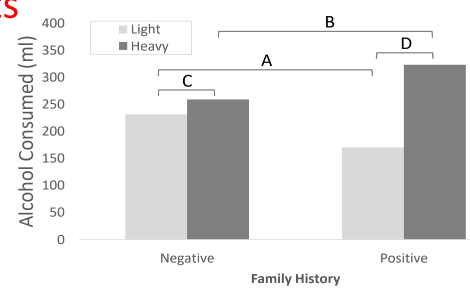
Main Effects & Interaction

```
data anova; set anova; if probf le .05 then sig=1; else sig=0; RUN;
proc sort data=anova; by ssize sample; RUN;
proc transpose data=anova out=anova2;
by ssize sample; id effect; var sig; RUN;
proc means data=anova2; by ssize; var x1 x2 x1_x2; RUN;
```

Simple Effects

```
data contrasts; set contrasts;
if x1=0 & x2=0 & _x1=0 & _x2=1 then simple='C';
if x1=0 & x2=0 & _x1=1 & _x2=0 then simple='A';
if x1=0 & x2=1 & _x1=1 & _x2=1 then simple='B';
if x1=1 & x2=0 & _x1=1 & _x2=1 then simple='D';
if missing(simple)=0;
if probt le .05 & estimate lt 0 & (simple~='A') then sig=1;
else if probt le .05 & estimate gt 0 & (simple='A')
then sig=1; else sig=0; RUN;

proc sort data=contrasts; by ssize sample simple; RUN;
proc transpose data=contrasts out=contrasts2;
by ssize sample; id simple; var sig; RUN;
proc means data=contrasts2; by ssize; var a b c d; RUN;
```



Estimate Power – (Family History)

--------------------------------- SSIZE=50 ---------------------------------

The MEANS Procedure

Variable	N	Mean	Std Dev	Minimum	Maximum
X1	10000	0.0520000	0.2220381	0	1.0000000
X2	10000	0.8661000	0.3405619	0	1.0000000
X1_X2	10000	0.5646000	0.4958341	0	1.0000000

--------------------------------- SSIZE=80 ---------------------------------

Variable	N	Mean	Std Dev	Minimum	Maximum
X1	10000	0.0486000	0.2150411	0	1.0000000
X2	10000	0.9769000	0.1502286	0	1.0000000
X1_X2	10000	0.7827000	0.4124290	0	1.0000000

--------------------------------- SSIZE=800 ---------------------------------

Variable	N	Mean	Std Dev	Minimum	Maximum
X1	10000	0.0556000	0.2291591	0	1.0000000
X2	10000	1.0000000	0	1.0000000	1.0000000
X1_X2	10000	1.0000000	0	1.0000000	1.0000000

THE ANALYSIS FACTOR

Estimate Power – (Modeling)

------------------------------------- SSIZE=50 -------------------------------------

The MEANS Procedure

Variable	N	Mean	Std Dev	Minimum	Maximum
X1	10000	0.0520000	0.2220381	0	1.0000000
X2	10000	0.8661000	0.3405619	0	1.0000000
X1_X2	10000	0.5646000	0.4958341	0	1.0000000

------------------------------------- SSIZE=80 -------------------------------------

Variable	N	Mean	Std Dev	Minimum	Maximum
X1	10000	0.0486000	0.2150411	0	1.0000000
X2	10000	0.9769000	0.1502286	0	1.0000000
X1_X2	10000	0.7827000	0.4124290	0	1.0000000

------------------------------------- SSIZE=800 -------------------------------------

Variable	N	Mean	Std Dev	Minimum	Maximum
X1	10000	0.0556000	0.2291591	0	1.0000000
X2	10000	1.0000000	0	1.0000000	1.0000000
X1_X2	10000	1.0000000	0	1.0000000	1.0000000

Estimate Power – (Interaction)

--------------------------------- SSIZE=50 ---------------------------------

The MEANS Procedure

Variable	N	Mean	Std Dev	Minimum	Maximum
X1	10000	0.0520000	0.2220381	0	1.0000000
X2	10000	0.8661000	0.3405619	0	1.0000000
X1_X2	10000	$\boxed{0.5646000}$	0.4958341	0	1.0000000

--------------------------------- SSIZE=80 ---------------------------------

Variable	N	Mean	Std Dev	Minimum	Maximum
X1	10000	0.0486000	0.2150411	0	1.0000000
X2	10000	0.9769000	0.1502286	0	1.0000000
X1_X2	10000	$\boxed{0.7827000}$	0.4124290	0	1.0000000

--------------------------------- SSIZE=800 ---------------------------------

Variable	N	Mean	Std Dev	Minimum	Maximum
X1	10000	0.0556000	0.2291591	0	1.0000000
X2	10000	1.0000000	0	1.0000000	1.0000000
X1_X2	10000	$\boxed{1.0000000}$	0	1.0000000	1.0000000

THE ANALYSIS FACTOR

Simple Effects

------------------------------------ SSIZE=50 -----

The MEANS Procedure

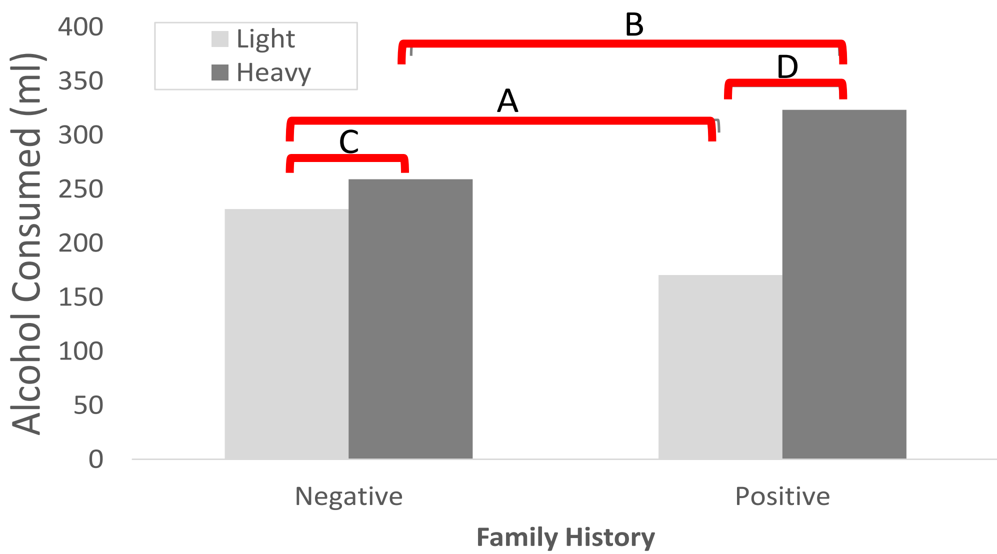
Variable	N	Mean	Std Dev
A	10000	0.3056000	0.4606841
B	10000	0.3461000	0.4757482
C	10000	0.1050000	0.3065687
D	10000	0.9560000	0.2051054

------------------------------------ SSIZE=80 -----

Variable	N	Mean	Std Dev
A	10000	0.4698000	0.4991121
B	10000	0.5082000	0.4999578
C	10000	0.1325000	0.3390505
D	10000	0.9966000	0.0582132

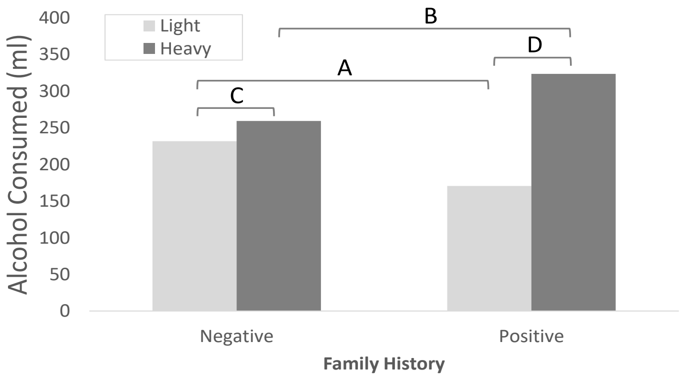
------------------------------------ SSIZE=800 -----

Variable	N	Mean	Std Dev
A	10000	0.9999000	0.0100000
B	10000	1.0000000	0
C	10000	0.7753000	0.4174055
D	10000	1.0000000	0



Power of Omnibus and Simple Effects as a Function of *N*

Effect	*N* = 80	*N* = 160	*N* = 800
Family History Main Effect	.05	.05	.06
Modeling Main Effect	.98	.99	1.00
Interaction	**.79**	**.97**	**1.00**
Simple Effect A	**.47**	**.76**	**1.00**
Simple Effect B	**.51**	**.80**	**1.00**
Simple Effect C	*.13*	*.22*	*.78*
Simple Effect D	1.00	1.00	1.00



Where Do You Get the Parameter Estimates and What if You're Wrong?

Determining Parameters

- Published similar data
  - Common not to report all statistics (e.g., random effects, residuals)
    - May need to contact authors
- Published semi-similar data
  - Similar variables
    - e.g., using a study of change in racial bias post-intervention to estimate parameters for change in sexism post-intervention
  - Similar design
    - e.g., using a baseline|intervention|follow-up|follow-up study of racial bias to estimate parameters for a baseline|intervention|follow-up|followup study of self-esteem
  - Between-person estimates of within-person processes
    - e.g., using an intervention|control between-subjects lab study of racial bias to estimate parameters for a baseline|intervention within-subjects study of racial bias

Educated Guesses

- Fixed Effects
  - Interaction effects usually smaller than main effects
  - Indirect effects (multiplicatively) smaller than direct effects
  - Between-person effects usually larger than within-person effects
  - Self-report (Actor) effects usually larger than other-report (Partner) effects
- Random Effects
  - They are usually there (i.e. non-zero)
  - Random effect variances usually smaller than residual variance
  - Higher cluster variance (in intercept and slopes) has a stronger impact on power when there are more clusters
- Residuals
  - Usually autocorrelated in longitudinal designs
  - Usually increase over time in longitudinal designs

Educated Guesses

- Use possible % of variance explainable as a reference for estimating effect sizes and error variance
  - Specific to area of research
    - Daily life – lots of within-person variability
      - Upper bound of explainable variance ~10%
      - Largely random effects
      - Individual fixed effects < 1%
    - Cognitive experiment – might expect to explain > 50% of the total variance
- Know your area and what is reasonable – use that as a reference

Sensitivity Analysis

- How bad can things be?
  - Even informed guesses are still guesses
- Maximize likelihood of being well-powered even if you guessed incorrectly
- Systematically vary parameters of interest & recalculate power
  - Efficiently balance # people and # repeated measurements
  - Robustness of hypothesized effects
- Start with most theoretically important parameter
  - How small can it get?
- Most uncertain parameter
    - How small/big can it get?
    - Usually random effects (if clustered data)

THE ANALYSIS FACTOR

Troubleshooting

- How do I know I did what I was intending to?
  1. Check means and variances of generated data and population parameters
  2. Check that the model recovers correct coefficients
     - Assuming, generating model = fit model

Example 1 Revisited

```
DATA EXAMPLE1;
RETAIN SUBNUM ;
RETAIN SEED 20160129;
RETAIN B0 1;
RETAIN B1 .15;
RETAIN SAMPLES 10000;
DO SAMPLE=1 TO SAMPLES;
  DO SUBNUM=1 TO 343;
    X = RANNOR(SEED);
    Y = B0 + B1*X + SQRT(.5)*RANNOR(SEED);
    OUTPUT;
  END;
END;
DROP SEED b0 b1 SAMPLES;
RUN;
```

Correct variance was (1-B1**2)

Variable	Label	N	Mean	Std Dev
Estimate	Parameter Estimate	10000	0.1503163	0.0385569
StandardizedEst	Standardized Estimate	10000	0.2078744	0.0522384
sig		10000	0.9725000	0.1635433

OH NO! Estimates don't match.

Conclusions

- Power analysis helps us design studies with high likelihoods of observing effects [big OR small] with a high degree of precision
- Cost-, resource-, and time-effective
- Classic power analysis tools generally not suitable for estimating factorial, multilevel, path, and latent factor models
- Simulation allows for flexible sample size estimates of any model using SPSS, SAS, MPlus, or R
  - If you can run the analysis, you can conduct the power analysis for it

Conclusions

- Requires a priori hypotheses with educated guesses about effect sizes and variability
  - Including random effects and residuals
- Similar published papers (even on very different topics) and pilot data can help estimate these parameters
- Sensitivity analysis helps make good decisions when there is uncertainty about the parameters
- Can be integrated into preregistration activities
- Can be adapted for Bayesian or other alternative hypothesis testing techniques

For More Information:

- Lane, S. P., & Hennes, E. P. (2018). Power struggles: Estimating sample size for multilevel relationship research. *Journal of Social and Personal Relationships, 35*, 7-31.
- Please request from the authors:
  - Lane, S. P., & Hennes, E. P. (R&R). Conducting sensitivity analyses to identify and buffer power vulnerabilities in studies examining substance use over time.
  - Lane, S. P., & Tonnsen, B. L. (R&R). Statistical power for longitudinal developmental trajectories: The (non)-impact of age matching within measurement occasions.
- Additional slides and syntax: https://henneslab.wixsite.com/scsj/research
- Please contact Sean at lane84@purdue.edu, or Erin Hennes at ehennes@purdue.edu, for additional materials or consulting.

THE ANALYSIS FACTOR

Thank You!

Collaborators:

Erin Hennes
Doug Steinley
Robert Proctor
Carol Song
Lan Zhao
Denis McCarthy
Katie Witkiewitz
Kristina Jackson
Niall Bolger
Ya-Hsin Hung