Psychology 205: Research Methods in Psychology
Simulations as tools for research

William Revelle

Department of Psychology
Northwestern University
Evanston, Illinois USA

November, 2016

Outline

Introduction
- Motivation and Cognition
- Personality and Motivation

How does the simulation work?
- What are the variables in the study?
- Specifying the variables

Searching the Literature

Experimental vs Subject Variables, and Dependent Variables
- EVs and SVs
- Dependent Variables

What should you test?

How do you run the simulation?
- Block Randomization
- Collecting the data

How do I analyze the results?
- Descriptive Statistics

Inferential Statistics

Background on the use of simulations

1. Psychological theories are ways of organizing observable phenomena in terms of a limited number of unobservable constructs.
2. In addition to describing known phenomena, theories allow for prediction of as yet unobserved phenomena.
3. Theories may be stated as informal descriptions, or may be stated in formal propositional logic or in mathematical equations.
4. Complex theories that involve many variables may be stated as dynamic processes that change over time and that can best be captured in computer simulations.

Our experiment: Prior work

1. This experiment simulates the complexity of a real research program by simulating the complex relationships between:
   - a set of observed characteristics of individuals,
   - how they react to situations in terms of their motivational state,
   - and how motivational state, in turn, affects
   - cognitive performance.
2. Prior work in the Personality, Motivation, and Cognition Laboratory at Northwestern has allowed us to formulate a complex model of human cognition in response to stress (Anderson and Revelle, 1982, 1994; Revelle and Loftus, 1992; Revelle and Oehlberg, 2008; Revelle and Anderson, 1992; Revelle et al., 1976, 1980; Wilt et al., 2011b).
3. This simulation is based upon that work.

Cognitive Performance

Four broad classes of cognitive performance tasks

1. Tasks requiring sustained attention
   - vigilance
   - proof reading
2. Tasks requiring working memory
   - memory scanning
   - n-back
   - list learning
3. Tasks requiring retrieval from long term memory
   - tests of knowledge and skill (semantic memory)
4. "Complex tasks" require integration of attention, working memory, long term memory
   - Graduate Record Exam

Varieties of motivation

1. Resource availability
   - Energetic Arousal (Thayer et al., 1988; Thayer, 1989; Rafaeli and Revelle, 2006; Watson and Tellegen, 1985; Wilt et al., 2011a).
     - varies diurnally (Baehr et al., 2000; Revelle et al., 1980).
     - increased by stimulant drugs
   - Tense Arousal (Gray, 1991; Gray and McNaughton, 2000; Wilt et al., 2011b)
2. Resource Allocation
   - Focus on task
   - breadth of allocation

Personality and the availability and allocation of cognitive resources

1. Impulsivity/Extraversion
   - General Approach Behavior (Gray, 1991; Revelle, 1997; Smillie et al., 2006; Wilt and Revelle, 2009)
   - Extraversion and sensitivity to cues for reward (Smillie et al., 2012, 2013)
2. Impulsivity as related to diurnal (energetic) arousal rhythm
3. Anxiety and the allocation of resources
   - Anxiety leads to narrowing of attention
   - Anxiety and focus of attention on threat salient cues

Prior work

1. The Personality-Motivation-Cognition lab at NU
    - Person characteristics
    - Situational Characteristics
    - Intervening motivational states
    - cognitive performance
    - (Anderson and Revelle, 1982, 1994; Revelle and Loftus, 1992; Revelle and Oehlberg, 2008; Revelle and Anderson, 1992; Revelle et al., 1976, 1980; Wilt et al., 2011b)
2. The simulation is a theory of the relationship between these four sets of variables
    - person characteristics
    - situational characteristics
    - intervening motivational states
    - cognitive performance

Our experiment

1. In a sense, the simulation is a theory of the relationship between these four sets of variables
    - person characteristics,
    - situational characteristics,
    - intervening motivational states,
    - and cognitive performance.
2. The parameters of the model have been set to reflect empirical estimates of the strength of various relationships.
3. Several nuisance variables have been added to more properly simulate the problems of experimental design.

Simulation as theory testing

1. This simulation of the theory may be used as a test of the theory as well as a tool for understanding the complexity of research.
2. That is, although one may want to study the full model, because of the limitations of one's time and energy, one may study only a limited aspect of the model.
3. The student's objective is two fold: to better understand a limited aspect of a particular psychological theory, and to try to understand what are the relationships that have been specified in the model.

Simulation experiment is web based

1. The simulation is a web based program that allows you to "collect" the data on the web and then save the resulting output file to your computer to do subsequent analyses.
2. The biggest question is what should you study.
3. To answer this, you need to consider the variables available.
4. The underlying model is a function of the IVs and SVs.
5. Your job is to try to estimate the underlying model.
6. The model is psychologically plausible and is based upon prior results.

IVs, SVs, and DVs

Independent variables that are under control of the experimenter may be categorized as experimental variables and subject variables.

Experimental variables (IVs) may be manipulated by the experimenter.

Subject variables (SVs) are characteristics of the subjects that may be measured but not manipulated.

Dependent variables (DVs) are those variables thought be caused by the IVs and SVs. They are causally downstream from IVs and SVs.

IVs, SVs in this study

1. Independent Variables
   - **Drug** Placebo or Caffeine (you need to specify how much and how administered).
   - **Time of Day** Subjects may be run between 8 am and 10 pm (22:00 hours)
2. Subject Variables
   - **Sex** Males (1) or Females (2)
   - **Anxiety** Traits are stable characteristics of subjects. Trait anxiety is the tendency to feel tense and nervous in many situations. You need to specify how you measure them.
   - **Impulsivity** Trait impulsivity is the general tendency to do and say things rapidly, without stopping to think.
   - **Subject Number** One subject is run per day, so as S# increases, the subjects are later in the quarter.

Dependent variables (DVs) in this study

1. Motivational state variables
   - Energetic arousal: Feelings of energy and alertness versus sleepiness and tiredness. How are you measuring it?
   - Tense arousal: Feelings of tension and frustration versus calmness and relaxation.
2. Cognitive performance may be organized in terms of the total complexity of the task and the specific combination of memory load and of attention. Some complex tasks show an inverted U shaped relationship with arousal, others show a positive monotonic relationship, others a negative monotonic relationship. (Humphreys and Revelle, 1984)
3. Cost: It is more expensive to select subjects at the extreme of the distributions of anxiety and impulsivity because this requires mass testing and then rejecting many subjects to get the special subjects.

Selecting variables

1. The values of the Independent and Subject variables may be specified by the experimenter for each subject, or may be allowed to vary randomly.
2. If allowed to vary randomly, the experimental variables will be assigned values in a uniform random distribution.
3. The subject variables may either be specified (this simulates choosing particular subjects based upon a pretest) or may be allowed to vary randomly.
4. If varying, they will be assigned values based upon samples from a normal distribution.
5. If subjects are selected for particular values on a personality dimension, this is the same as rejecting many potential subjects and thus the Cost of running grows more rapidly than the simple number of subjects who participate.

Science does not occur in a vacuum – search the literature

Two powerful search engines

1. NUSearch (Northwestern Library requires VPN if off campus)
   - http://search.library.northwestern.edu/ NU Search
   - Citation linker (if you have a citation and you want to get it)
   - Use the DOI (Digital Object Identifier) from a citation if you have it.
   - This will then bring up a page of where to find it (e.g. EBSCOhost psych articles
2. Google Scholar (better if logged in through NU VPN)
   - Just a key word search and then ask to get it at NU
   - Also can search by citations to an article (this allows time travel)
   - The “Find it@NU” allows you to log into journals.

Help in searching the Library

1. http://libguides.northwestern.edu/psycmethods Getting started
2. http://libguides.northwestern.edu/c.php?g=114851&p=748373 Finding articles
   - Annual Review of Psychology
   - PsychINFO (has Boolean Search and can screen for year)
   - Web of Science (Allows you to search for who is citing the article. This allows you to find out what the recent literature is.)
3. Combining PsychInfo/Web of Science/Google Scholar
   - Find some classic article
   - Check out current citations to it
   - Download it, read the abstract, read the artilcle
   - Repeat

Experimental variables

1. Drug has two levels (0=Placebo or 1=Caffeine). Caffeine is known to act as a central nervous system stimulant although it has some side effects such as tremor (Revelle et al., 1976, 2012, 1980).
2. Time of Day has 15 levels (8 AM ... 10 PM or 8 ...22). Although most cognitive psychologists do not examine the effects of time of day on cognitive performance, there is a fairly extensive literature suggesting that performance does vary systematically across the day (Revelle et al., 1980; Revelle, 1993).

Subject variables

1. Sex of subject sometimes interacts with characteristics of the experiment (sex of experimenter, stress of experiment, type of task) and has sometimes been associated with levels of anxiety. In this study, Sex varies randomly taking on the values of 1 or 2. (Using the mnemonic of the number of X chromosomes, that is 1=M and 2=F).
2. Trait anxiety is a stable personality trait associated with feelings of tension, worry, and somatic distress. Trait anxious individuals are more sensitive to cues for punishment and non-reward and are also more likely to experience negative affect than are less trait anxious individuals (Gray, 1991; Wilt et al., 2011b). In this simulation, anxiety can take on values from 0-10.
3. Impulsivity (see next page)

Subject Variables part 2

1. Impulsivity is a stable personality trait associated with making up one's mind rapidly and doing and saying things with out stopping to think. It has been shown in prior work to relate to an inability to sustain performance, particularly in the morning (Anderson and Revelle, 1982, 1983). Theories of impulsivity have also suggested that impulsivity is related to a general sensitivity to cues for reward and to a greater propensity towards positive affect (Gray, 1991; Revelle, 1997; Wilt and Revelle, 2009; Zinbarg and Revelle, 1989). In this simulation, impulsivity can take on values from 0-10.

Dependent Variables

1. Energetic arousal reflects self reports of feelings of energy, activity, and alertness. EA has been shown to increase with exercise and to decrease with sleep deprivation (Thayer et al., 1988; Thayer, 1989). EA is also associated with feelings of positive affect (Rafaeli and Revelle, 2006; Watson and Tellegen, 1985; Wilt et al., 2011a).
2. Tense arousal reflects feelings of tension, frustration, and fear (Thayer, 1989) and is moderately associated with feelings of negative affect (Rafaeli and Revelle, 2006; Watson and Tellegen, 1985).
3. Performance in this simulation reflects accuracy on a simple decision task. A perfect score is 100, and performance deteriorates from that as a function of condition and motivational state. Abstractly, this may be thought of as accuracy on a vigilance task, or the ability to make accurate judgments on some sustained processing task (Anderson and Revelle, 1982, 1983, 1994; Humphreys and Revelle, 1984).

Dependent variables (part 2)

1. The cost of any experiment is a function of the number of subjects (it increases by 1 for every subject) and also of the scarcity of the subjects. Thus, if you choose to run just very high (10) and very low (1) anxiety subjects, this will require more prescreening to identify such subjects, and thus the cost will be higher than if you just chose average levels of anxiety, or if you just allowed anxiety to randomly vary. It is important to report the cost of the study you carry out.

What should you test?

1. Any experiment pits power against practicality. That is, the more subjects that are studied, the more statistical power that one has to detect an effect.
2. However, subjects are not an unlimited resource. They are hard to recruit and they are time consuming to run. This is reflected in the *cost* of the experiment.
3. In addition, for a particular number of subjects, as the number of variables that are examined increases, the potential number of higher order relationships (interactions) increases dramatically at the same time that the power to detect these interaction decreases because of the limited number of subjects in any one condition.

A word of caution

1. A reasonable approach is do have some theoretical reason to believe that a certain relationship exists, and then perhaps conduct a series of "pilot" studies to determine the sensitivity of certain parameter values.
2. The goal of this project is to try to determine at least some of the relationships that have been built into the model. You will be evaluated on principles of experimental design, not on the significance of the results.

Running the simulation

The simulation is available as a web based simulation that starts at http://personality-project.org/revelle/syllabi/205/simulation/simulation.experiment.php. It consists of three pages:

1. A set of instructions describing the experiment (somewhat redundant with this handout) and a request to specify how many subjects you want to run. Enter this number (e.g., 100) to continue on to the next page.
2. A list of each of the subjects that you asked to run. For each pseudo participant you have the option of letting the computer randomly assign them to a condition, or you may assign them to a condition. For each participant and for each variable you may take the default option (random assignment) or specify by clicking the appropriate radio button. When you are finished selecting the conditions for all the subjects, select the "submit" button. This takes you to the next page where you will be shown the data for all the subjects.

Intro to the experiment

The simulation is a web based program that allows you to "collect" the data on the web and then save the resulting output file to your computer to do subsequent analyses. The biggest question is what should you study. To answer this, you need to consider the variables available.

What are the variables you can specify in this simulation?

Independent variables that are under control of the experimenter may be categorized as experimental variables and subject variables. Experimental variables may be manipulated by the experimenter. Subject variables are characteristics of the subjects that may be measured but not manipulated.

In this experiment the Experimental Variables include

1. Drug condition (placebo or caffeine),
2. Time of Day. Given the realities of volunteer subjects, Time of Day is assumed to only vary between 8 am and 10 PM (22.00 hours).

The **Subject Variables** are that are "assessed" are:

1. Sex (Male=1, Female=2)
2. Trait Anxiety (0-10)
3. Trait Impulsivity (0-10)
4. Subject Number reflects when the subject appears in the quarter.

The **Dependent Variables** are measures of motivational state

1. Energetic Arousal
2. Tense Arousal)
3. Performance (accuracy on some attention task)

The values of the IVs and DVs may be specified by the experimenter for each subject, or may be allowed to vary randomly. If allowed to vary randomly, the experimental variables will be assigned values in a uniform random distribution. The subject variables may either be specified (this simulates choosing particular subjects based upon a pretest) or may be allowed to vary randomly. If varying, they will be assigned values based upon samples from a normal distribution. If subjects are selected for particular personality types, this is the same as rejecting many potential subjects and thus the Cost of running grows more rapidly than the simple number of subjects who participate.

It is a good idea to think carefully about your design before you run it.

Subject Number increases for every subject run in a particular experiment. Currently, it can not exceed 1000, but most runs will use less than 100. You will be asked to specify the number of subject that you want to run.

How many subjects do you want to run?

Choosing your subjects

This form will allow you to run up to the 8 subjects that you specified. For each simulated participant you need to specify the experimental conditions. When you are finished with specifying all the subjects, you can enter submit (at the end of the page). If you specify less than 8 participants, the last N will be filled with random participants. You can edit these out later in the statistical analysis if you choose.

Enter the conditions for the subject 1

Sex Random or Male Female
Drug Random or Placebo Caffeine
Time of Day Random or Fixed 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
Anxiety Random or Fixed 0 1 2 3 4 5 6 7 8 9 10

Impulsivity Random or Fixed 0 1 2 3 4 5 6 7 8 9 10

Enter the conditions for the subject 2
Sex Random or Male Female
Drug Random or Placebo Caffeine
Time of Day Random or Fixed 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
Anxiety Random or Fixed 0 1 2 3 4 5 6 7 8 9 10

Impulsivity Random or Fixed 0 1 2 3 4 5 6 7 8 9 10

Enter the conditions for the subject 3
Sex Random or Male Female
Drug Random or Placebo Caffeine
Time of Day Random or Fixed 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
Anxiety Random or Fixed 0 1 2 3 4 5 6 7 8 9 10

Impulsivity Random or Fixed 0 1 2 3 4 5 6 7 8 9 10

Enter the conditions for the subject 4
Sex Random or Male Female
Drug Random or Placebo Caffeine
Time of Day Random or Fixed 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
Anxiety Random or Fixed 0 1 2 3 4 5 6 7 8 9 10

Impulsivity Random or Fixed 0 1 2 3 4 5 6 7 8 9 10

Assigning subjects

1. The assignment of subjects to experimental conditions may be done using various random processes. Flipping a coin, using a table of random numbers, using the `sample` or `runif` functions in R are easy ways to generate random sequences. But to guarantee equal numbers of subjects in all conditions and to avoid end of experiment effects, it is convenient to *block randomize* subjects to conditions.
2. If you want to assign subjects to conditions using *block randomization* you can use the `block.random` function in the psych package to do so. For more information on using the `block.random` function, either see the supplementary material on the syllabus, or by using the help function in R ?block.random.

Using `block.random`

Using `block.random` in an experiment where you want to block randomize 2 factors, sex and drug and you want to run 48 subjects:

```
library(psych)   #make it active
my.cond <- block.random(n=48,c(sex=2,drug=2))
headtail(my.cond) #to  show just the first and last 4 cases
#my.cond   (without the # comment will show all the cases.
```

```
> headtail(my.cond)
    blocks sex drug
S1       1   1    2
S2       1   1    1
S3       1   2    1
S4       1   2    2
...    ... ...  ...
S45     12   2    1
S46     12   1    2
S47     12   1    1
S48     12   2    2
```

More on block randomization

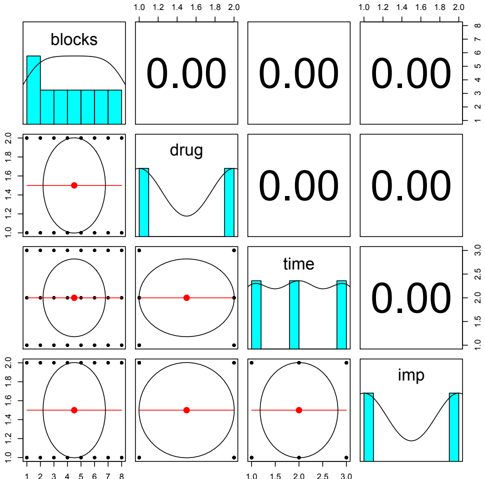
Now, consider an experiment with 96 subjects and two drug conditions, three time conditions, and two levels of impulsivity

```
my.cond <- block.random(n=96,c(drug=2,time=3,imp=2))
headtail(my.cond) #to  show just the first and last 4 cases
#my.cond   (without the # comment will show all the cases.
pairs.panels(my.cond) #this will show the lack of correlation between
```

```
    blocks drug time imp
S1       1    2    3   2
S2       1    1    1   1
S3       1    1    2   1
S4       1    1    2   2
...    ...  ...  ... ...
S93      8    2    1   2
S94      8    1    1   1
S95      8    2    2   2
S96      8    2    3   2
```

Visualizing block randomization

Although not necessary to do in order to use the block randomized conditions, it is useful to visualize what has happened by using the `pairs.panels` function on the output of `block.random`



Intro to the experiment

snum sex drug time anxiety impulsivity arousal tension performance cost
1 2 0 10 4 4 -7 14 58 1
2 2 0 14 0 4 78 4 72 1
3 2 0 12 5 8 -21 18 43 1
4 2 1 20 5 5 174 41 85 1
5 1 0 9 3 3 9 11 47 1
6 2 0 17 3 5 70 11 70 1
7 2 1 18 5 8 52 44 62 1
8 1 1 8 5 2 65 37 72 1
9 2 0 13 3 7 9 14 54 1
10 1 1 16 7 2 125 60 82 1
11 1 0 17 1 3 98 6 55 1
12 1 0 22 2 7 30 8 47 1
13 2 0 11 6 8 14 15 46 1
14 1 0 20 0 4 62 5 60 1
15 2 1 19 2 6 127 16 87 1
16 1 0 15 7 7 34 26 69 1
17 1 0 9 2 8 -26 8 29 1
18 1 1 21 1 2 62 14 71 1
19 2 0 22 5 0 32 22 56 1
20 2 0 13 7 5 119 23 79 1
21 1 1 9 3 6 74 21 72 1
22 1 1 15 2 3 87 20 74 1
23 1 1 18 6 7 113 41 82 1
24 1 1 12 4 9 32 23 68 1

After your data are collected

1. The final page has columns labels for the variables and then one row of data for each subject. You may select the entire page and copy and paste it into the text editor of your choice to save it for later data analysis. Or you may paste it directly into R using the `read.clipboard` function from the **psych** package.
2. See the supplementary handout on using R for data analysis in research methods.
3. It is important to save this page in your favorite text editor if you want to do further analyses of these data at a later time.

Analyzing the results – overview

1. The simulation produces a page of data with column labels. This page may be selected and saved to your hard disk (remember the name) for later analysis. Alternatively, you can copy the page (use the copy command in your browser) and then paste it into R using the `read.clipboard` function. (Make sure you have made the *psych* active first).
2. The analyses you choose to do depend upon what you studied, but in general you will probably want to do the following steps. (These are discussed in more detail in the tutorial prepared by Katharine Funkhouser: Analyzing the simulation experiment or in slightly more detail in the tutorial for 205.
3. What follows is an even shorter introduction.

Getting the data from the browser page: copy to the clipboard and then read it

R code

```
> my.data <- read.clipboard()
> my.data
```

	snum	sex	drug	time	anxiety	impulsivity	arousal	tension	performance	cost
1	1	2	0	10	4	4	-7	14	58	1
2	2	2	0	14	0	4	78	4	72	1
3	3	2	0	12	5	8	-21	18	43	1
4	4	2	1	20	5	5	174	41	85	1
5	5	1	0	9	3	3	9	11	47	1
6	6	2	0	17	3	5	70	11	70	1
7	7	2	1	18	5	8	52	44	62	1
8	8	1	1	8	5	2	65	37	72	1
9	9	2	0	13	3	7	9	14	54	1
10	10	1	1	16	7	2	125	60	82	1
11	11	1	0	17	1	3	98	6	55	1
12	12	1	0	22	2	7	30	8	47	1
13	13	2	0	11	6	8	14	15	46	1
14	14	1	0	20	0	4	62	5	60	1
15	15	2	1	19	2	6	127	16	87	1
16	16	1	0	15	7	7	34	26	69	1
17	17	1	0	9	2	8	-26	8	29	1
18	18	1	1	21	1	2	62	14	71	1
19	19	2	0	22	5	0	32	22	56	1
20	20	2	0	13	7	5	119	23	79	1
21	21	1	1	9	3	6	74	21	72	1
22	22	1	1	15	2	3	87	20	74	1
23	23	1	1	18	6	7	113	41	82	1

Get the data and describe them

First, copy your browser output page to the clipboard (copy). Then read them in and describe them.

Two things to remember:

1. library(psych) to make the psych package active. You just need to do this once per time that you start R.
2. All function calls use parentheses: make sure that you close your parenthetical statements.

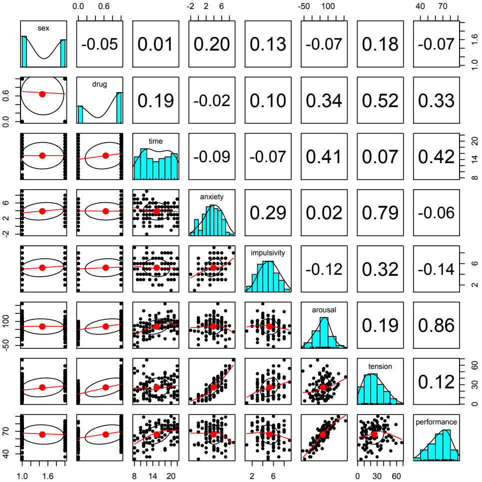
R code

```
library(psych)  #just do this once per R run
sim.data <- read.clipboard() #note the  parentheses
describe(sim.data)
```

	var	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
snum	1	100	50.50	29.01	50.5	50.50	37.06	1	100	99	0.00	-1.24	2.90
sex	2	100	1.47	0.50	1.0	1.46	0.00	1	2	1	0.12	-2.01	0.05
drug	3	100	0.64	0.48	1.0	0.68	0.00	0	1	1	-0.57	-1.69	0.05
time	4	100	15.26	4.35	15.0	15.31	5.93	8	22	14	-0.01	-1.29	0.43
anxiety	5	100	3.87	2.25	4.0	3.95	2.97	-2	9	11	-0.25	-0.45	0.23
impulsivity	6	100	5.23	1.69	5.0	5.20	1.48	1	9	8	0.03	-0.53	0.17
arousal	7	100	68.03	48.23	71.0	68.46	43.74	-57	210	267	-0.08	0.17	4.82
tension	8	100	25.79	14.54	25.0	24.80	17.79	2	68	66	0.54	-0.39	1.45
performance	9	100	65.72	12.94	69.0	66.41	13.34	33	90	57	-0.45	-0.53	1.29

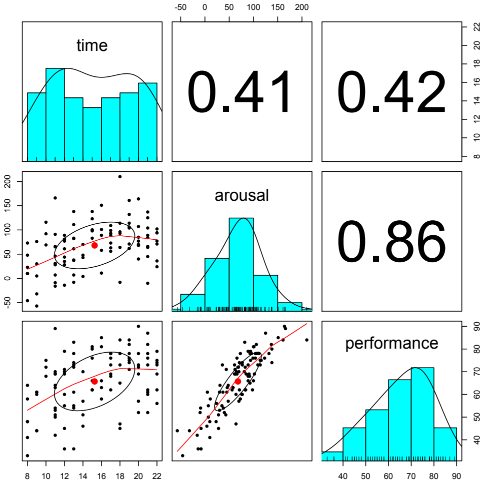
Show the correlations using pairs.panels

`pairs.panels(sim.data[2:9])` Select variables 2-9



Just show a selected set of variables

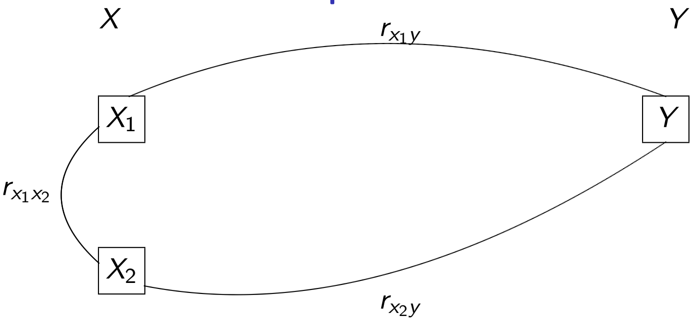
`pairs.panels(sim.data[c(4,7,9)])` select variables 4, 7, and 9



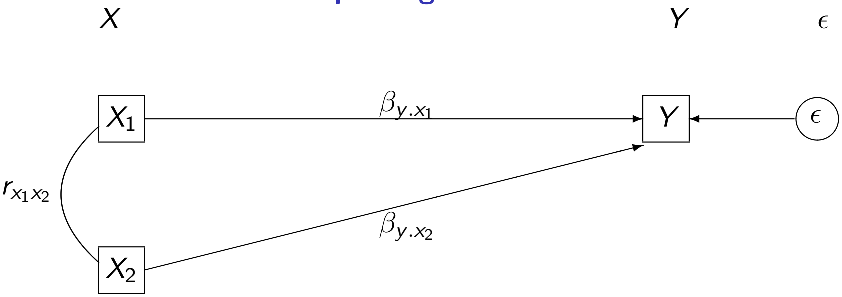
Many types of models are special cases of General Linear Model

1. Y = bX (X is continuous) Regression
2. Y = bX (X has two levels) t-test
3. Y = bX (X has > 2 levels) F-test
4. $Y = b_1X_1 + b_2X_2 + b_3X_3$ (X_i is continuous) Multiple regression
5. $Y = b_1X_1 + b_2X_2 + b_3X_{12}$ (X_i is continuous) Multiple regression with an interaction term
   - In this case, we need to zero center the X_i so that the product is independent of the Xs.
6. $Y = b_1X_1 + b_2X_2 + b_3X_{12}$ (X_i is is categorical) Analysis of Variance
7. $Y = b_1X_1 + b_2X_2 + b_3X_{12} + Z$ (X_i and Z are continuous) Analysis of Covariance

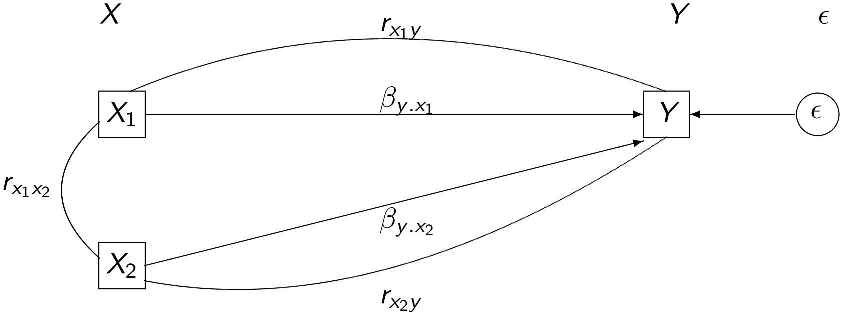
Multiple correlations



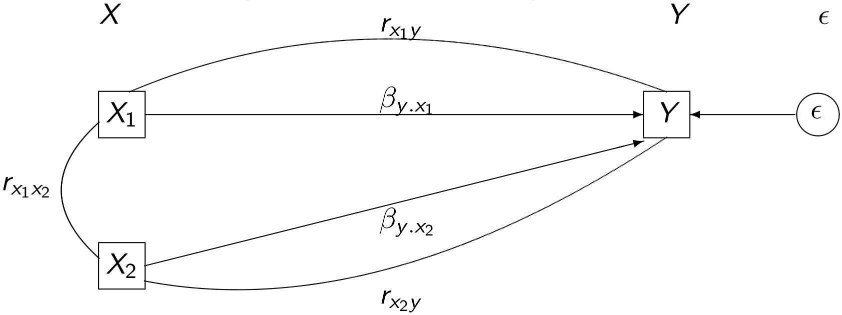
Multiple Regression



Multiple Regression: decomposing correlations



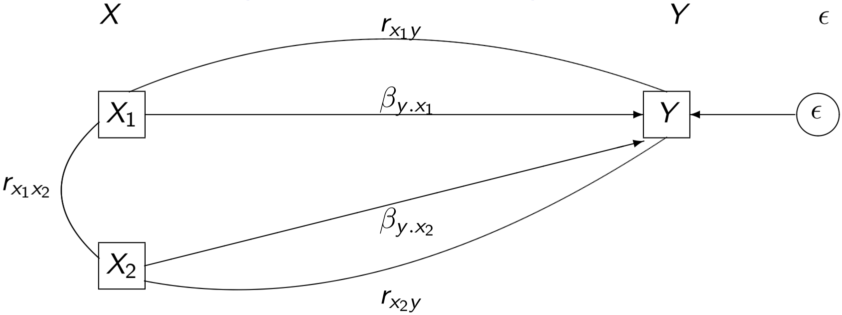
Multiple Regression: decomposing correlations



$$r_{x_1 y} = \overbrace{\beta_{y.x_1}}^{\color{blue}{direct}} + \overbrace{r_{x_1 x_2}\beta_{y.x_2}}^{\color{red}{indirect}}$$

$$r_{x_2 y} = \underbrace{\beta_{y.x_2}}_{\color{blue}{direct}} + \underbrace{r_{x_1 x_2}\beta_{y.x_1}}_{\color{red}{indirect}}$$

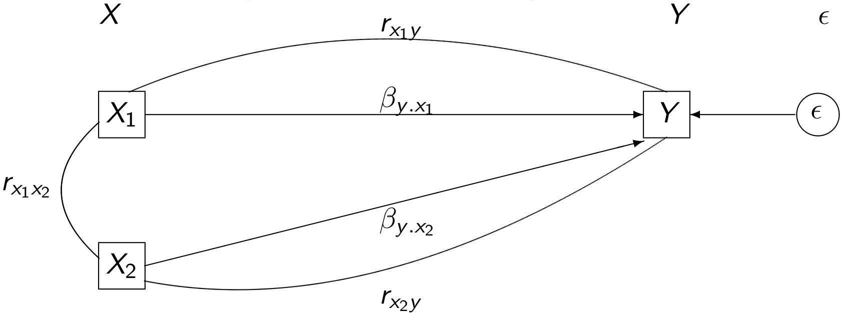
Multiple Regression: decomposing correlations



$$r_{x_1 y} = \overbrace{\beta_{y.x_1}}^{direct} + \overbrace{r_{x_1 x_2}\beta_{y.x_2}}^{indirect} \qquad \beta_{y.x_1} = \frac{r_{x_1 y} - r_{x_1 x_2} r_{x_2 y}}{1 - r^2_{x_1 x_2}}$$

$$r_{x_2 y} = \underbrace{\beta_{y.x_2}}_{direct} + \underbrace{r_{x_1 x_2}\beta_{y.x_1}}_{indirect} \qquad \beta_{y.x_2} = \frac{r_{x_2 y} - r_{x_1 x_2} r_{x_1 y}}{1 - r^2_{x_1 x_2}}$$

Multiple Regression: decomposing correlations



$$r_{x_1y} = \overbrace{\beta_{y.x_1}}^{\color{blue}{direct}} + \overbrace{r_{x_1x_2}\beta_{y.x_2}}^{\color{red}{indirect}}$$

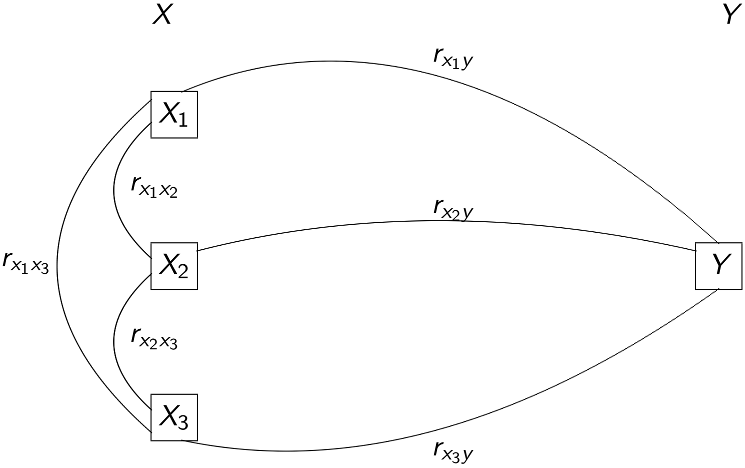
$$r_{x_2y} = \underbrace{\beta_{y.x_2}}_{\color{blue}{direct}} + \underbrace{r_{x_1x_2}\beta_{y.x_1}}_{\color{red}{indirect}}$$

$$\beta_{y.x_1} = \frac{\color{blue}{r_{x_1y}} - \color{red}{r_{x_1x_2}r_{x_2y}}}{1 - \color{red}{r^2_{x_1x_2}}}$$

$$\beta_{y.x_2} = \frac{\color{blue}{r_{x_2y}} - \color{red}{r_{x_1x_2}r_{x_1y}}}{1 - \color{red}{r^2_{x_1x_2}}}$$

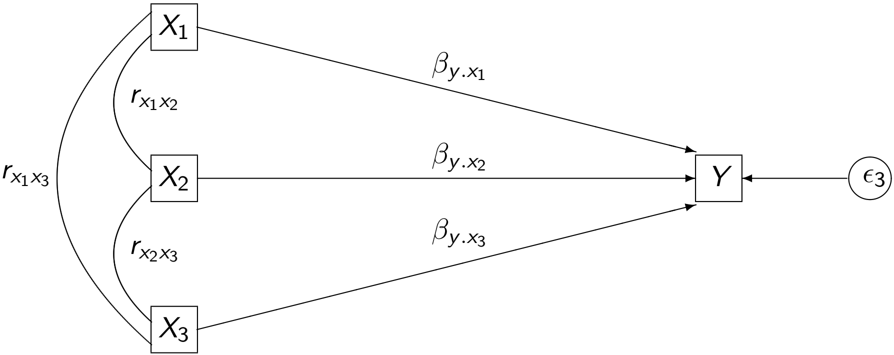
$$R^2 = r_{x_1y}\beta_{y.x_1} + r_{x_2y}\beta_{y.x_2}$$

What happens with 3 predictors? The correlations

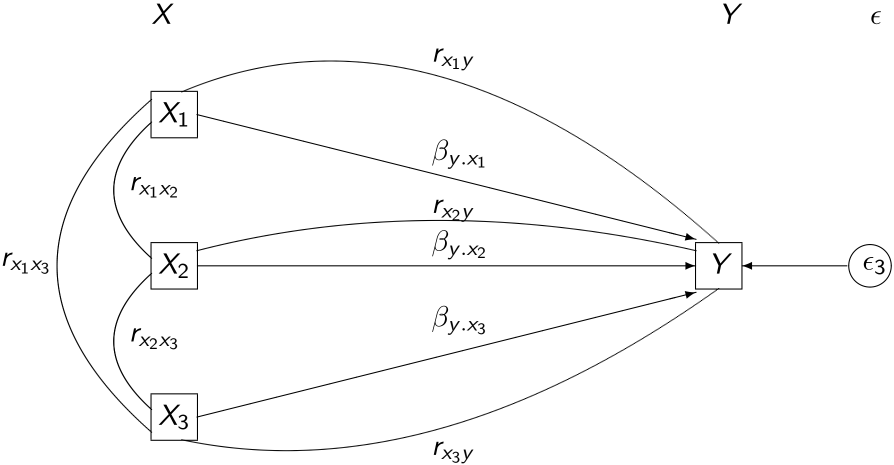


What happens with 3 predictors? β weights

X Y ϵ



What happens with 3 predictors?



$$r_{x_1 y} = \underbrace{\beta_{y.x_1}}_{\textit{direct}} + \underbrace{r_{x_1 x_2}\beta_{y.x_1} + r_{x_1 x_3}\beta_{y.x_3}}_{\textit{indirect}} \quad r_{x_2 y} = \ldots \quad r_{x_3 y} = \ldots$$

The math gets tedious

Multiple regression and linear algebra

- Multiple regression requires solving multiple, simultaneous equations to estimate the direct and indirect effects.
  - Each equation is expressed as a $r_{x_i y}$ in terms of direct and indirect effects.
  - Direct effect is $\beta_{y.x_i}$
  - Indirect effect is $\sum_{j \neq i} beta_{y.x_j} r_{x_j y}$
- How to solve these equations?
- Tediously, or just use linear algebra.

A simple way to do regressions

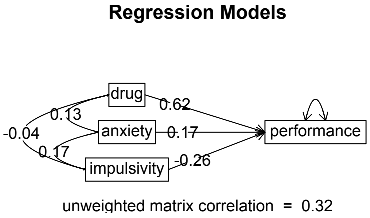
R code

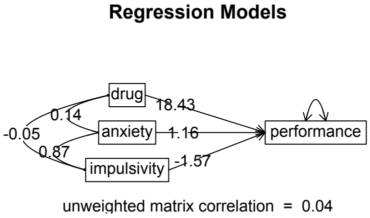
```
#as correlations
setCor("performance",c("drug","anxiety","impulsivity"),
       data=my.data)
#as unstandardized regression coefficients
setCor("performance",c("drug","anxiety","impulsivity"),
   data=my.data,std=FALSE)
```

```
Multiple Regression from raw data
Beta weights
            performance
drug               0.62
anxiety            0.17
impulsivity       -0.26
Multiple R
            performance
performance        0.71
multiple R2
            performance
performance        0.51

Multiple Regression from raw data
Beta weights
            performance
drug              18.43
anxiety            1.16
impulsivity       -1.57
Multiple R
```

Standardized and raw path diagrams





Compare the inferential statistics

```
 t of Beta Weights
            performance
drug               3.92
anxiety            1.04
impulsivity       -1.61

Probability of t <
            performance
drug            0.00085
anxiety         0.31000
impulsivity     0.12000

SE of Beta weights
            performance
drug               4.70
anxiety            1.11
impulsivity        0.98
```


```
 t of Beta Weights
            performance
drug               3.92
anxiety            1.04
impulsivity       -1.61

Probability of t <
            performance
drug            0.00085
anxiety         0.31000
impulsivity     0.12000
```

The General Linear Model

R code

```
model = lm(y ~ x1 + x2 + x1*x2,data=my.data)
```

But the product term is correlated with X_1 and X_2 and so we need to zero center (subtract out the mean) from the predictors.

R code

```
cen.data.df <- data.frame(scale(sim.data,scale=FALSE))
model = lm(arousal ~ drug * time,data=cen.data.df)
summary(model)  #to show the results
```

The basic moderated multiple regression

```
cen.data.df <- data.frame(scale(sim.data,scale=FALSE))
model = lm(arousal ~ drug * time,data=cen.data.df)
summary(model)

Call:
lm(formula = arousal ~ drug * time, data = cen.data.df)

Residuals:
     Min       1Q   Median       3Q      Max
-111.886  -25.533    1.889   24.796  120.735

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  -0.3015     4.3534  -0.069 0.944933
drug         27.8027     9.1221   3.048 0.002978 **
time          3.9382     1.0078   3.908 0.000173 ***
drug:time     0.7470     2.0731   0.360 0.719393
---
Signif. codes:  0 Ô***Õ 0.001 Ô**Õ 0.01 Ô*Õ 0.05 Ô.Õ 0.1 Ô Õ 1

Residual standard error: 42.72 on 96 degrees of freedom
Multiple R-squared:  0.2392,       Adjusted R-squared:  0.2155
F-statistic: 10.06 on 3 and 96 DF,  p-value: 7.932e-06
```

1. Center the data to allow for interactions
2. Specify the linear model in terms of the Dependent Variable as a function of a number of Independent Variables
3. Summarize it
   - Each row of the regression is a 1 by 96 degrees of freedom test and is reported (e.g.,) as a $t_{96} = 3.05, p < .01$

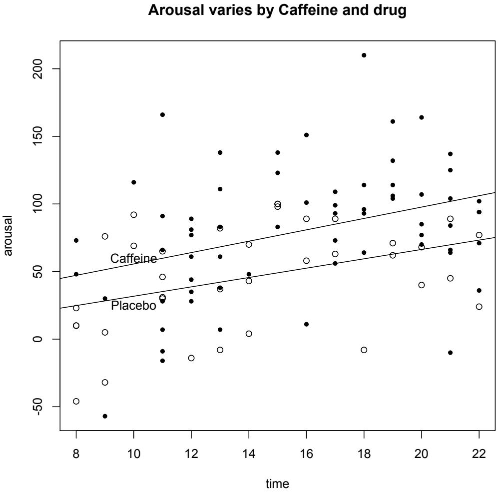
Multiple regression is hard to understand from just the table

A picture is worth a 1000 words.

```
with(sim.data,plot(arousal ~ time,pch=21-drug,
      main= 'Arousal varies by Caffeine and drug'))
by(sim.data,sim.data$drug,function(x)
        abline(lm(arousal ~time,data=x)))
text(10,25,"Placebo")
text(10,60,"Caffeine")
```

1. Draw the basic graph with different plot characters (pch) for different drug conditions
2. Add the lines for the two drug conditions
3. Add labels for each line

Show the graph

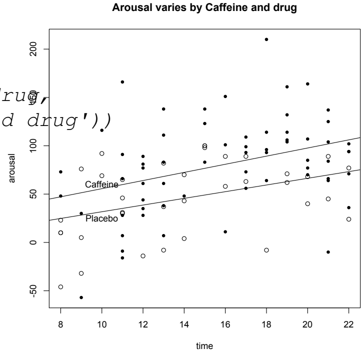


The figure in more detail

```
with(sim.data,plot(arousal ~ time,pch=21-drug,
      main= 'Arousal varies by Caffeine and drug'))

by(sim.data,sim.data$drug,function(x)
        abline(lm(arousal ~time,data=x)))

text(10,25,"Placebo")
text(10,60,"Caffeine")
```



Statistics and graphics in R: general comments

1. Build up the commands, first do a simple one, then a more complex one.
2. When you make a mistake, use the up arrow on the keyboard to do the command again.
3. Copy your code to a text editor so that you can save it.
4. Annotate it for your self.
5. Copy the output to your text editor to print it (use a fixed width font).
6. Use the ? about a function (e.g., ? setCor, ?lm) and then try the examples.
7. Use the various tutorials on the syllabus.
8. Analyzing the simulation experiment or in slightly more detail in the tutorial for 205.
9. Don't Panic.

Sections of an APA style paper

1. Title page
2. Abstract (separate page - write last)
3. Introduction (starts a new page, but uses the title of the paper
4. Method (does not start a new page)
    - Materials
    - Procedure
5. Results
6. Discussion and Conclusion
7. References
8. Tables
9. Figure Captions
10. Figures

Structure of a paper: The hour glass

1. Introduction
   - The problem
   - Prior work
   - Overview
2. Method
   - Materials
   - Procedure
3. Results
   - Use paragraphs for different findings
   - Words, Numbers, Statistics
   - Figures and tables are not the subject nor object of sentences
4. Discussion
   - Wake up Grandma
   - General discussion
   - Take home message

Title Page

1. Title
2. Author(s)
3. affiliation
4. running head

Abstract

1. In less than 100-150 words tell
   - What was the problem
   - What was done
   - Who were the participants
   - What was found
   - What does it mean
2. Write the abstract last once you know what you have found

Introduction

1. Theoretical question being examined
    - Why is it interesting?
    - Why is it important?
2. Review of previous work
    - What has already been found?
    - What unsolved problems are raised by prior research?
    - Are there methodological flaws in previous work that needs to be addressed?
3. Overview of study
    - Brief paragraph describing basic design
    - Variables (both constructs and observed) of interest

Methods/Procedures

1. Enough information to allow someone to replicate the study if they chose to do so
    - Subjects/participants
    - who were they
    - how were they chosen
    - what special characteristics do they have
2. Apparatus/materials
    - any special equipment or forms
3. Procedure
    - what was said by the experimenter
    - what was asked of the participant

Results

1. Any finding worth discussing is introduced here
   - say it in words (doing X led to an increase in Y)
   - say it in numbers (mean Y for low X = , mean Y for high X =
   - say it in statistics ($F_{a,b} = xx.xx$, $p < .0y$
2. references are made to tables and figures
   - Use parenthetical references to figures and table
   - Don't make them subject or objects of sentences
3. Any number to be introduced in paper needs to be mentioned here

Discussion and Conclusion

1. What do the results mean
   - In a manner that a generalist can understand (Wake up Grandma)
   - In a manner than a specialist will appreciate
2. How do the results relate to the original hypothesis?
   - Why are they different (if they are)?
3. Implications for further study
   - What should be the next set of studies to be conducted?
4. Final paragraph gives the "take home message"

Final sections

1. References
   - literature cited (and read by you)
2. Tables
   - See Leary, Plonsky, etc.
3. Figure Captions (one page for captions for all of the figures).
   - Captions are more than titles. They can explain what the graphs mean.
   - If you have any thing like error bars that need to be explained, do so.
4. Figures (do not include figure captions on these figures)

Additional Comments

1. Stylistic Considerations
   - Write for the intelligent but uniformed reader
   - Remember topic sentences and what is important in each section
2. Special things to consider - scientific readers are different. They do not read straight through, but rather
   - abstract
   - First paragraph of intro
   - First paragraph of discussion
   - Last paragraph of discussion
   - then, and only then, the whole paper
3. So, "punch up" those critical paragraphs

Guides to writing an APA paper

1. APA manual of style - The official reference
2. Leary (Chapter 15) detailed example
3. Plonsky "cheat sheet" to the APA paper http://www.uwsp.edu/psych/apa4b.htm
4. Also possible to use sophisticated typesetting language: LaTex with the apa.cls style sheet at http://personality-project.org/revelle/syllabi/205/apa.style.html

Anderson, K. J. and Revelle, W. (1982). Impulsivity, caffeine, and proofreading: A test of the Easterbrook hypothesis. *Journal of Experimental Psychology: Human Perception and Performance*, 8(4):614–624, doi:10.1037/0096-1523.8.4.614.

Anderson, K. J. and Revelle, W. (1983). The interactive effects of caffeine, impulsivity and task demands on a visual search task. *Personality and Individual Differences*, 4(2):127–134, doi:10.1016/0191-8869(83)90011-9.

Anderson, K. J. and Revelle, W. (1994). Impulsivity and time of day: Is rate of change in arousal a function of impulsivity? *Journal of Personality and Social Psychology*, 67(2):334–344, doi:10.1037/0022-3514.67.2.334.

Baehr, E. K., Revelle, W., and Eastman, C. I. (2000). Individual differences in the phase and amplitude of the human circadian temperature rhythm: with an emphasis on morningness-eveningness. *Journal of Sleep Research*, 9(2):117–127, doi:10.1046/j.1365-2869.2000.00196.x.

Gray, J. A. (1991). The neuropsychology of temperament. In Strelau, J. and Angleitner, A., editors, *Explorations in temperament: International perspectives on theory and measurement*, pages 105–128. Plenum Press, New York, NY.

Gray, J. A. and McNaughton, N. (2000). *The Neuropsychology of anxiety: An enquiry into the functions of the septo-hippocampal system*. Oxford University Press, Oxford.

Humphreys, M. S. and Revelle, W. (1984). Personality, motivation, and performance: A theory of the relationship between individual differences and information processing. *Psychological Review*, 91(2):153–184, doi:10.1037/0033-295X.91.2.153.

Rafaeli, E. and Revelle, W. (2006). A premature consensus: Are happiness and sadness truly opposite affects? *Motivation and Emotion*, 30(1):1–12, doi:10.1007/s11031-006-9004-2.

Revelle, W. (1993). Individual differences in personality and motivation: 'non-cognitive' determinants of cognitive performance. In Baddeley, A. and Weiskrantz, L., editors, *Attention: Selection, awareness, and control: A tribute to Donald Broadbent*, pages 346–373. Clarendon Press/Oxford University Press, New York, NY.

Revelle, W. (1997). Extraversion and impulsivity: The lost dimension? In Nyborg, H., editor, *The scientific study of human nature: Tribute to Hans J. Eysenck at eighty*, pages 189–212. Pergamon/Elsevier Science Inc, Amsterdam, Netherlands.

Revelle, W., Amaral, P., and Turriff, S. (1976). Introversion-extraversion, time stress, and caffeine: effect on verbal performance. *Science*, 192:149–150, doi:10.1126/science.1257762.

Revelle, W. and Anderson, K. J. (1992). Models for the testing of theory. In Gale, A. and Eysenck, M., editors, *Handbook of Individual Differences: Biological Perspectives*, chapter 4, pages 81–113. John Wiley and Sons, Chichester, England.

Revelle, W., Condon, D., and Wilt, J. (2012). Caffeine. In

Ramachandran, V. S., editor, *Encyclopedia of Human Behavior, 2nd Edition.*, pages 423–429, Oxford. Elsevier.

Revelle, W., Humphreys, M. S., Simon, L., and Gilliland, K. (1980). Interactive effect of personality, time of day, and caffeine: A test of the arousal model. *Journal of Experimental Psychology General*, 109(1):1–31, doi:10.1037/0096-3445.109.1.1.

Revelle, W. and Loftus, D. A. (1992). The implications of arousal effects for the study of affect and memory. In Christianson, S.-A., editor, *The handbook of emotion and memory: Research and theory*, pages 113–149. Lawrence Erlbaum Associates, Inc, Hillsdale, NJ, England.

Revelle, W. and Oehlberg, K. (2008). Integrating experimental and observational personality research – the contributions of Hans Eysenck. *Journal of Personality*, 76(6):1387–1414, doi:10.1111/j.1467-6494.2008.00526.x.

Smillie, L. D., Cooper, A., Wilt, J., and Revelle, W. (2012). Do extraverts get more bang for the buck? refining the affective-reactivity hypothesis of extraversion. *Journal of Personality and Social Psychology*, 103(2):306–326, doi:10.1037/a0028372.

Smillie, L. D., Geaney, J., Wilt, J., Cooper, A. J., and Revelle, W. (2013). Aspects of extraversion are unrelated to pleasant affective reactivity: Further examination of the affective reactivity hypothesis. *Journal of Research in Personality*, 47(5):580–587, doi:10.1016/j.jrp.2013.04.008.

Smillie, L. D., Pickering, A. D., and Jackson, C. J. (2006). The new reinforcement sensitivity theory: Implications for personality measurement. *Personality and Social Psychology Review*, 10(4):320–335.

Thayer, R. E. (1989). *The biopsychology of mood and arousal*. Oxford University Press, The biopsychology of mood and arousal. xi, 234 pp. New York, NY.

Thayer, R. E., Takahashi, P. J., and Pauli, J. A. (1988). Multidimensional arousal states, diurnal rhythms, cognitive and social processes, and extraversion. *Personality and Individual Differences*, 9(1):15–24.

Watson, D. and Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin*, 98(2):219–235, doi:10.1037/0033-2909.98.2.219.

Wilt, J., Funkhouser, K., and Revelle, W. (2011a). The dynamic relationships of affective synchrony to perceptions of situations. *Journal of Research in Personality*, 45:309–321, doi:10.1016/j.jrp.2011.03.005.

Wilt, J., Oehlberg, K., and Revelle, W. (2011b). Anxiety in personality. *Personality and Individual Differences*, 50(7):987–993, doi:10.1016/j.paid.2010.11.014.

Wilt, J. and Revelle, W. (2009). Extraversion. In Leary, M. R. and Hoyle, R. H., editors, *Handbook of Individual Differences in Social Behavior*, chapter 3, pages 27–45. Guilford Press.

Zinbarg, R. and Revelle, W. (1989). Personality and conditioning:

A test of four models. *Journal of Personality and Social Psychology*, 57(2):301–314, doi:10.1037/0022-3514.57.2.301.