

The scientific possibilities of online experiments

Overview

- Online populations are diverse
- Replication of experimental results
- More data can be fundamentally different
- Case studies

Diverse Population

Ease of recruiting specific populations

- Online populations are more diverse.

Ease of recruiting specific populations

- Online populations are more diverse. It is easy to recruit populations of
 - People of different ages
 - Both into categories as well as a covariate

Ease of recruiting specific populations

- Online populations are more diverse. It is easy to recruit populations of
 - People of different ages
 - Both into categories as well as a covariate
 - People from different countries
 - Easy cross-cultural analyses between a few countries

Ease of recruiting specific populations

- Online populations are more diverse. It is easy to recruit populations of
 - People of different ages
 - Both into categories as well as a covariate
 - People from different countries
 - Easy cross-cultural analyses between a few countries
 - People with different native languages
 - Native vs non-native speakers analyses

Ease of recruiting specific populations

- Online populations are more diverse. It is easy to recruit populations of
 - People of different ages
 - Both into categories as well as a covariate
 - People from different countries
 - Easy cross-cultural analyses between a few countries
 - People with different native languages
 - Native vs non-native speakers analyses
 - People with diverse musical experience and ability
 - Etc.

Replication and Power

Online experiments and replication

- Psychology is facing a crisis of increasing failures to replicate published studies
- Online experiments help solve this issue

Replicating Research Online

- Fast data collection decreases the opportunity cost of replicating experiments
- Replicating with large samples increases power and avoids some issues with significance testing
- Replicating existing studies before extending them helps avoid dead end research projects
- Our lab has begun self-replicating our own studies before publication

Online experiments and replication

1. Online experiments are easy to replicate
 - Designed to be cross-platform
 - Not specific to any lab environment
2. Online experiments allow for powerful replication of many lab-based experiments
 - Easy to adjust sample size based on initial effect size and expected power of the replication
 - Rapid data collection of large samples
 - Different population than many lab studies

Sampling

- Slow small sampling procedures encourage “sampling to significance”:
 - i.e. keep sampling until $p < 0.05$
 - This breaks the assumptions of every frequentist statistic: t test, ANOVA, regression, etc.

Large datasets are Different

Large samples fundamentally have more power

- High powered studies help mitigate the file drawer effect
- Reduce Type II errors for planned comparisons
- Even more important for conducting and interpreting post-hoc analyses of data

Large datasets are Different

Large samples allow for more between-subject conditions

- Switch from categorical to continuous independent variables
 - This can improve the interpretation of effects
- More independent variables can be manipulated
 - Large interaction analyses become tractable

Large datasets are Different

Large samples have different statistical properties that invite sophisticated analyses

- Individual differences analyses
- Cluster and network detection
- Outlier and contaminant detection
- Bootstrapping techniques
- Etc.

Case Studies

Large Sample Case Study

Large Sample Case Study

Traditional lab study (N = 66):

Condition	Measure
High	0.68
Low	0.51

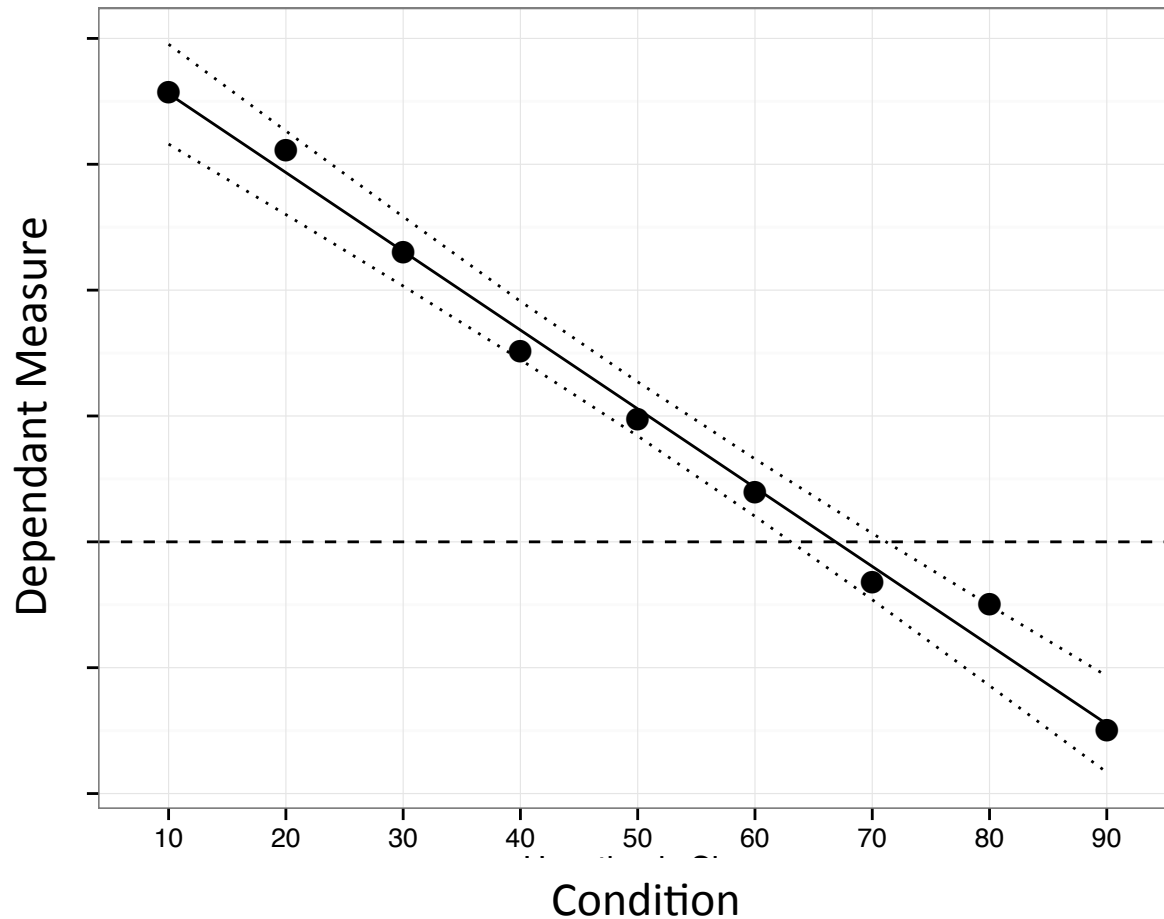
$$F(1, 65) = 25, p < 0.005$$

Large Sample Case Study

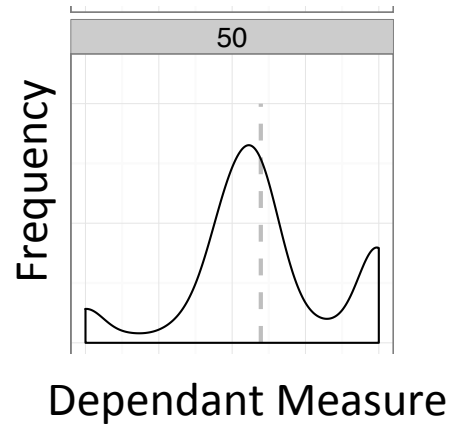
Online study (N = 400) with 9 conditions

Large Sample Case Study

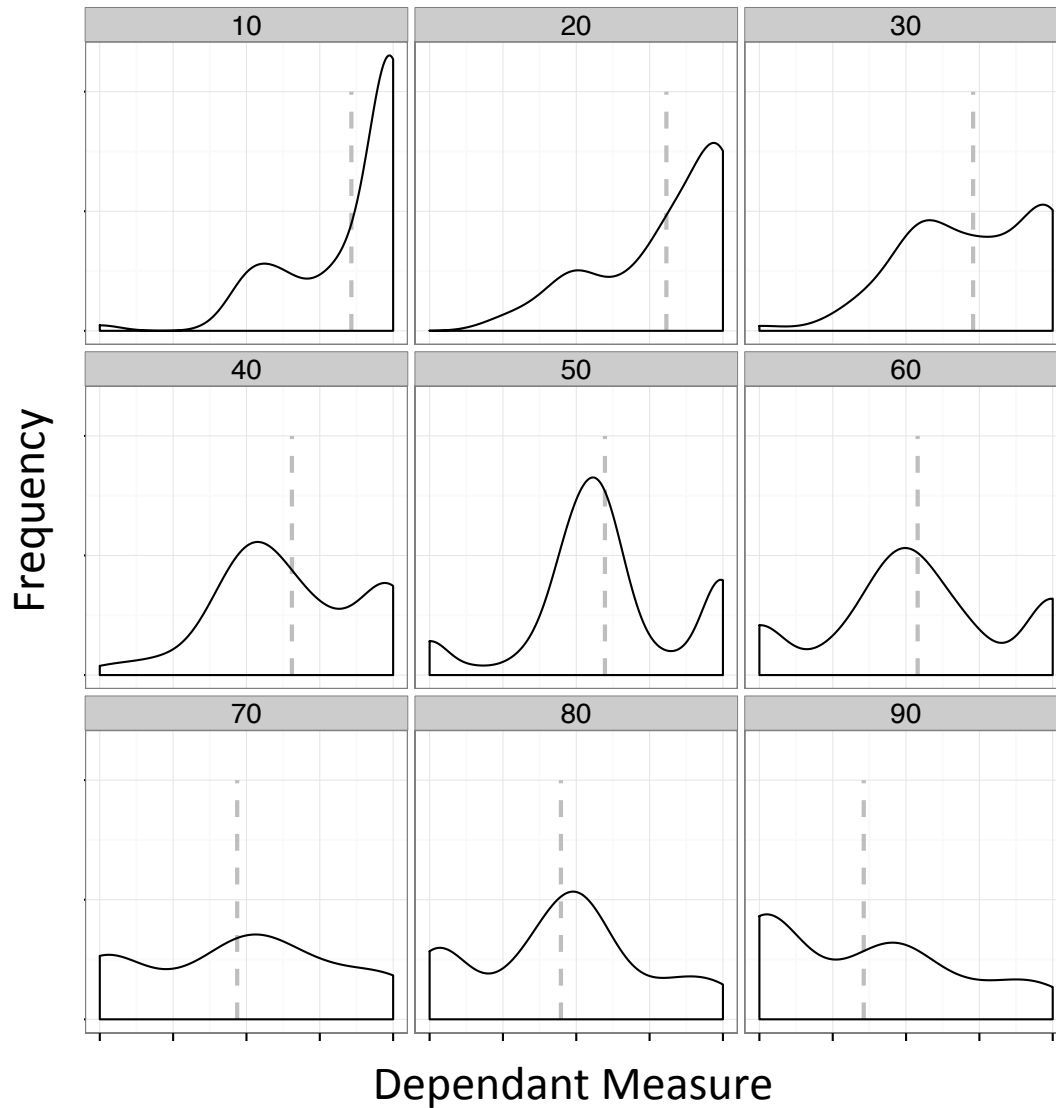
Online study (N = 400) with 9 conditions



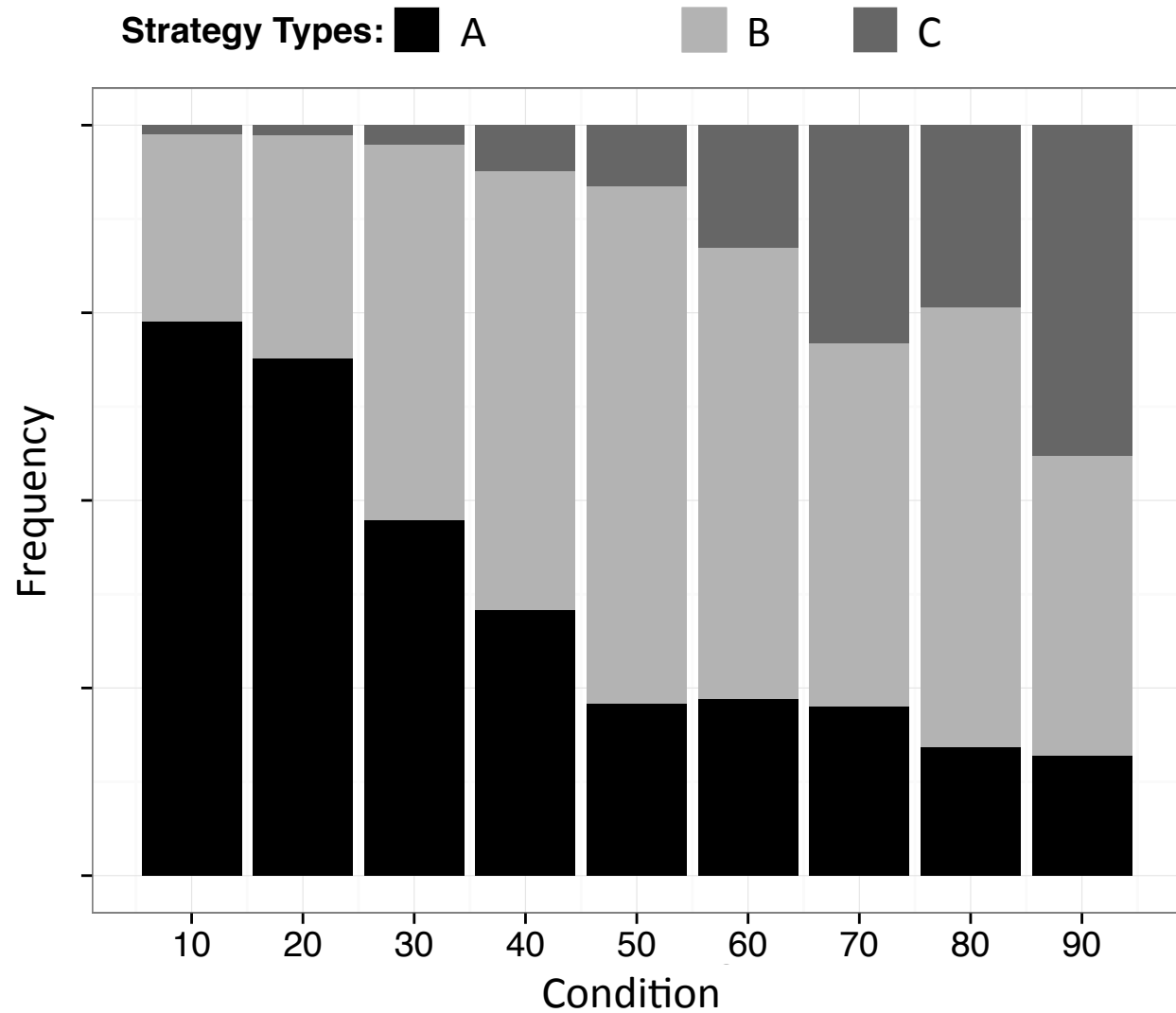
Large Sample Case Study



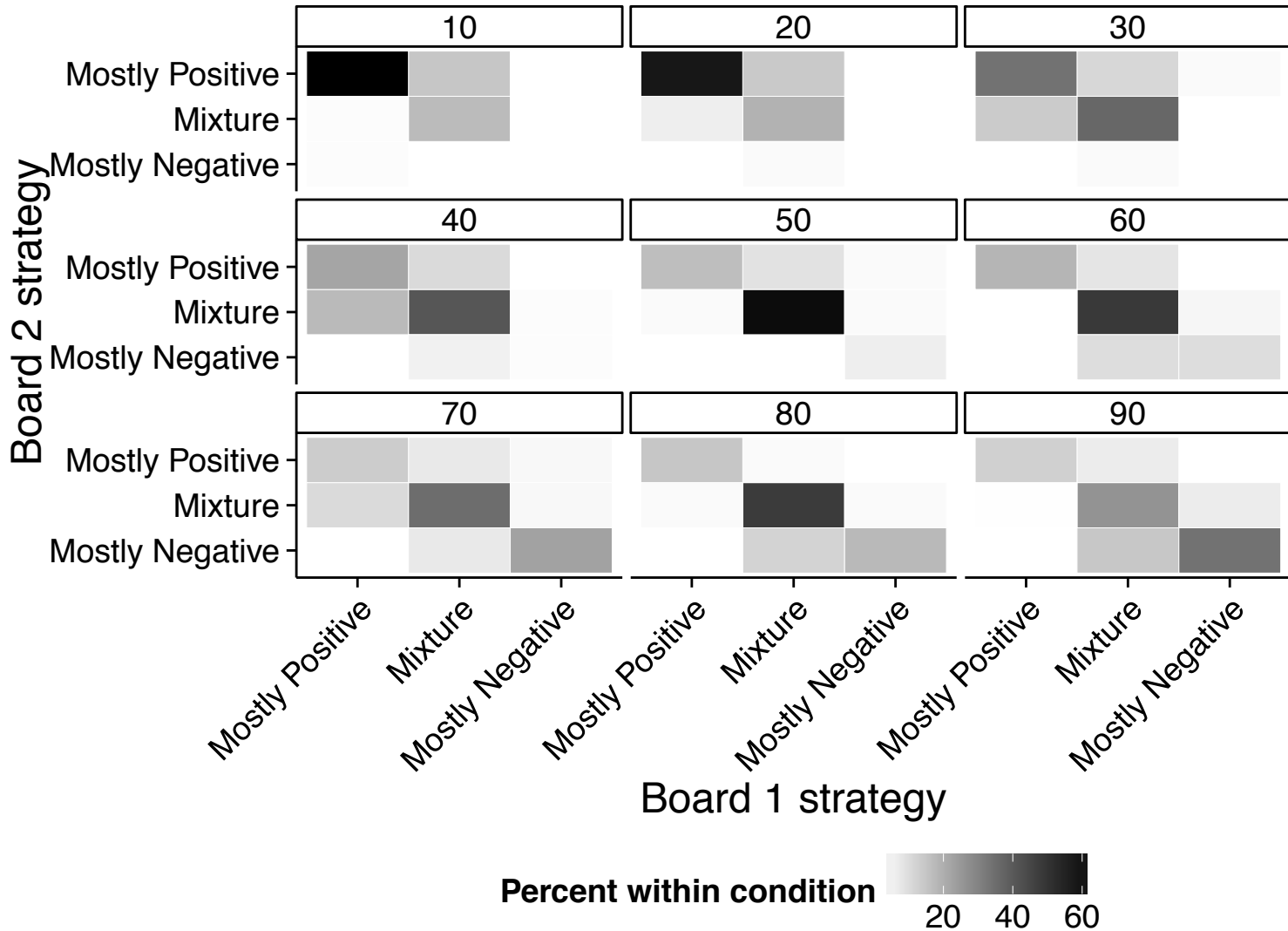
Large Sample Case Study



Large Sample Case Study



Large Sample Case Study

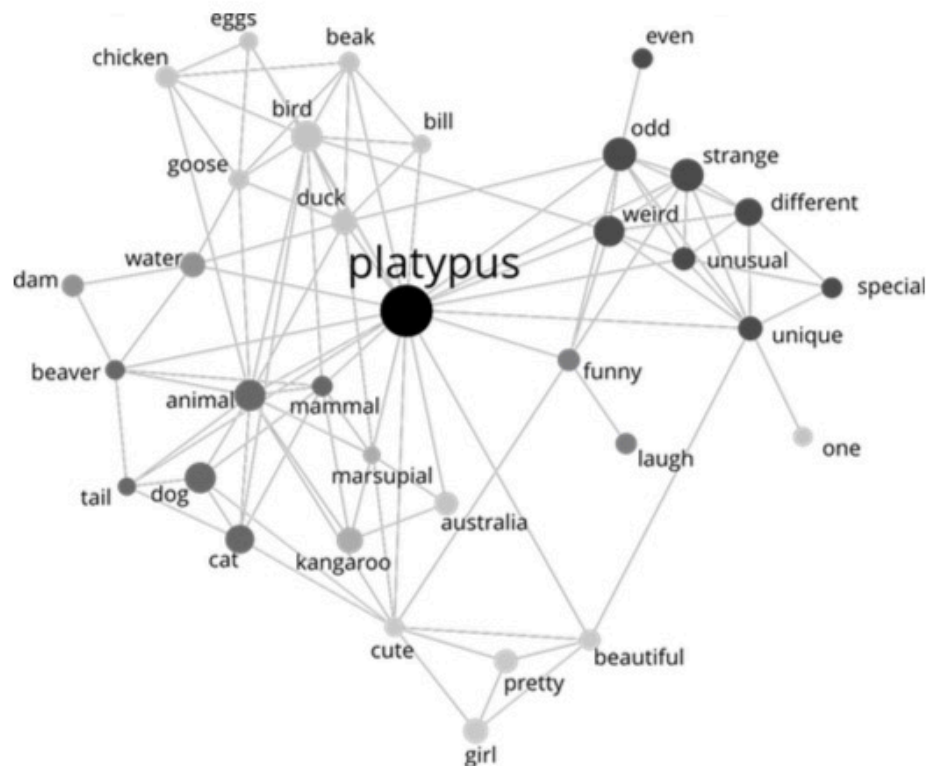


Large datasets can be fundamentally different

Example: the Dutch Word Association Lexicon
De Deyne, et al. (2013)

Word Associations

- Typical experiments use < 200 target words
- Create a small network of words



Word Associations

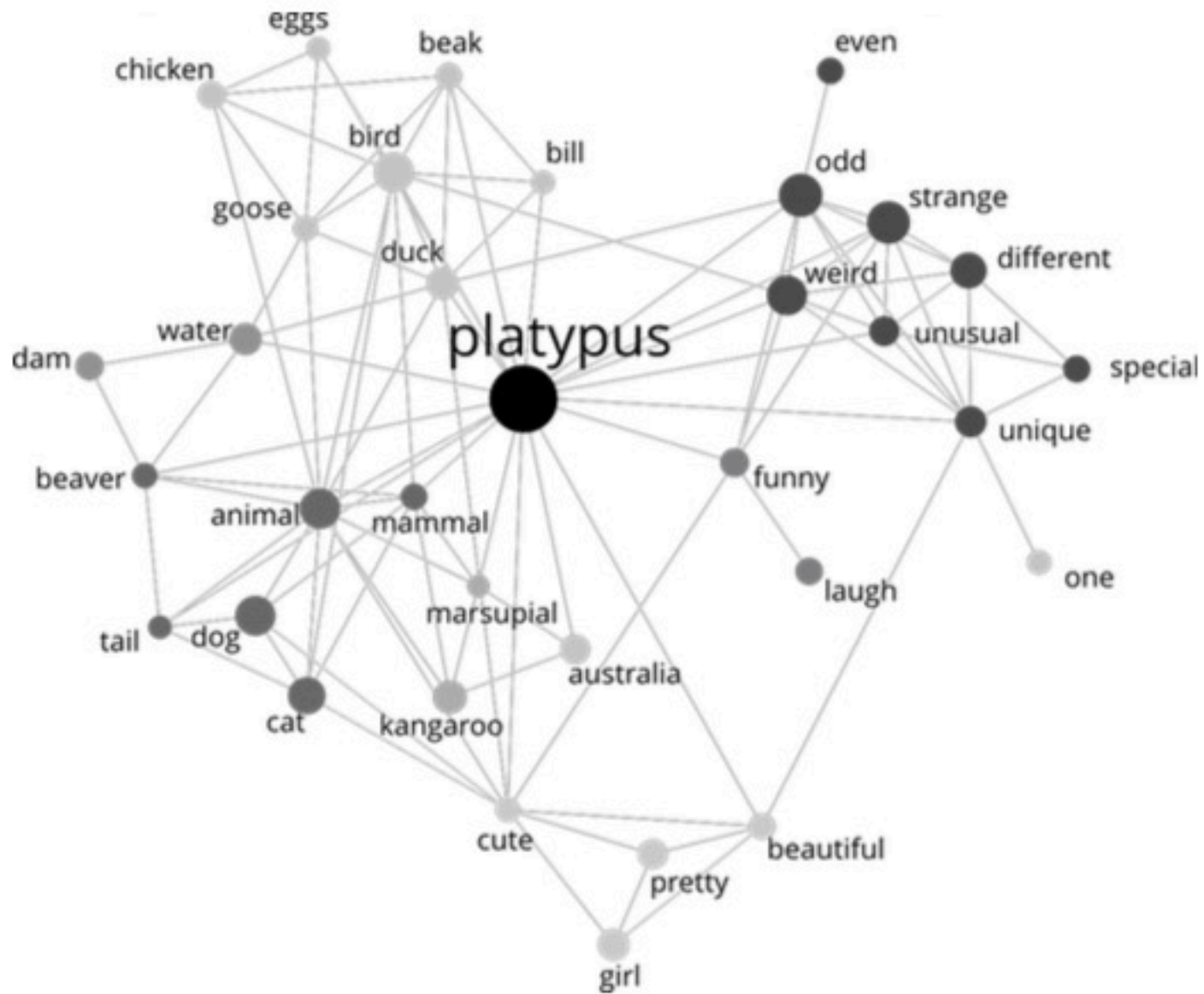
- Typical experiments use < 200 target words
- Create a small network of words
- Use distances in the network to predict many semantic aspects of language

Dutch Word Association Lexicon

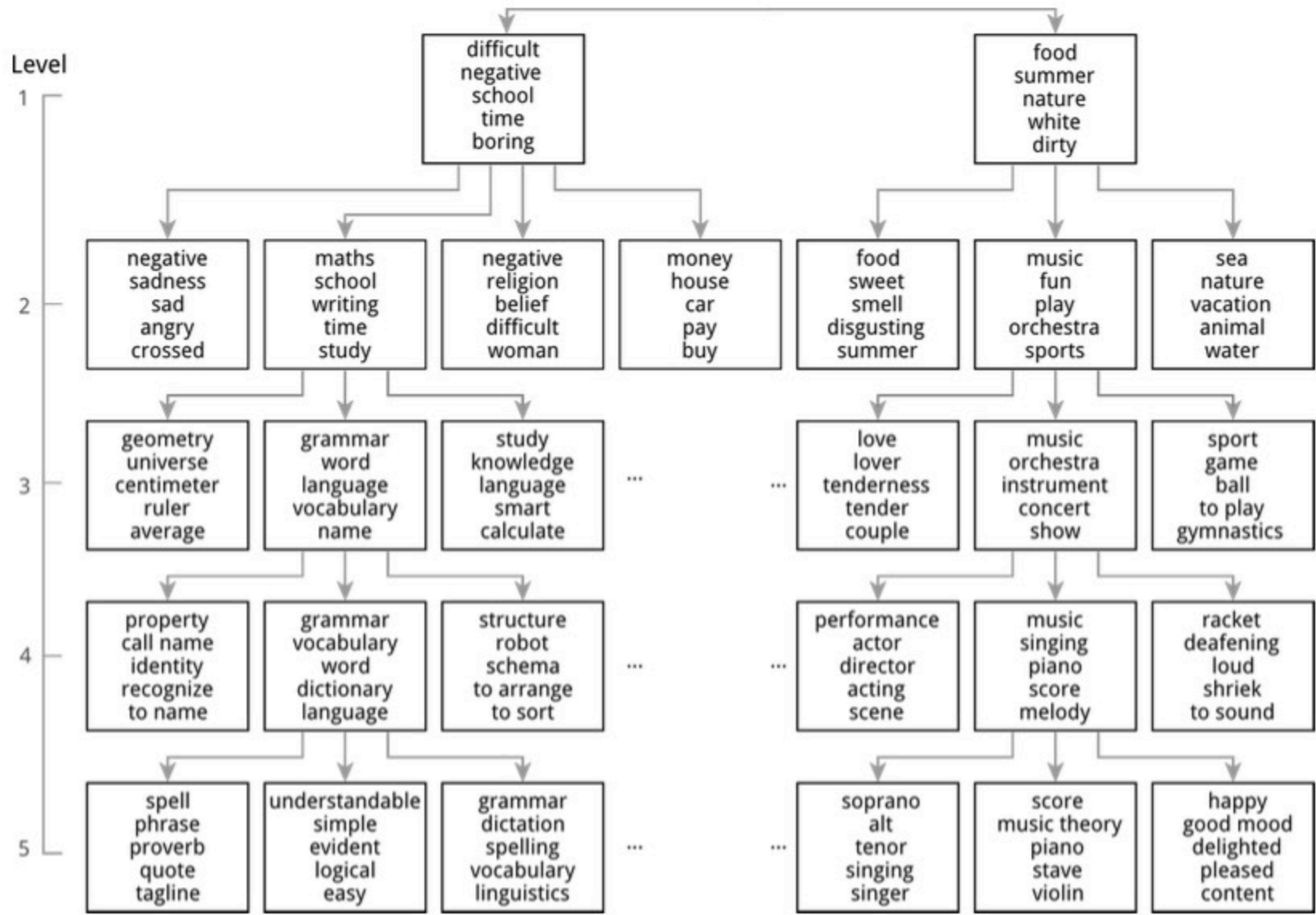
- Started with a set of 1,000 target words
- Recruited participants online (free)
 - Spread virally
- Added new target words based on popular participant responses (snowball procedure)

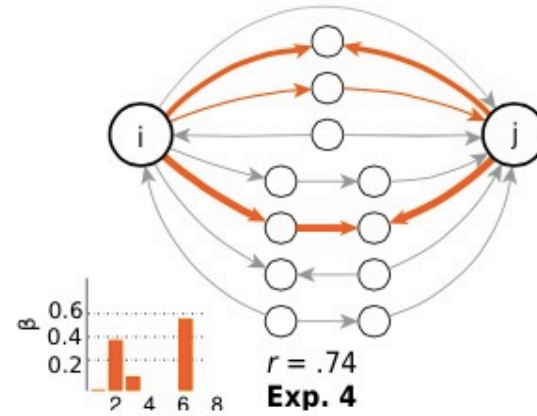
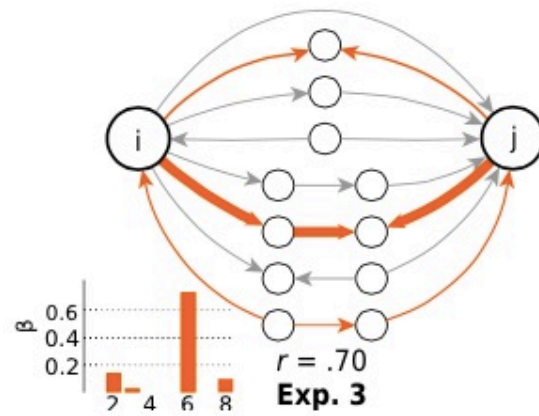
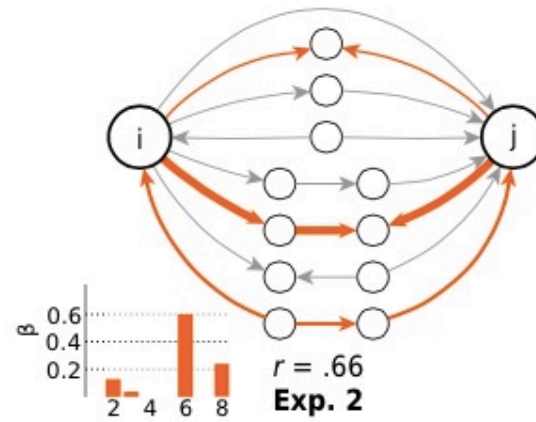
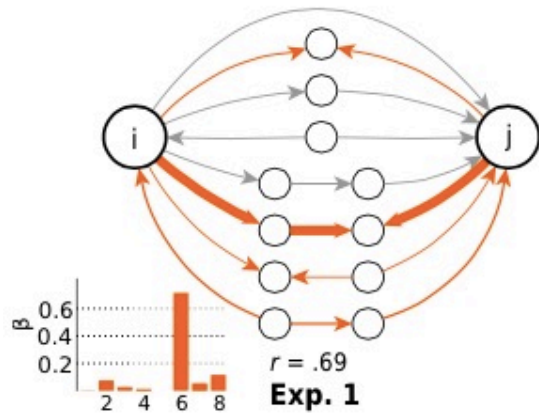
Dutch Word Association Lexicon

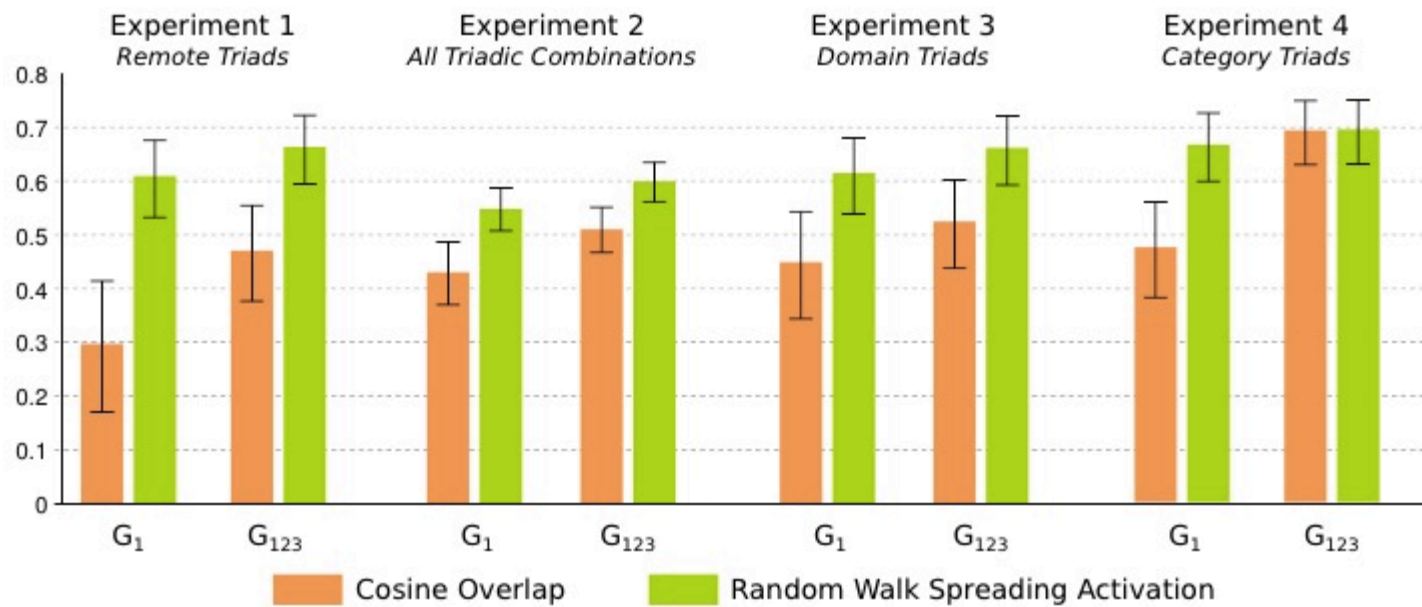
- Word associates recorded for 11,000+ unique target words
 - 3,000,000 unique responses
 - 70,000 unique participants
- Networks analyses using 11,000 words is fundamentally different than 200 words
 - More dense with connections
 - Sub-networks can be compared and contrasted



De Deyne, et al. (2013)







Online experiments enable scale

- Example: Inferring priors

Online experiments open up new possibilities

- Easy access to diverse populations
- Increased power of experiments
- Fundamentally different analyses