Data Analysis

Richard W. Foltin, Ph.D.

Columbia University
Department of Psychiatry
New York State Psychiatric Institute

Missing Data

- Missing data can be imputed if "missingness" is random
- Of course this is not always true as missing data are often related to the independent manipulation of dose
  - Drug-induced impairment, illness or other subjective effect hinders collection of data

Did We Do An Experiment?

- Order of testing conditions was not assigned at random
  - We did not give a participant the largest dose of cocaine until we had tested a lower cocaine dose in that participant
- Our participants were not a random sample of cocaine users in Baltimore
  - Sample of Convenience
  - Inferential statistics require random sampling

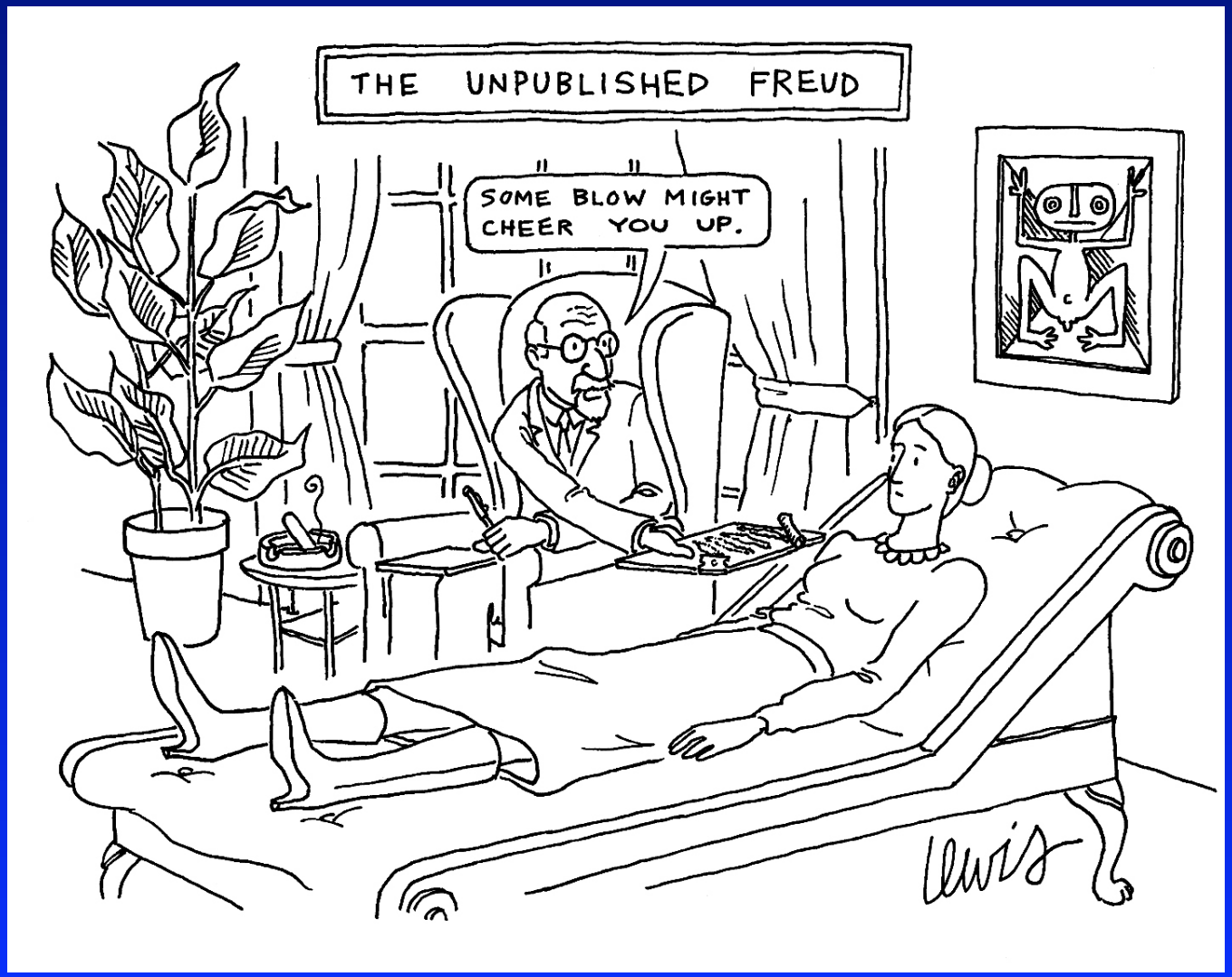
Discussion Point

Do our designs violate the assumptions of random assignment to condition and random testing order?

Sampling Problem: Example 1

Excellent health statistics - smokers are less likely to die of age related illnesses.'

Sampling Problem: Example 2



Discussion Point

Should we be concerned about our sampling techniques and the inferences about the target population that can be based on our sample?

Scales of Measurement

Scale	Property
Nominal	Identity
Ordinal	Identity + Magnitude
Interval	Identity + Magnitude + Equal Distance
Ratio	Identity + Magnitude + Equal Distance + true 0 point

Parametric Statistics can only be accomplished with Interval and Ratio scales

Subjective-effects Questionnaires

- Likert
  - Are the distances between points equal?
  - Is the distribution normal?
- "Approximately interval"
- Using derived scores based on average of multiple measures makes the distances closer to equal (increases number of points) and the distribution approximately normal

Discussion Point

Should we be concerned about conducting parametric statistics on data that might be on an ordinal or approximately interval scale of measurement?

A Favorite Article

- **Physiologic, subjective, and behavioral effects of amphetamine, methamphetamine, ephedrine, phenmetrazine, and methylphenidate in man**
  - **Martin WR, Sloan JW, Sapira JD, Jasinski DR.**
  - **Clin Pharmacol Ther, Vol. 12 (2): 245-58, 1971**

- **Comprehensive Assessments (up to 12 hr)**
  - **5 Physiological measures**
  - **Caloric content of lunch**
  - **Subjective Effects- ARCI, liking**
  - **Drug identification**
  - **Observer ratings**
  - **Urinary catecholamines**

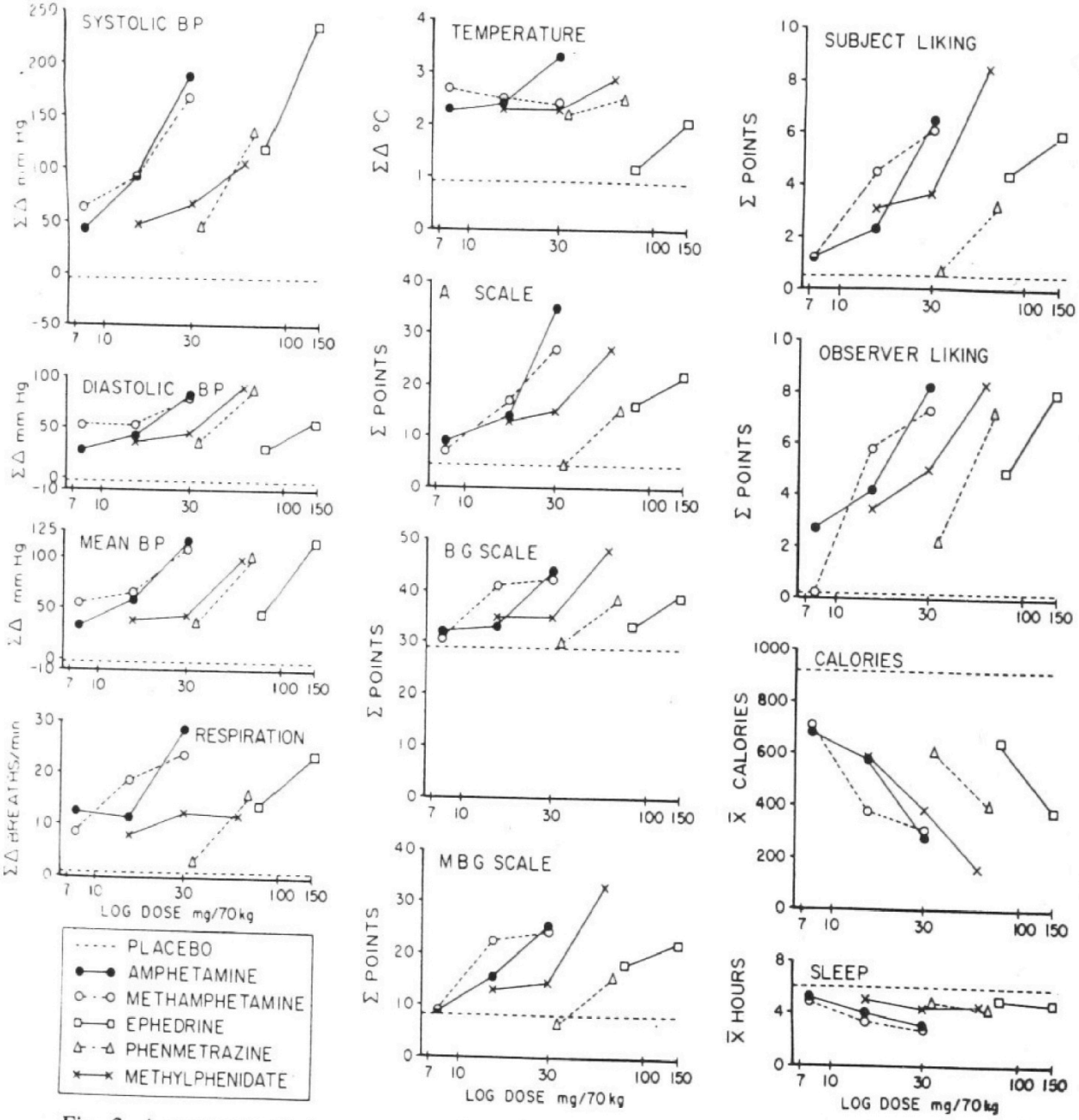
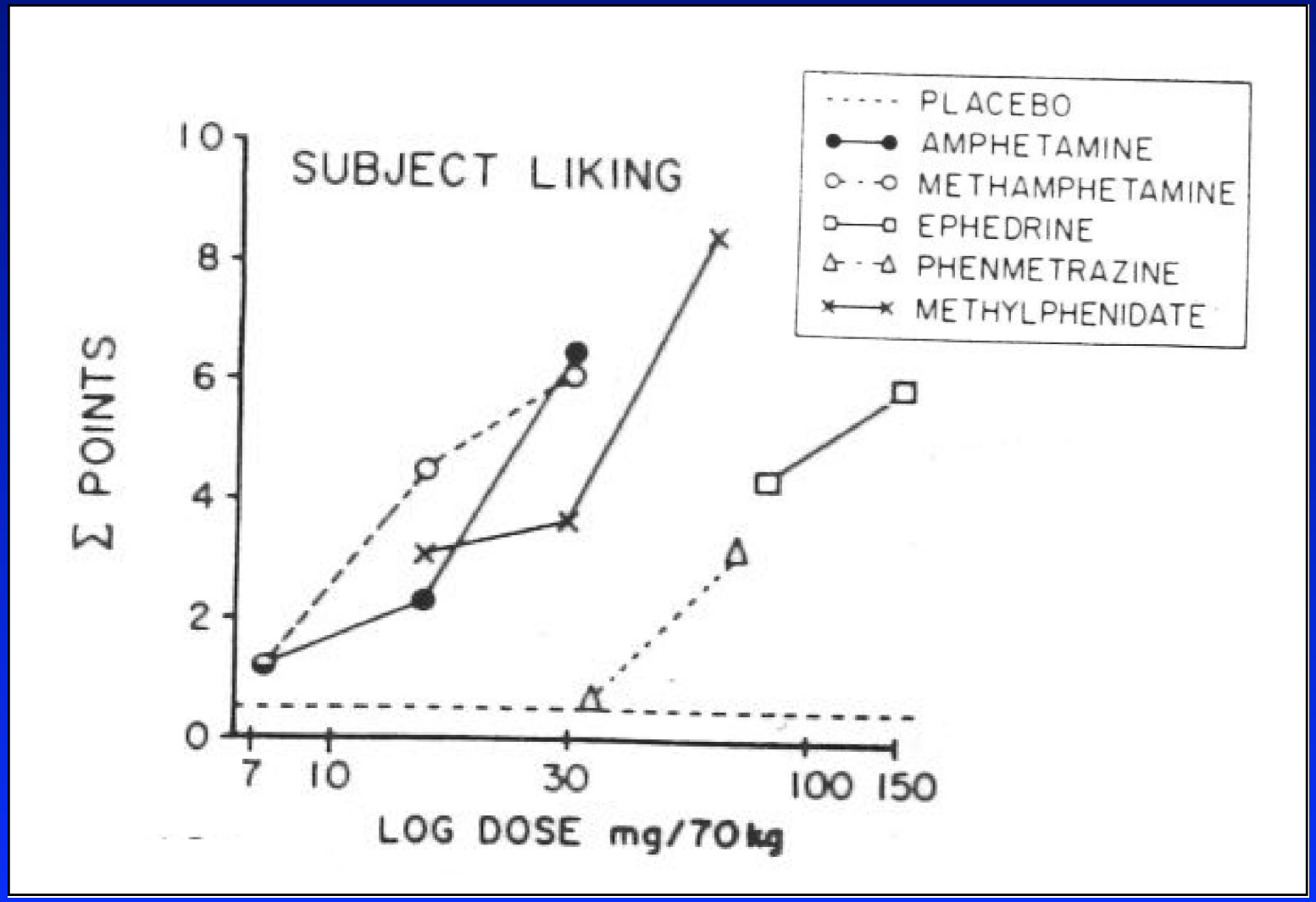


Fig. 2. A summary of dose-response relationships obtained for several physiologic and subjective parameters with amphetamine-like drugs. These data provided the basis on which relative potency data presented in Table I were calculated.

Participant Ratings of "Liking"



> 950 Data Points

- 12 participants
- "A valid bioassay was one in which there was a significant regression (dose-effect; RWF added), no significant deviation from linearity, no significant difference in preparations, and a significant treatment effect. (p 248)"
  - ANOVAs done across weeks
- Dose-effects compared by calculating relative potency and confidence intervals
  - Amphetamine was the standard and given a potency of 1
- Presented correlations among measures

Griffiths et al. Data Presentation

Vol. 271

THE JOURNAL OF PHARMACOLOGY AND EXPERIMENTAL THERAPEUTICS
Copyright © 1994 by The American Society for Pharmacology and Experimental Therapeutics
JPET 271:683-694, 1994

Vol. 271, No. 2
Printed in U.S.A.

Tandospirone and Alprazolam: Comparison of Behavioral Effects and Abuse Liability in Humans[1]

SUZETTE M. EVANS,[2] JOSEPH R. TROISI II[3] and ROLAND R. GRIFFITHS

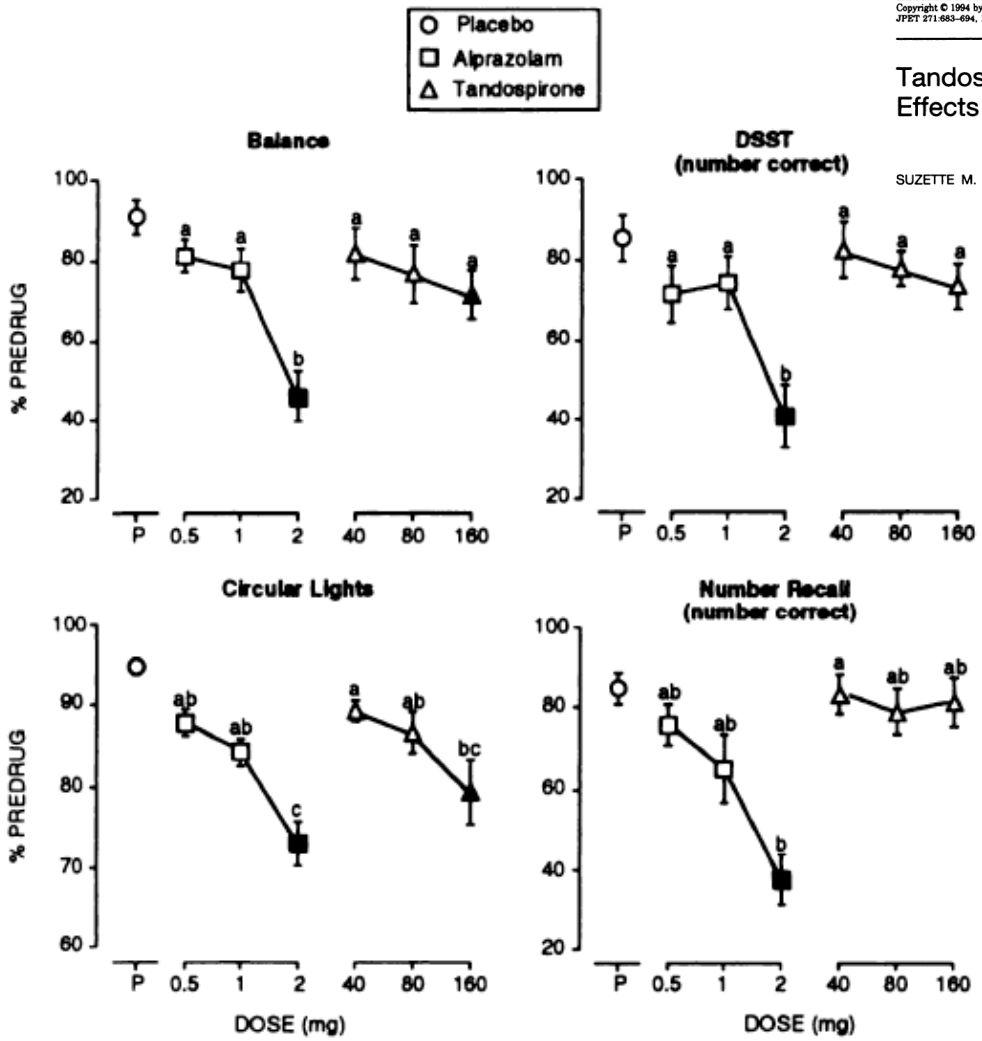


Fig. 2. Dose effects for the peak magnitude of performance impairment. Peak effect data for the balance, circular lights, DSST (number correct) and enter and recall (number correct) tasks are expressed as percentages of predrug scores. *X*-axes: dose in milligrams (log scale); ○ at "P" designate placebo values. Data points show means of 14 subjects. Vertical bars show ± 1 S.E.M. and the absence of bars indicates 1 S.E.M. fell within the area of the data symbol. Filled symbols indicate a significant difference from placebo ($P \leq .05$). Letters "a", "b" and "c" indicate comparisons among the six drug doses. Within the same panel, any two means designated with the same letter are not significantly different from each other at $P \leq .05$ (Tukey's *post hoc* tests).

Too Much Information?

- Multiple families of dependent measures
  - Self-reported mood
  - Physiological
  - Observer-rated
  - Performance
- Multiple measures within each “family”
- Multiple drug doses
  - Placebo + 2 or more active doses
- Multiple time points

Validity of Measures

- Predictive validity
  - Do our measures predict (correlate with) real world outcome?
- External validity
  - Can our outcomes be generalized to other real-world groups and conditions?
  - Significant risk with small sample sizes
- Construct validity
  - Do our measures assess abuse liability as a construct?
- Participant motivation
  - Does this vary between groups?

So Many Tests, So Little Time

- Experimentwise or familywise error rate
  - Bonferroni
- But, all of outcomes essentially measure same thing
  - Do not correct when tests are related (Motulsky, 1995)
- Life time error rate

Discussion Point

How do we analyze significant interactions and dose effects, and compare one drug to another?

What's Really Different?

- Orthogonal comparisons
  - Each comparison provides independent information
- Limit number of comparisons to the degrees of freedom
- Planned comparisons
  - Hypothesis driven
  - Key feature is *a priori* nature, not independence (Winer, 1962)
  - Do not require a significant main effect or interaction and the omnibus F results should not be presented
- Post-hoc or unplanned comparisons (data sifting)
  - Significance level must be adjusted

Discussion Point

When is water-boarding an appropriate statistical test?

Data Reduction

- Simple dichotomous (nominal) outcome
  - Abuse liability YES or NO
  - Outcome not actually measured in study
- Multiple regression to predict abuse
  - Problem is predictors are related
    - » Ice cream sales predict drownings

A better understanding of the structure of the relationships among the dependent variables could be used to guide data reduction

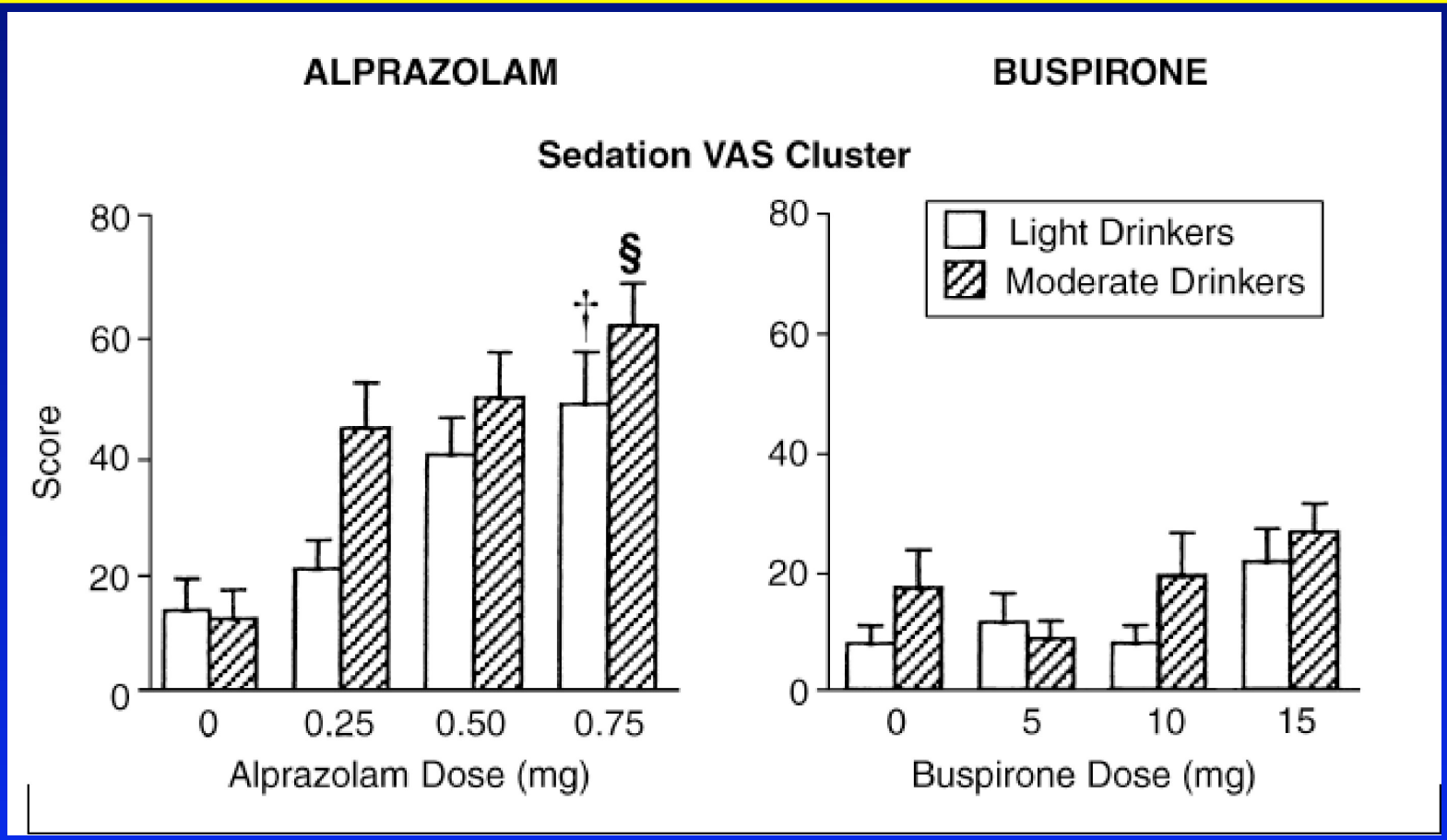
Reduction Approaches

- Factor analysis or principal components analysis
  - Like doing regression with multiple variables
  - Maximizing variance of the relationship among your variables (Eigenvalues)
  - Each subsequent factor maximizes relationship on the variance that remains, i.e., is orthogonal, and accounts for less variance
- Cluster analysis
  - Grouping objects of similar kind into respective categories
  - Makes no assumptions about structure
  - Maximizes association of objects within a group and minimizes association of objects between groups
  - Exploratory tool

Marijuana/THC, Alprazolam, Cocaine Clusters

- Personality
  - Self-confident, friendly
  - Not sensitive
- Physical complaints
  - Nausea
- Sedation
  - Tired, sedated
  - MJ includes bad drug effect
- On edge/anxious or Negative Affect
  - MJ has Irritability cluster & anxious cluster
- Depressed/confused (MJ)
- Good Drug Effect
  - Stimulated, High, Good Drug Effect
  - APZ adds forgetful
- Quality of Drug (cocaine)
  - Quality, potency, liking

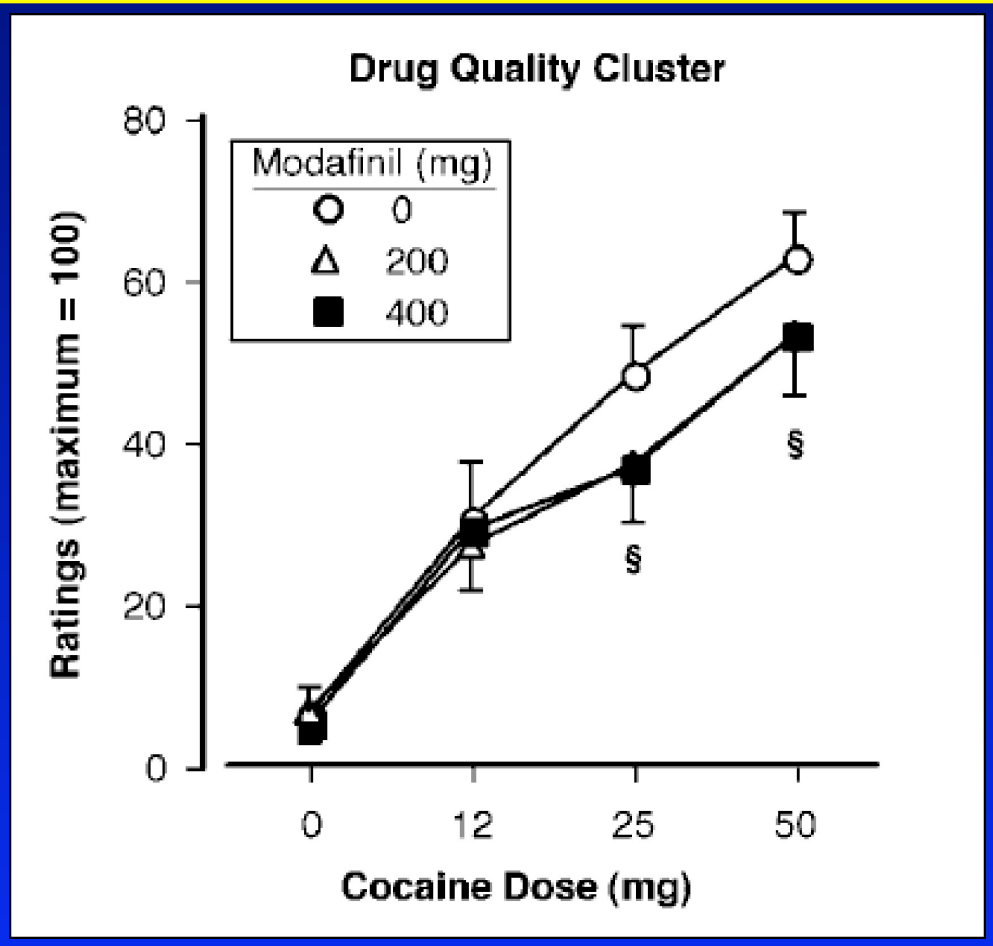
Alprazolam: Sedated Cluster



S.M. Evans and F.R. Levin

The effects of alprazolam and buspirone in light and moderate female social drinkers *Behavioural Pharmacology* 2002; **13**:1–14

Modafinil: Cocaine Quality

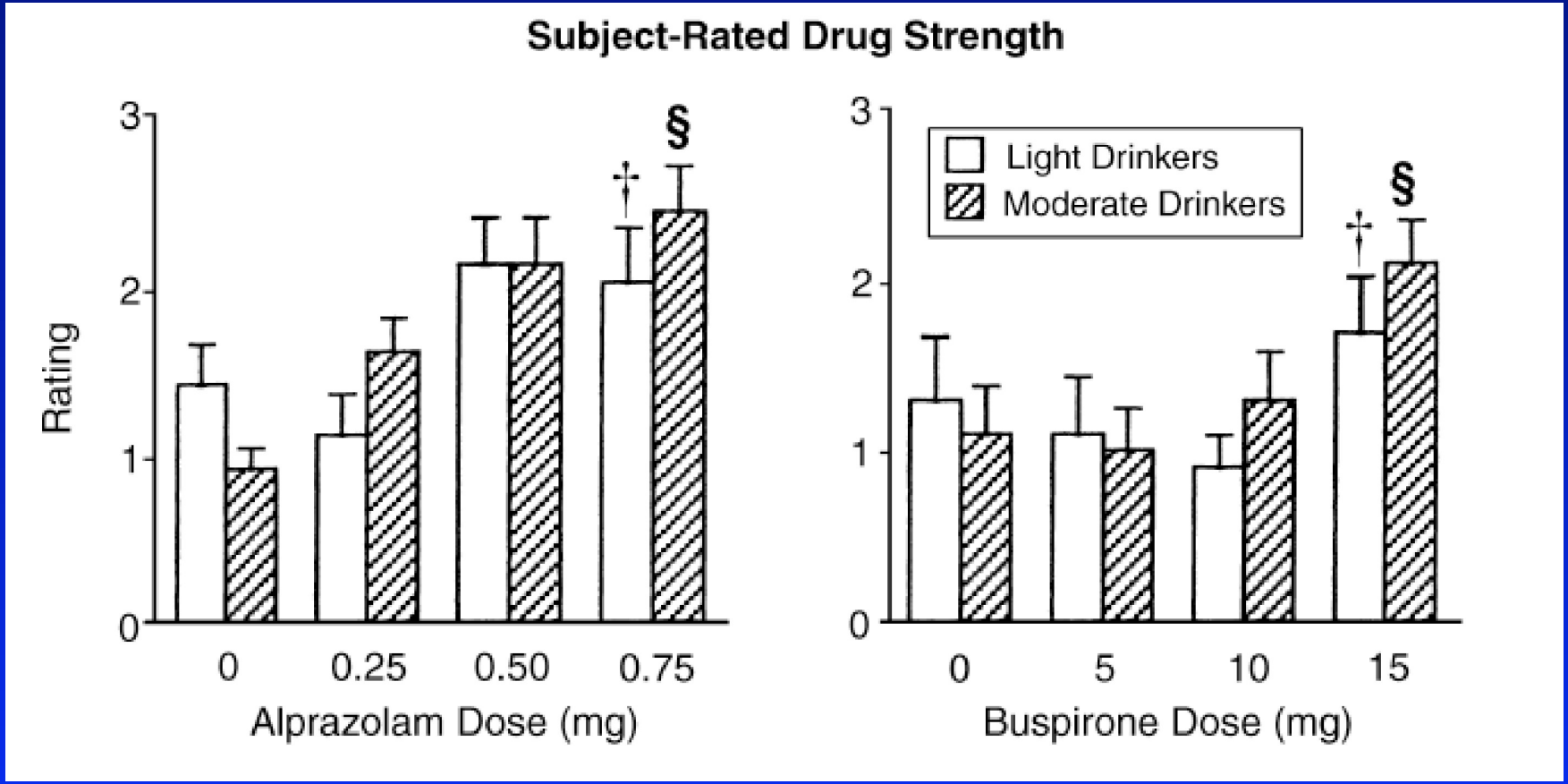


Hart, C.L., Haney M., Vosburg, S.K, Rubin, E., & **Foltin R.W.** Human smoked cocaine self-administration is decreased by modafinil, Neuropsychopharmacology33:761-768, 2008.

Discussion Point

What techniques should be used to decrease the number of dependent variables, if any?

Alprazolam: Drug Strength



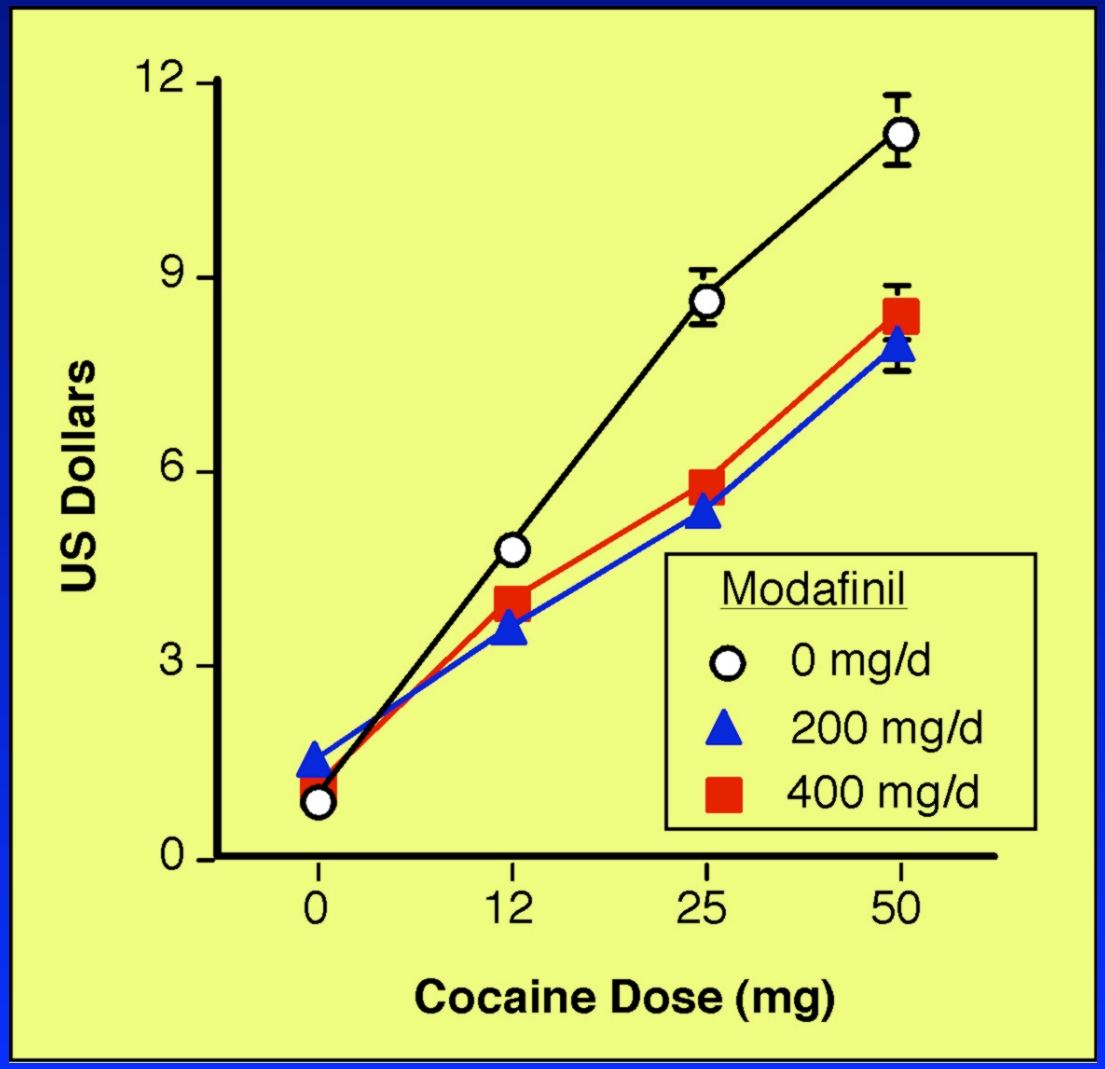
Discussion Point

How do we demonstrate that an appropriate dose range was tested?

Discussion Point

Could we, if we had a concurrent measure of equieffective doses, use a single dependent variable to predict abuse liability?

Modafinil: "Pay"



Neuropsychopharmacology (2008) 33, 1179–1191

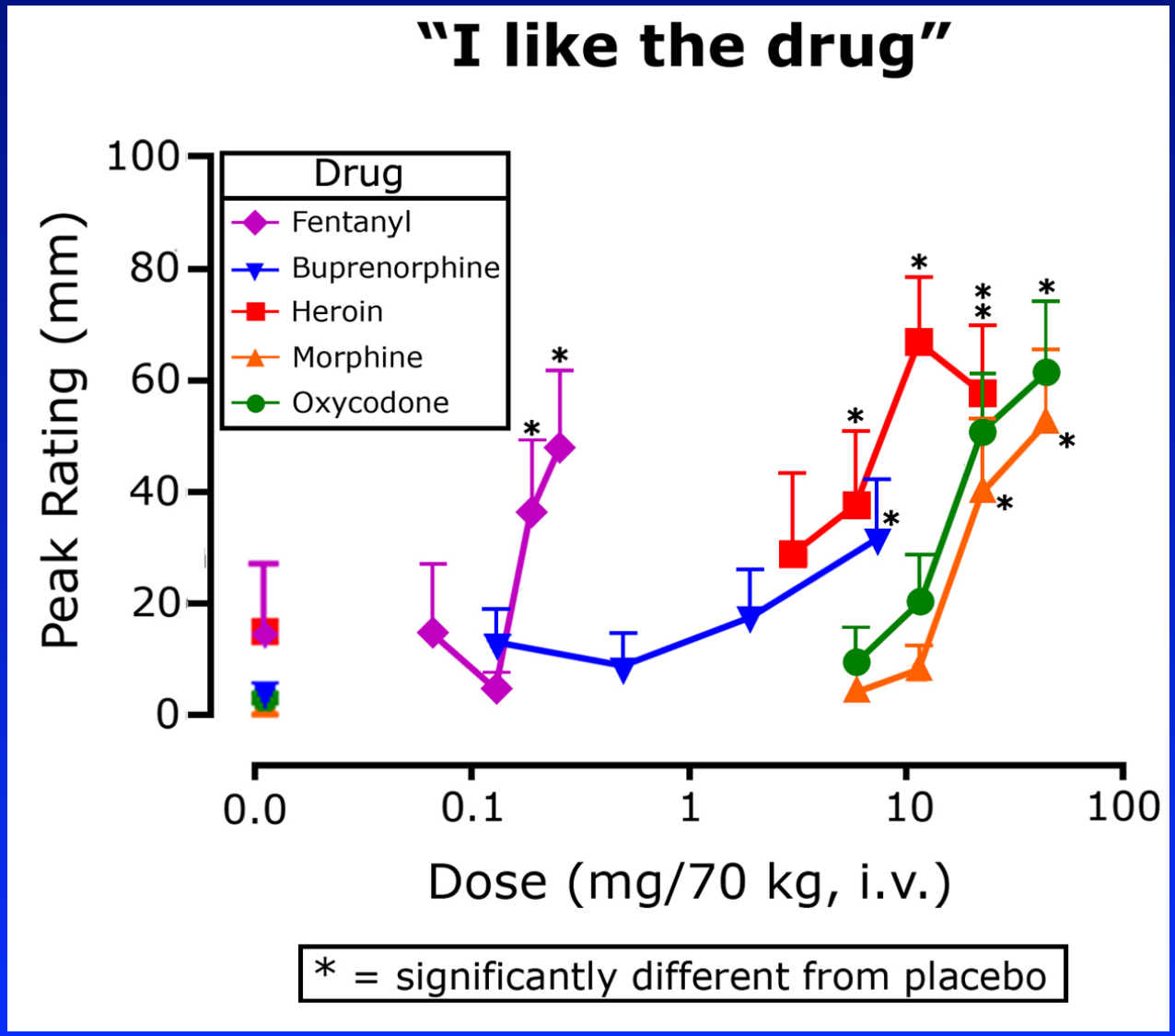
www.neuropsychopharmacology.org

Abuse Liability of Prescription Opioids Compared to Heroin in Morphine-Maintained Heroin Abusers

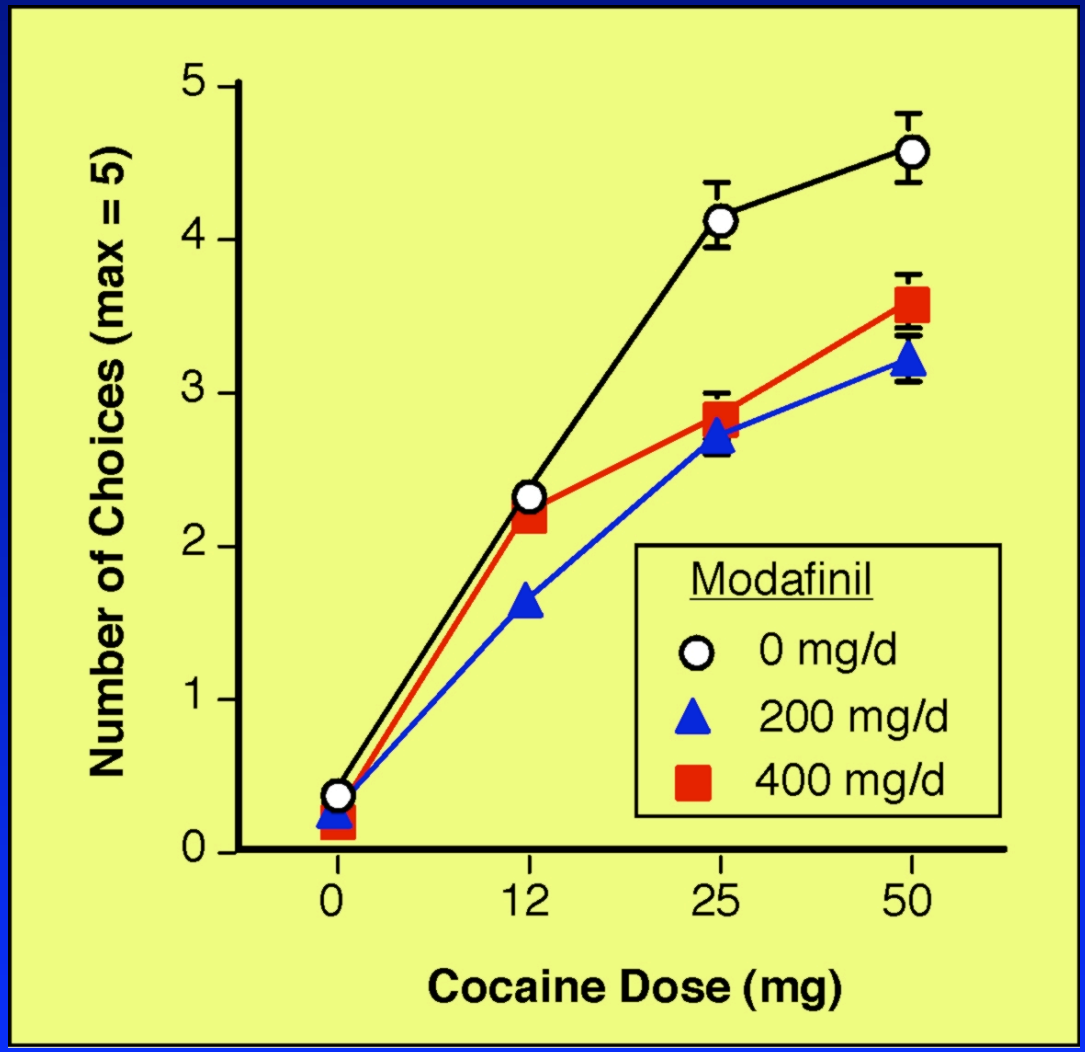
Sandra D Comer*[,1,2], Maria A Sullivan[1,2], Robert A Whittington[3], Suzanne K Vosburg[1,2] and William J Kowalczyk[1,2]

[1]Division on Substance Abuse, New York State Psychiatric Institute, College of Physicians and Surgeons of Columbia University, New York, NY, USA; [2]Department of Psychiatry, College of Physicians and Surgeons of Columbia University, New York, NY, USA; [3]Department of Anesthesiology, College of Physicians and Surgeons of Columbia University, New York, NY, USA

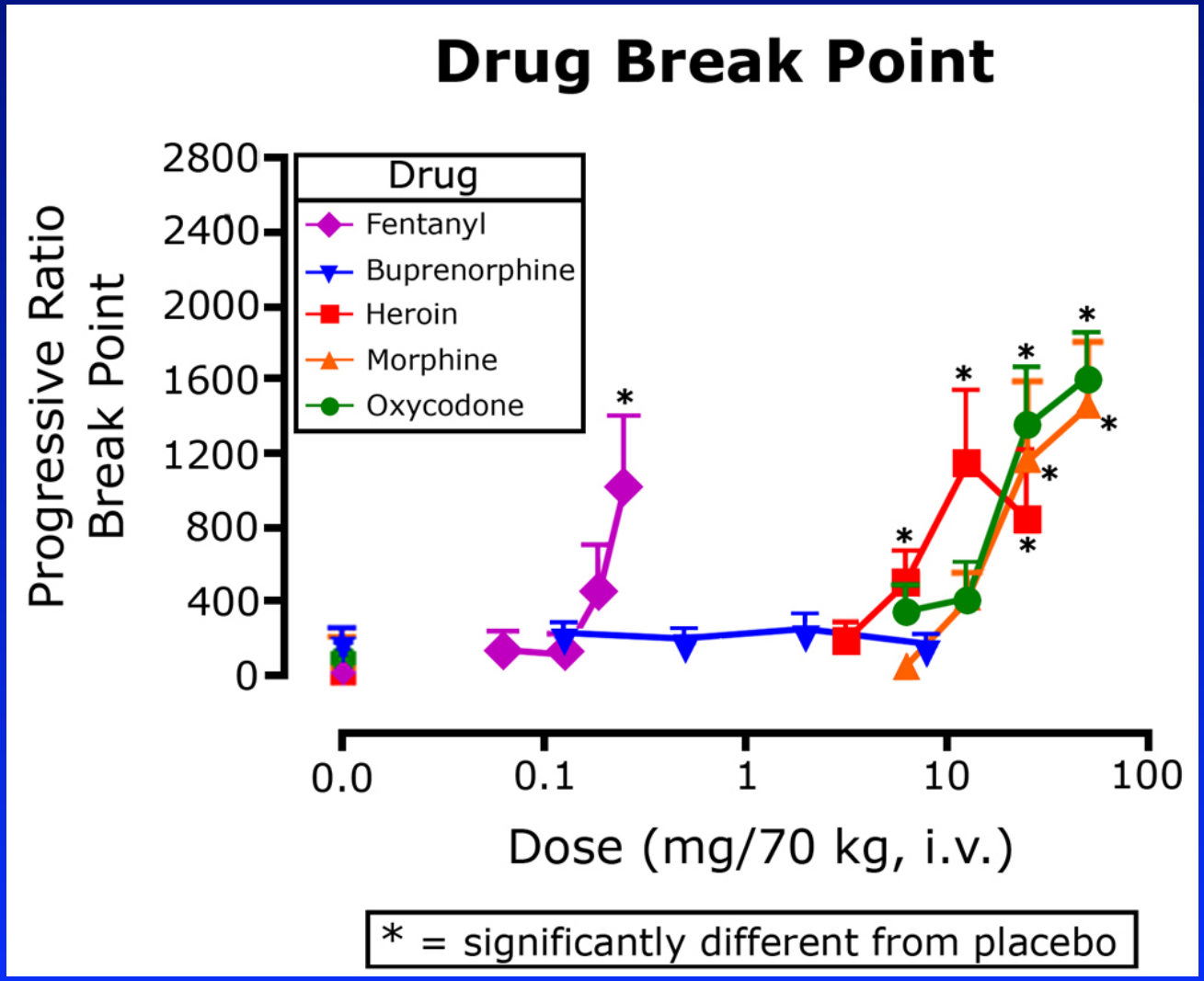
Drug Liking


"I like the drug"
Peak Rating (mm)
100
80
60
40
20
0
Drug
Fentanyl
Buprenorphine
Heroin
Morphine
Oxycodone
0.0
0.1
1
10
100
Dose (mg/70 kg, i.v.)
* = significantly different from placebo

Modafinil: Cocaine Choice (Each Cocaine Choice Cost $5 of Participant Earnings)



Reinforcing Efficacy with Alternative to Drug Taking



Where Do We stand?

- Happy with our sample
- Dealt with missing data
- Convinced ourselves that the data can be analyzed (approximately interval if using parametric statistics)
- Our measures are valid
- We did an experiment (random assignment to condition)
- Developed hypothesis-driven data questions (comparisons)
- Reduced our data to a reasonable number of measures

But, how do we draw a conclusion?

Discussion Point

Is comparison with a drug of established abuse potential the best approach, and how similar is a difference?

Abuse Liability Odds Ratio

- Odds ratio (OR) estimates strength of relationship between a variable and target outcome (retrospectively)
  - Ratio of the odds, not the percentages
- A drug for which liking scores (a derived variable based on reduction, e.g., liking, willingness to take again) are >70 (out of 100 after derived variables are standardized to 100) has 7 times greater odds of being abused than (the odds of) a drug with a liking score of <30.
  - Could compare doses within drug
- Not relative risk, which is the ratio of the probabilities of two events
  - Estimated using population samples

Discussion Point

What is the outcome measure?

Suggestions

- Develop a set of measures to use across laboratories
- Develop a data reduction scheme to reduce the measures to a few
  - This will address issues about scale of measurement
- Retrospectively assess external validity of previous studies to guide data reduction

- Develop a simple test for a positive signal

- Conduct an experiment to validate data reduction technique and predictive validity

Talking Points

- Designs
- Samples
- Scale of Measurement
- Correction for Multiple Tests
- Best Comparisons
- Data Reduction
- Effective Dose Range
- Single-measure Option
- Abuse Liability Criteria/ Outcome Measure