LIBR
Laureate Institute for Brain Research

Pragmatic Computational Psychiatry: Integrating Computational Approaches and Risk-Prediction Models

Martin P. Paulus
Crane Huang
Katia M. Harlé
Laureate Institute for Brain Research
6655 S Yale Ave, Tulsa, OK 74136-3326
mpaulus@laureateinstitute.org
http://www.laureateinstitute.org

Outline

- The Challenges Biological Psychiatry
  - The impact of neuroscience on Clinical Psychiatry
  - Solving good problems
  - Complexity of psychiatric disorders
- Pragmatism and Risk Prediction Models
- Examples - Predicting Outcomes
  - Predicting Development of Stimulant Problem Use
  - Utility of Computational models
- Future Directions

On the next slide is a list of major discoveries in neuroscience research that have changed clinical psychiatry…

LIBR

Laureate Institute for Brain Research

LIBR
Laureate Institute for Brain Research

Failed Promises of Biological Psychiatry

- Despite profound advances from molecular to systems neuroscience, these insights have had virtually no influence on practical psychiatry.
- The development of new therapeutics based on neuroscience approaches to understanding the pathophysiology of these illnesses has stalled (T. R. Insel, 2012).
- Third, with the development of a new diagnostic classification for mental disorders (APA, 2013), neuroscience had virtually no impact on the disorder categories.
- There are no clinical tools for prognosis, diagnosis, treatment monitoring that are based on neuroscience approaches (Prata, Mechelli, & Kapur, 2014).

LIBR
Laureate Institute for Brain Research

Possible Reasons for the Failure to Deliver:

- Are mental health conditions are fundamentally not reducible to biological processes.
- Do we have sufficiently developed technologies and approaches to map psychiatric diseases onto biological processes?
- Making biology useful for clinical psychiatry is an extremely difficult problem.
- Operational, institutional, and procedural aspects of biological research in psychiatry have not provided the appropriate environment and incentives to develop biological approaches that solve clinical problems.
- The focus on a misguided search for “mechanisms” that underlie psychiatric illnesses has hindered progress.

Mental Health Practice in 2015:

Presentation:

"Depressed Mood"

Interview Assessment

Self Report

DSM 5 Diagnosis

Intervention:

- Cognitive Behavioral Therapy
- Medication Treatment
- Lifestyle Changes
- Device-based Treatment

?

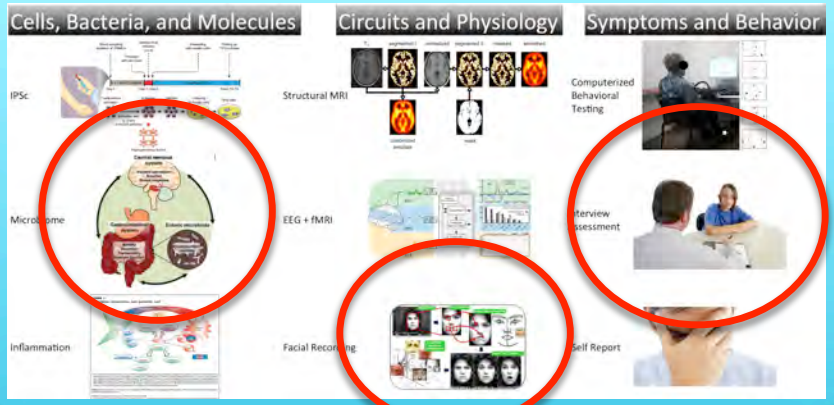
LIBR
Laureate Institute for Brain Research

Mental Health Practice of the Future:

Presentation:

"Depressed Mood"

Multi-level Assessment:



Intervention:

Cognitive Behavioral Therapy

Medication Treatment

Lifestyle Changes

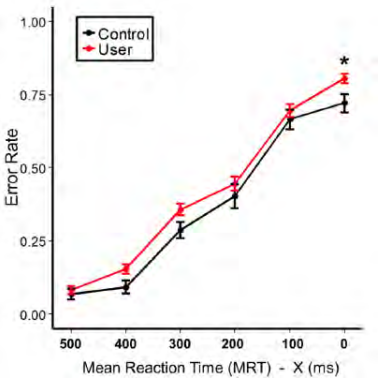
Device-based Treatment

LIBR
Laureate Institute for Brain Research

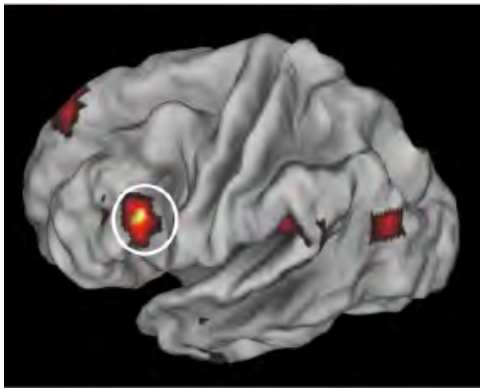
Computational Psychiatry

Old Approach:

Behavior

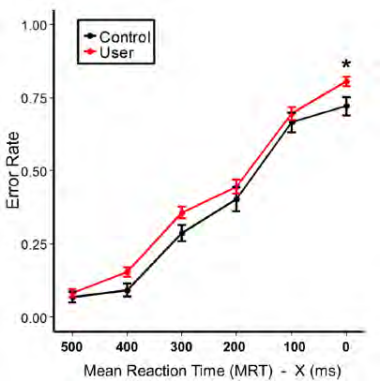


Brain Processing

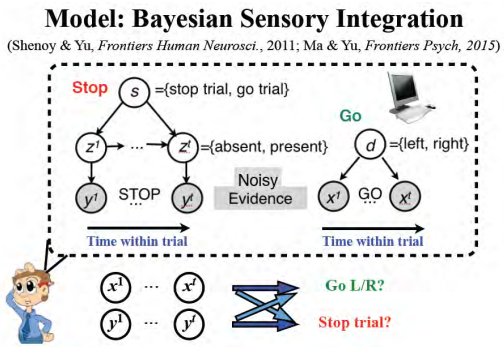


New Approach:

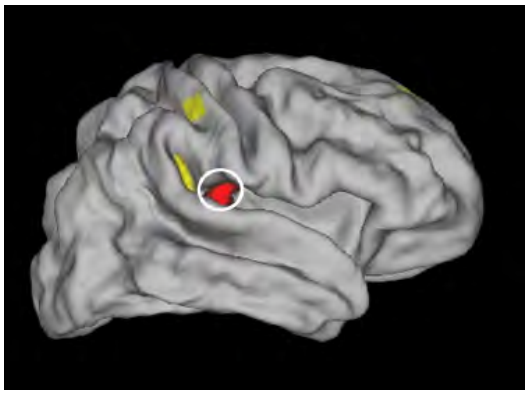
Behavior



Processing Model:



Model-derived Brain Processing



We need YOU!

We also need GOOD problems!

Research report

Depression is not a consistent syndrome: An investigation of unique symptom patterns in the STAR*D study

Eiko I. Fried [a,*], Randolph M. Nesse [b]

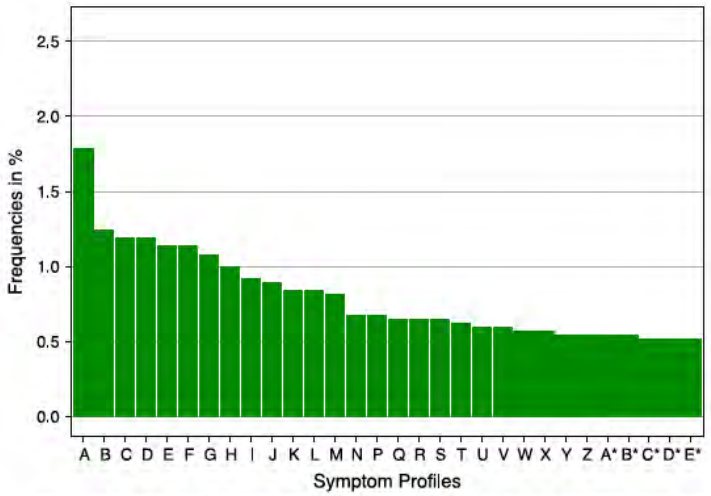


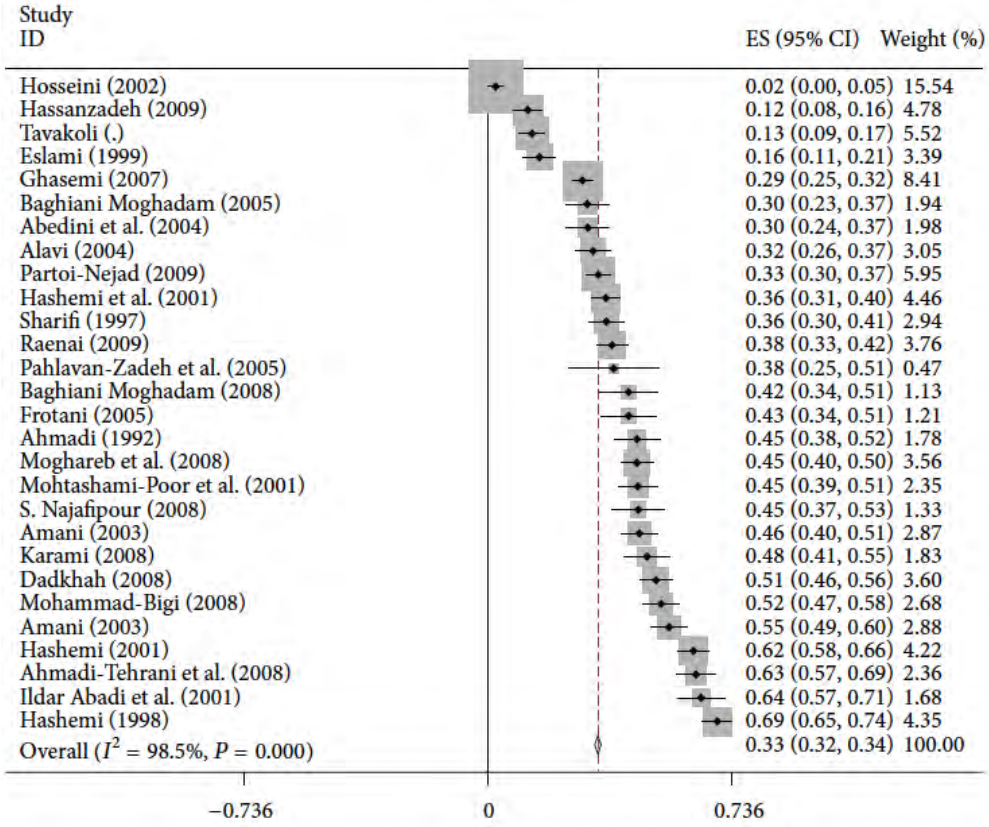
Fig. 1. Frequencies of the 30 most common depression symptom profiles during the beginning of the first treatment stage of the STAR*D study (n=3703).

Table 2
Detailed information about the 10 most frequent symptom profiles.

	Sad	Ene	Con	Ins	Int	App	Bla	Wei	Agi	Ret	Sui	Hyp	Freq(%)	Profile description
A													1.78	No symptoms
B	x	x	x	x	x	x	x	x	x	x			1.24	All but Sui and Hyp
C	x	x	x	x	x	x		x					1.19	Mixed profile
D	x	x	x	x	x	x	x	x					1.19	Mixed profile
E	x	x	x	x	x								1.13	Mixed profile
F	x	x	x	x	x		x						1.13	Mixed profile
G				x									1.08	Only Ins
H	x	x	x	x	x	x	x	x	x				1.00	All but Ret, Sui and Hyp
I	x	x	x	x									0.92	Mixed profile
J	x	x	x	x	x	x		x	x	x			0.89	All but Hyp, Bla and Sui

Prevalence of Depression among University Students: A Systematic Review and Meta-Analysis Study

Diana Sarokhani,[1,2] Ali Delpisheh,[3,4] Yousef Veisani,[1,4] Mohamad Taher Sarokhani,[1,5] Rohollah Esmaeli Manesh,[6] and Kourosh Sayehmiri[4,7]



LIBR
Laureate Institute for Brain Research

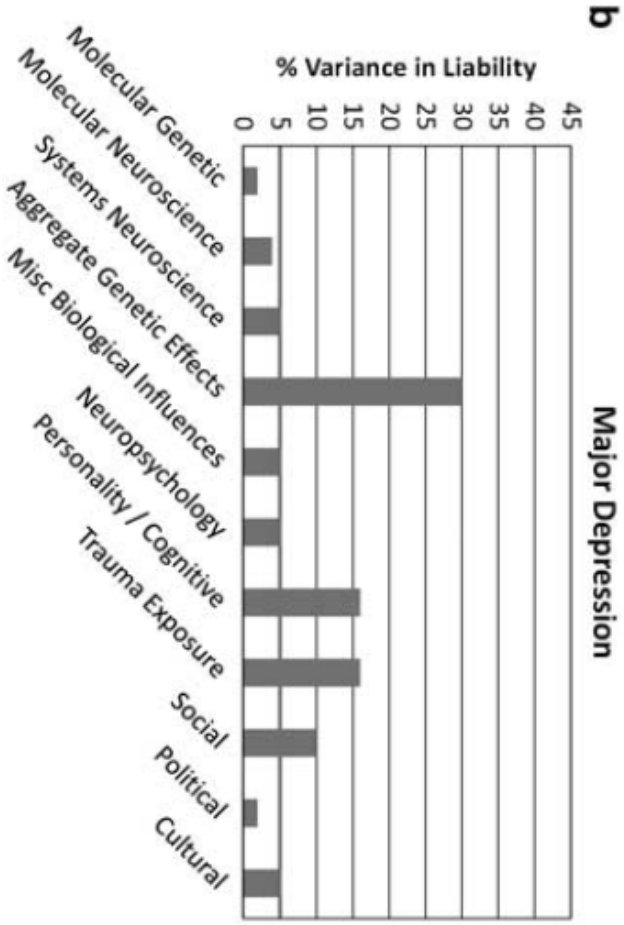
The Complexity of Psychiatric Disorders

Kendler's Difference Makers (Kendler 2012):

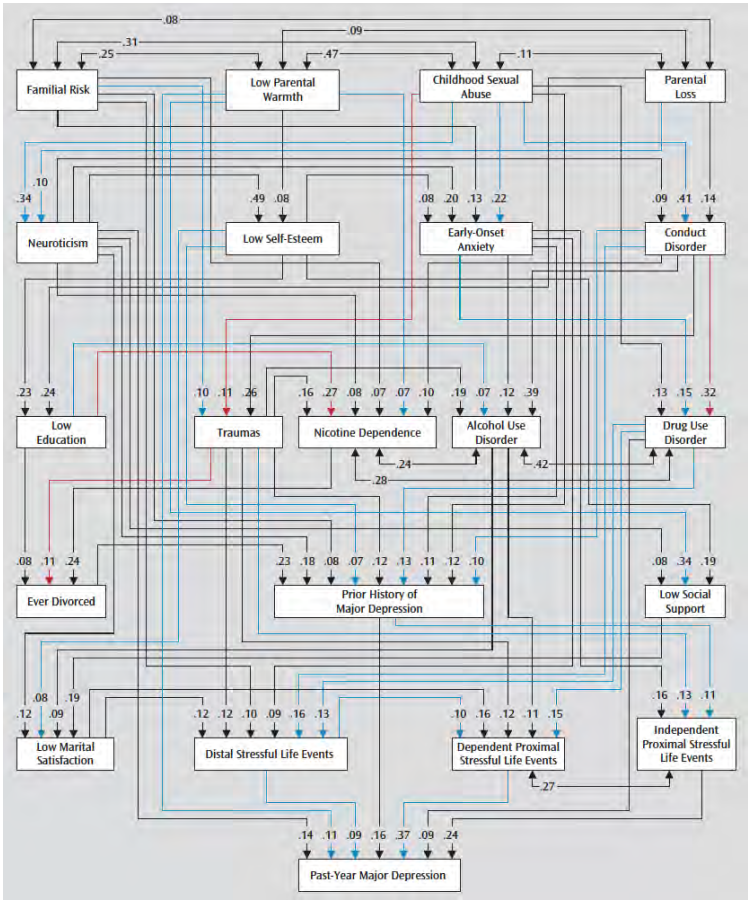
- Being exposed to a *difference-maker* increases the probability of illness but the difference-maker need not be necessary or sufficient.
- Difference-makers are distributed across:
  - biological,
  - psychological
  - and social–cultural domains, and
  - these levels are actively inter-twined with each other in etiologic pathways.

Example – Major Depressive Disorder

- "The commonly expressed wish to develop an **etiologically based nosology** for psychiatric disorders is **deeply problematic**."



The complexity of depression: (Kendler and Gardner 2014)



- *Our findings are broadly congruent with a typology of major depression developed from a psychoanalytic perspective by Blatt (61)….*
- *Blatt proposed that major depression takes two forms:*
  - *"anaclitic" and*
  - *"introjective."*
- *The former arises from deficiencies in caring relationships and unmet dependency needs (e.g., "I am unlovable"), and the latter emerges from the inability to meet internal demands for self-worth and achievement (e.g., "I am a failure").*

LIBR
Laureate Institute for Brain Research

“IT’S DIFFICULT TO MAKE PREDICTIONS, ESPECIALLY ABOUT THE FUTURE”

Niels Bohr, Samuel Goldwyn, Robert Storm Petersen, Yogi Berra, Mark Twain, Nostradamus, Anonymous

Risk Prediction Models

- Risk prediction models use:
  - predictors (covariates) to estimate
  - the absolute probability or risk that a certain outcome is present (diagnostic prediction model)
  - or will occur within a specific time period (prognostic prediction model) in an individual with a particular predictor profile (Moons, Kengne et al. 2012)

Components of Risk Prediction Models (Gerds, Cai et al. 2008)

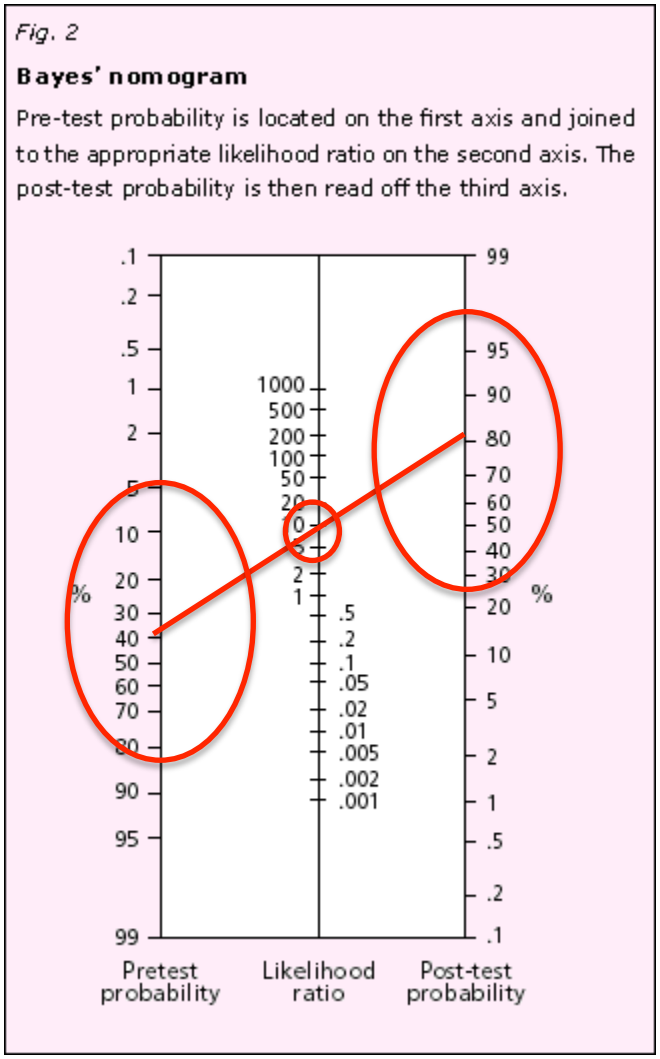
- Samples of *n* subjects
- For each subject *k* markers
- An individual subject status at some later time *t*
- A model that takes the sample and markers and assigns a probability *p* of the status at time t for each individual.
- Risk prediction models can be derived with many different statistical approaches.
- To compare them, measures of predictive performance are derived from ROC methodology and from probability forecasting theory.
- These tools can be applied to assess single markers, multivariable regression models and complex model selection algorithms (Gerds, Cai et al. 2008).

How does this work in Practice?

- Fagan's Nomogram:
  - Pre-test probability
  - The likelihood ratio (positive / negative)
  - Post-test probability

- *Determines how a test helps to make better predictions.*

Fagan TJ (1975) Letter: Nomogram for Bayes theorem. N Engl J Med 293: 257.



Prediction is not equal Significance

Why significant variables aren't automatically good predictors

Adeline Lo[a], Herman Chernoff[b,1], Tian Zheng[c], and Shaw-Hwa Lo[c,1]

- "What constitutes a good variable for classification and what constitutes a good variable for significance depend on different properties of the underlying distributions."

The key difference between finding a subset of variables to be highly significant versus finding it to be highly predictive is that the former uses assumptions on the exact distributions of the variables, whereas the latter requires knowledge of both **control and case distribution**.

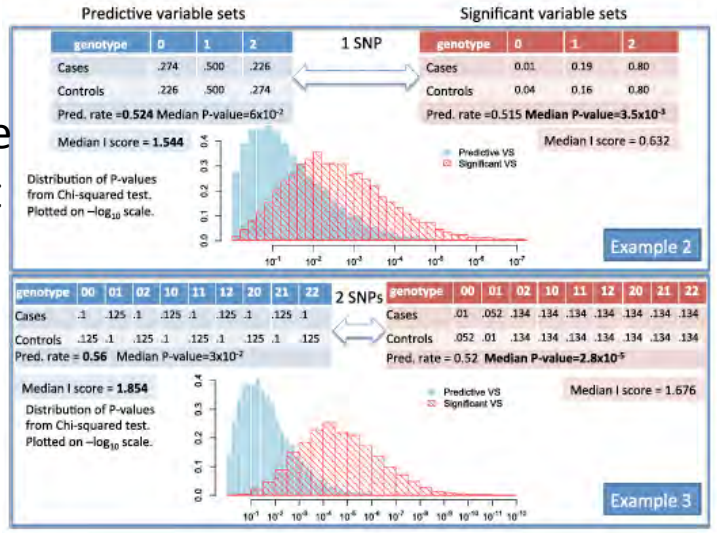


Fig. 2. Reversals of predictive and significant variable sets in SNP examples. Example 2 has one explanatory variable (1 SNP) for which the probabilities under cases and controls are listed in the tables. Example 3 has two explanatory variables (2 SNPs) for which the probabilities under cases and controls are listed in the tables. Left-hand-side tables (in blue) are for more predictive variable sets, whereas right-hand-side tables (in red) are for more significant variable sets. The prediction rate (proportion of correct predictions) of each variable set (of size 1 or 2) can be directly computed using the genotype frequencies specified. Using sample sizes of 500 cases and 500 controls, we simulate $B = 1{,}000$ random case-control data sets by simulating genotype counts among cases and controls using the genotype frequencies specified. I score and the χ^2 test statistic were computed for each simulated data set. Simulation details can be found in the *Supporting Information*.

Predicting Problem Use:

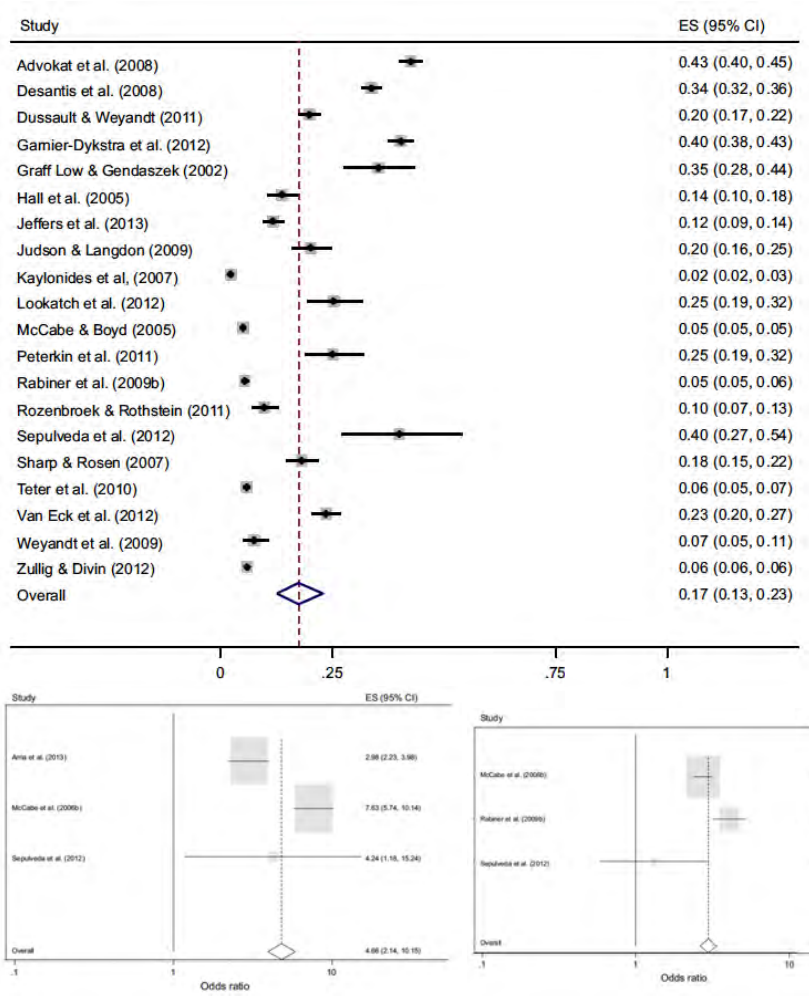
Bayesian Neural Adjustment of Inhibitory Control

Predicts Emergence of Problem Stimulant Use

Katia M. Harlé[1], Jennifer L. Stewart[2], Shunan Zhang[4], Susan F. Tapert[1,3],
Angela J. Yu*[4], Martin P. Paulus*[1,3,5]

Stimulant Use

- **17% of college student misuse stimulants** Benson, Flory et al. 2015.
  - Increase risk for
    - Alcohol related problems
    - Marijuana use
- 59% of current prescription holders had given away or sold their stimulant medication during their lifetime. Gallucci, Martin et al. 2015
- Self-reported non-medical use correlated with levels of anxiety and stress, internal restlessness, and perceived safety of the medications. Verdi, Weyandt et al. 2014



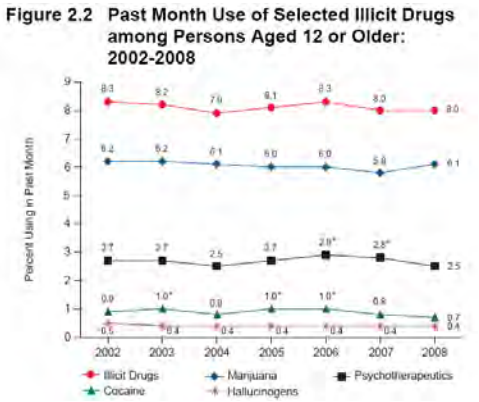
Stimulant Dependence

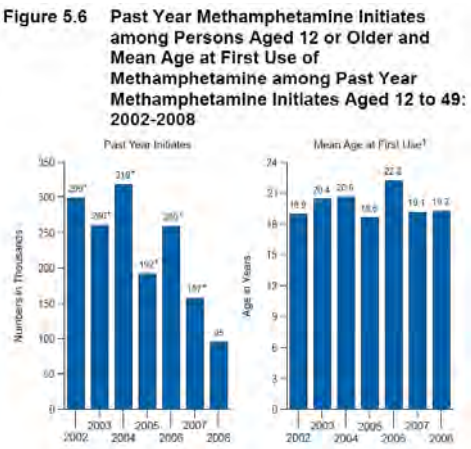
- Stimulants
  - Cocaine
  - Methamphetamine
  - Amphetamine

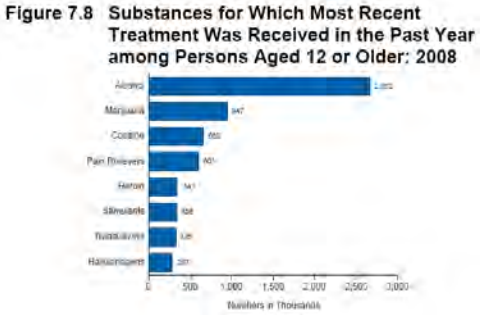
- 12 – 15% ever tried stimulants

- 1-3% have stimulant dependence

- 50% of sober stimulant dependent individuals relapse within a year.







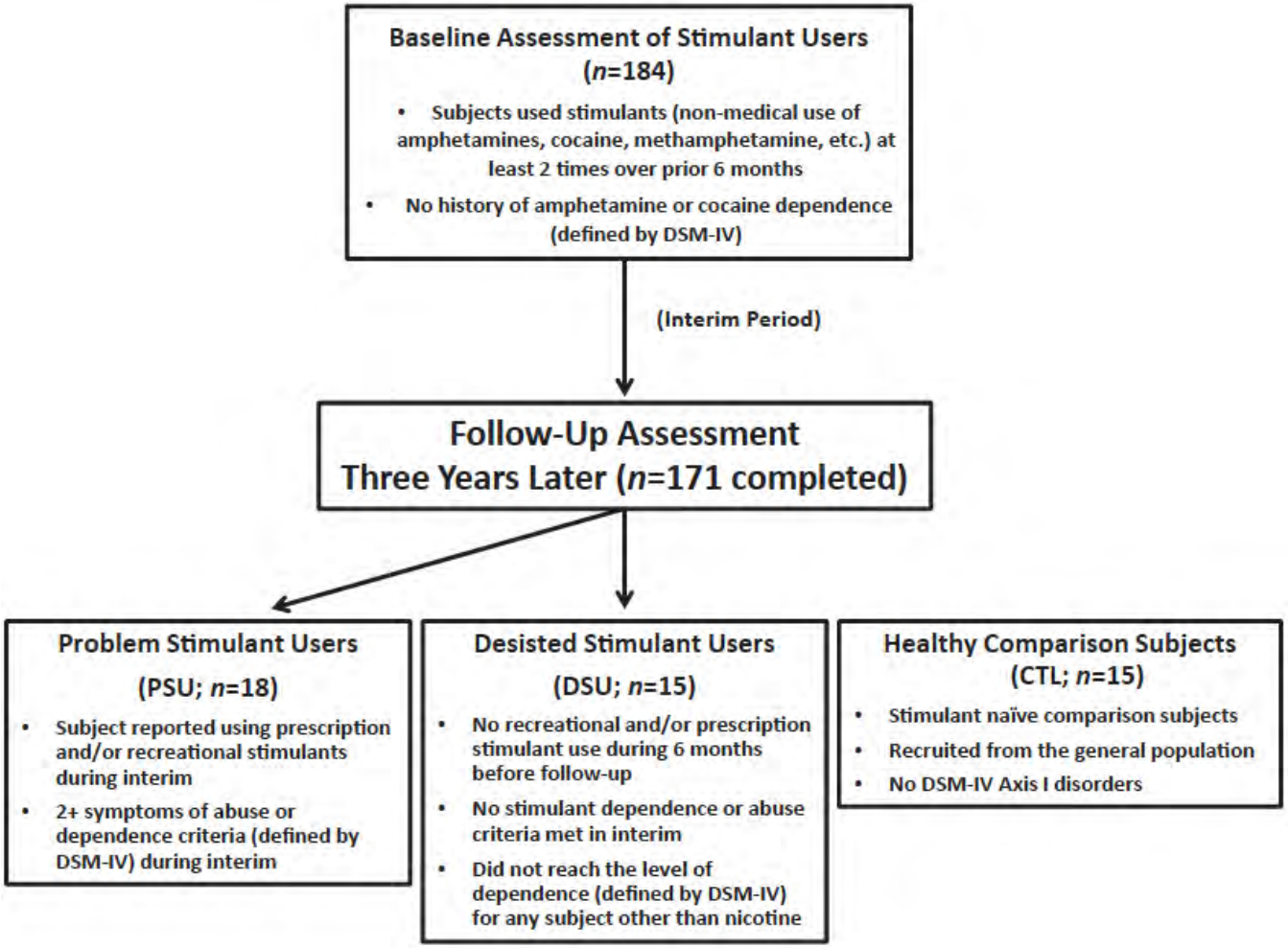
LIBR
Laureate Institute for Brain Research

Goal of the Study

- Only about 1 in 7 individuals who experiment with illicit substances goes on to develop a substance use disorder.
- Current approaches are not able to predict **who is at highest risk to develop problem use**.

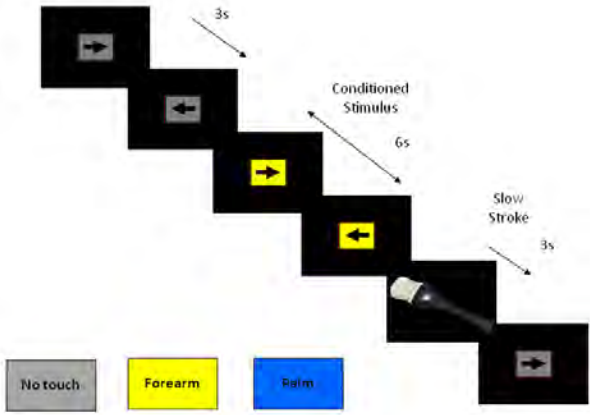
- *Can we use fMRI and computational psychiatry to:*
  - *Predict problem use?*
  - *Identify process dysfunction?*
  - *Identify modifiable risk factors?*

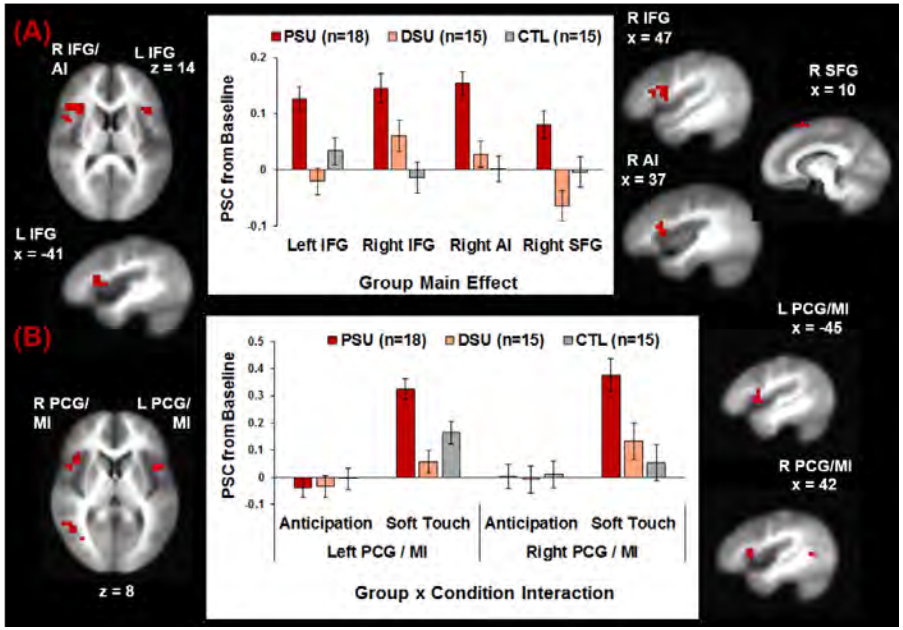
Longitudinal Follow up Stimulant Users:



Positive Interoceptive Processing

- Problem Users show exaggerated positive valence interoceptive processing:





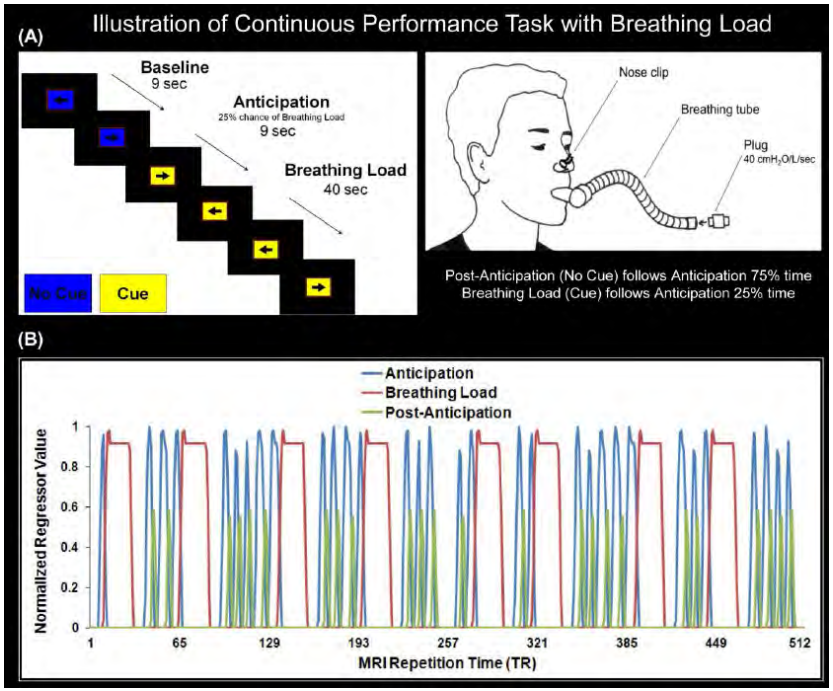
Hyperactivation to pleasant interoceptive stimuli characterizes the transition to stimulant addiction☆

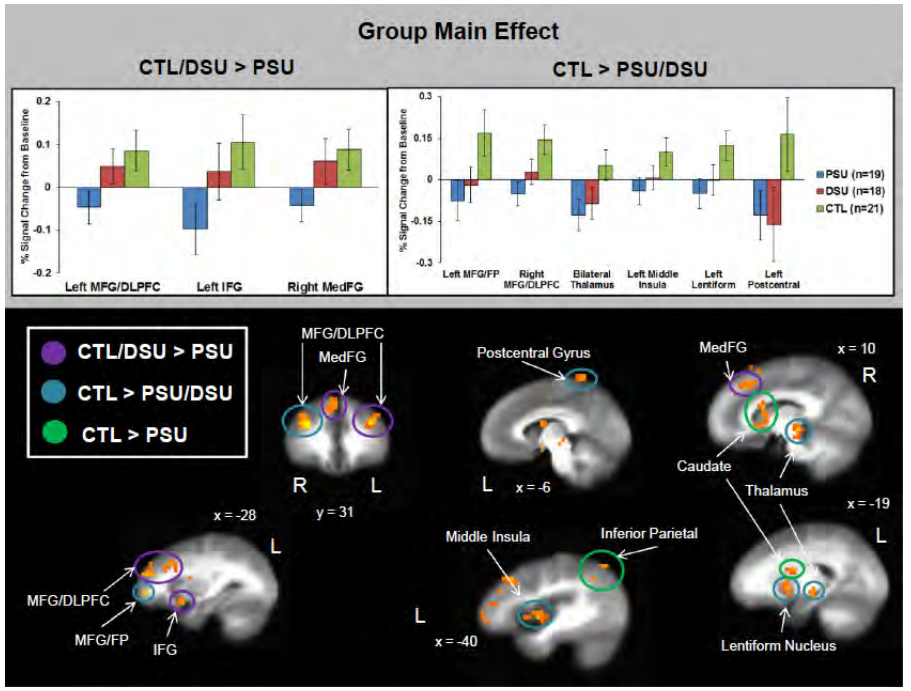
Jennifer L. Stewart[a,b,*], April C. May[b], Susan F. Tapert[b,c], Martin P. Paulus[b,c]

LIBR
Laureate Institute for Brain Research

Aversive Interoceptive Processing

- Problem Users show attenuated aversive interoceptive processing:



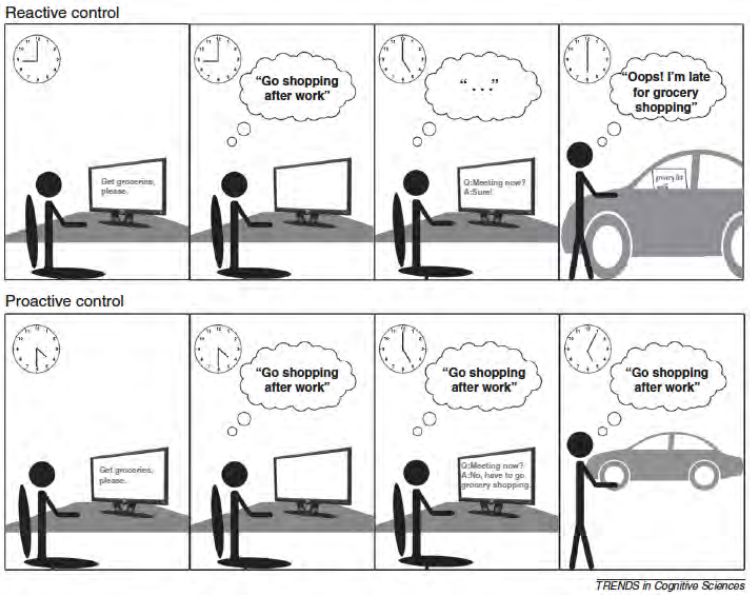


Do you feel alright? Attenuated neural processing of aversive interoceptive stimuli in current stimulant users

JENNIFER L. STEWART,[a] ASHLEY L. JUAVINETT,[b] APRIL C. MAY,[a] PAUL W. DAVENPORT,[c] AND MARTIN P. PAULUS[a,d]

Inhibition – Cognitive Control

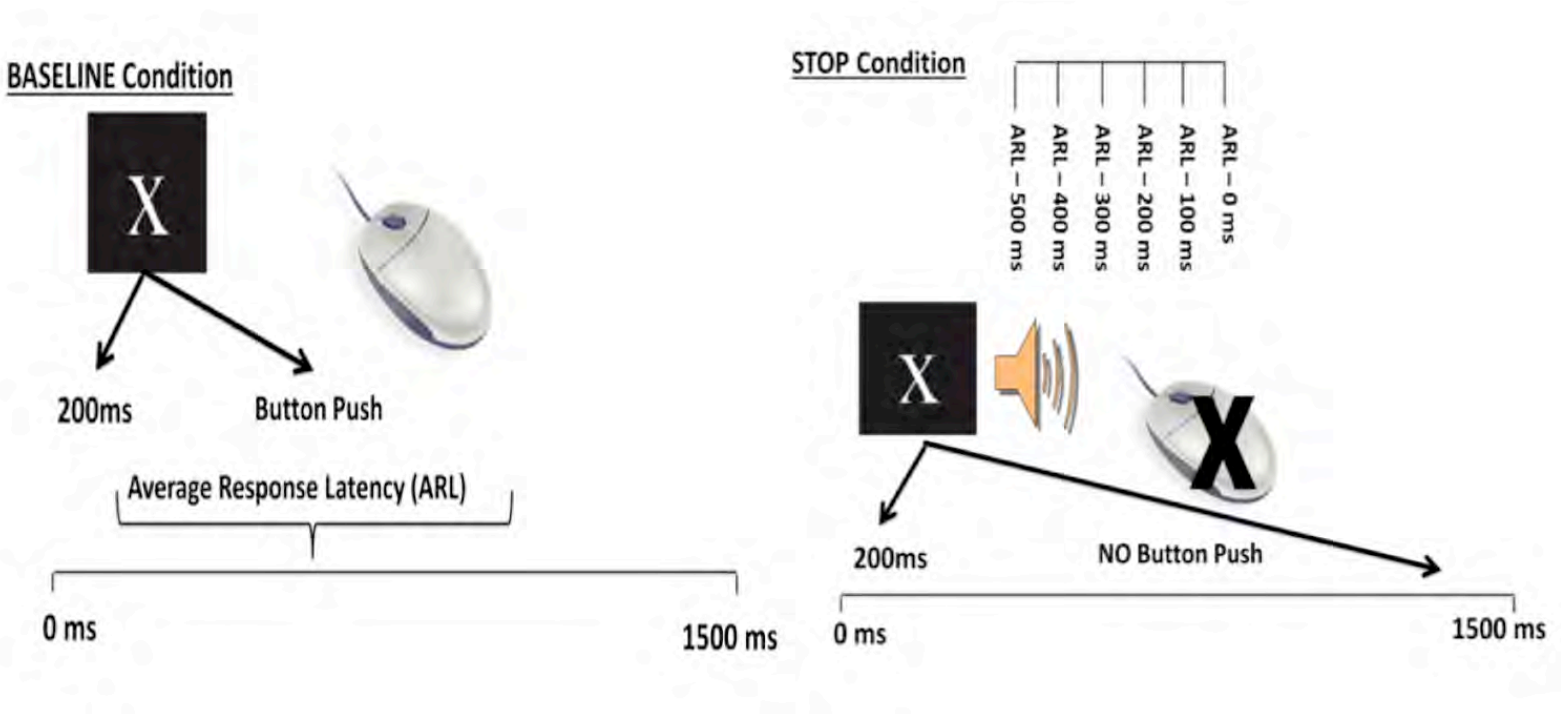
- Pro-active / Reactive Cognitive Control Framework:



The variable nature of cognitive control: a dual mechanisms framework

Todd S. Braver[1,2*]

Stop Signal Task



Stop Signal Paradigm – Inhibitory Control

- Stop-signal paradigm:
  - subjects perform a go task (e.g. reporting the identity of a stimulus), occasionally, the go stimulus is followed by a stop signal, which instructs subjects to withhold the response Logan, Cowan et al. 1984.
  - performance in the stop-signal paradigm is modeled as a race between a 'go process' and a 'stop process' Verbruggen and Logan 2008
  - Post-stop-signal slowing is not a reflexive byproduct of the stop signal Bissett and Logan 2012 but an active strategic process.

Methods

- Over a five-year period, potential participants were recruited via Internet ads, newspapers, and flyers (Reske et al., 2011)
  - 1,025 individuals underwent detailed phone screens
- 184 OSU met study inclusion criteria:
  - 2+ off-prescription uses of cocaine or amphetamines over the past six months;
  - Exclusions: lifetime stimulant dependence; lifetime stimulant use for medical reasons; treatment of substance-related problems.
- fMRI session was completed within two weeks of the baseline interview session:
  - Participants were instructed to abstain from illicit substance use $\geq$ 72 hours prior to this session
  - Abstinence was determined by urine toxicology screen
  - Stop-signal task while undergoing fMRI.

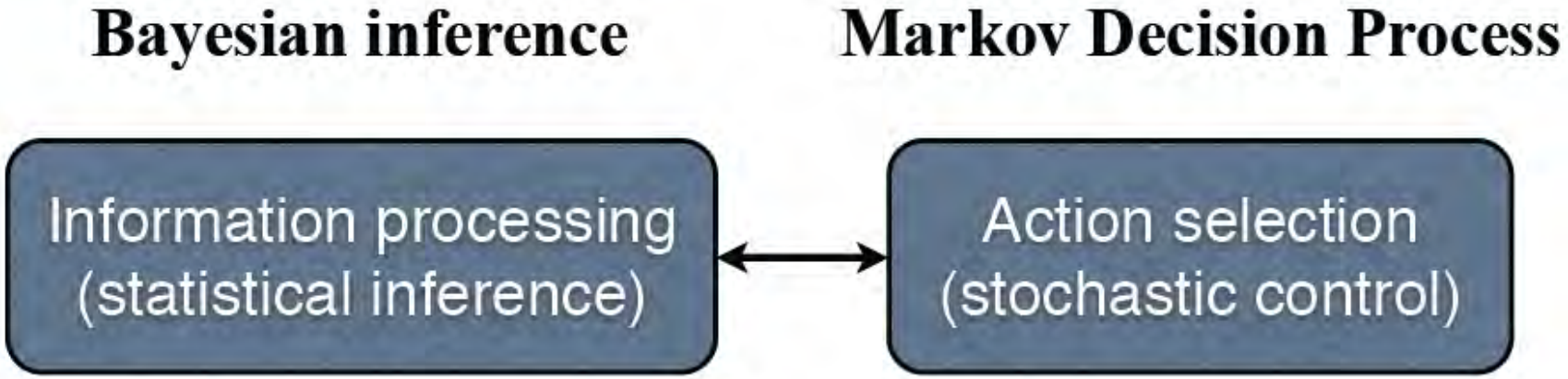
Analysis Approach

- Behavior during Stop Signal Task:
  - Bayesian Ideal Observer Model
  - Model-parameter-based regressors – neural activation
- Cross-validated robust regression and random forest analyses:
  - Identified potential predictive neural regions by conducting voxelwise logistic regression analyses in a randomly selected 'training' subset.
  - Test subset was used to assess the relative predictive power of these activation clusters identified based on the training sample.
  - For each predictive regions two sets of coefficients were obtained:
    - based on the test sample (used for final cut points and accuracy estimation),
    - based on the training sample.
  - Comparative analyses:
    - To obtain a more selective predictive model, we included all these activations as well as the three baseline lifetime drug use measures as independent variables in a random forest analysis predicting 3-year clinical outcome.

Cohort:

	Problem Stimulant Users (PSU) n=38		Desisted Stimulant Users (DSU) n=50		
	Mean	SD	Mean	SD	t-test
Demographics					
Age	20.7	1.6	21.0	1.3	p=.41(.83)
Education	14.6	1.4	14.9	1.2	p=.33(.98)
Verbal IQ (WTAR)	109.8	6.1	108.6	8.6	p=.47(.71)
Alcohol (typical drinks/week)	18.6	13.7	18.2	13.8	p=.91(.11)
Nicotine (typical cigarettes/day)	2.3	3.7	2.9	4.5	p=.87(.15)[a]
Attention/Hyperactivity (from SSAGA II)					
ADHD Attention Symptoms	1.4	2.5	0.5	0.99	p=.26(1.13)[a]
ADHD Hyperactivity Symptoms	1.2	2.2	0.7	1.4	p=.56(.59)[a]
Conduct Symptoms	1.5	1.7	1.6	1.5	p=.86(.17)
Personality/Mood					
BIS	66.8	9.6	64.5	9.0	p=.24(1.18)
SSS	25.0	4.9	24.7	4.7	p=.73(.35)
BDI	1.7	1.7	3.4	3.9	p=.37(.89)[a]
Lifetime Drug Uses (Baseline)					
Cocaine	26.2	40.6	22.3	46.2	p=.27(1.10)[a]
Prescription Stimulants	29.1	38.7	24.7	78.0	**p=.03**(2.22)[a]
Cannabis	784.9	1094.6	811.4	1158.1	p=.52(.64)[a]
Interim Drug Uses (Baseline - Follow-Up)					
Cocaine	279.0	605.9	5.8	18.7	**p<.001**(5.06)[a]
Prescription Stimulants	60.6	86.6	6.5	28.2	**p<.001**(7.20)[a]
Marijuana	580.9	924.7	762.9	1520.0	p=.54(0.62)[a]

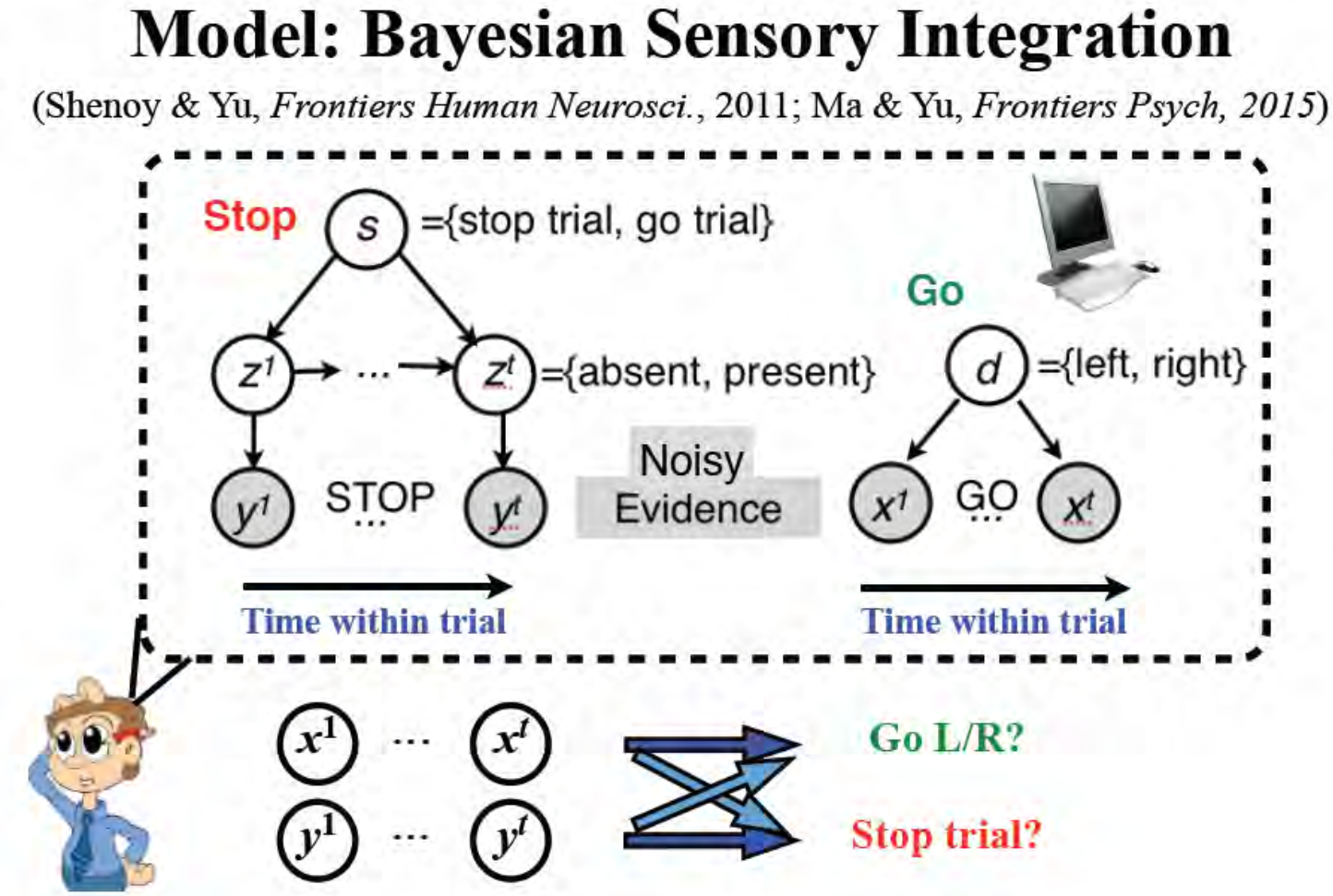
Dynamic Bayesian Modeling



- inferring state of the world
- prediction what happens next
- learning statistical contingencies

- decision policy: mapping between perceived world states and actions
- action selection should be context-dependent/goal directed

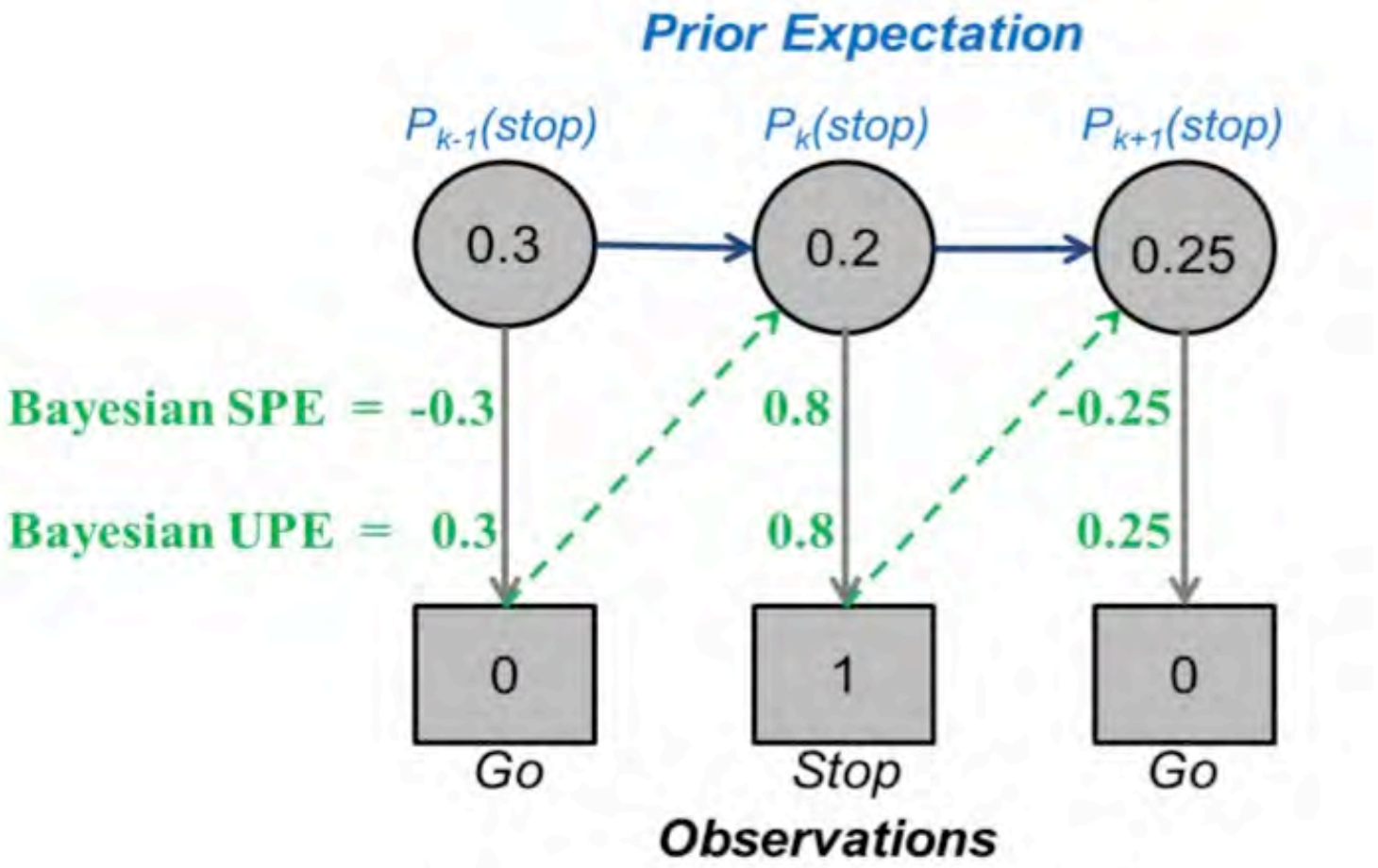
Dynamic Belief Updating (Yu et al.)



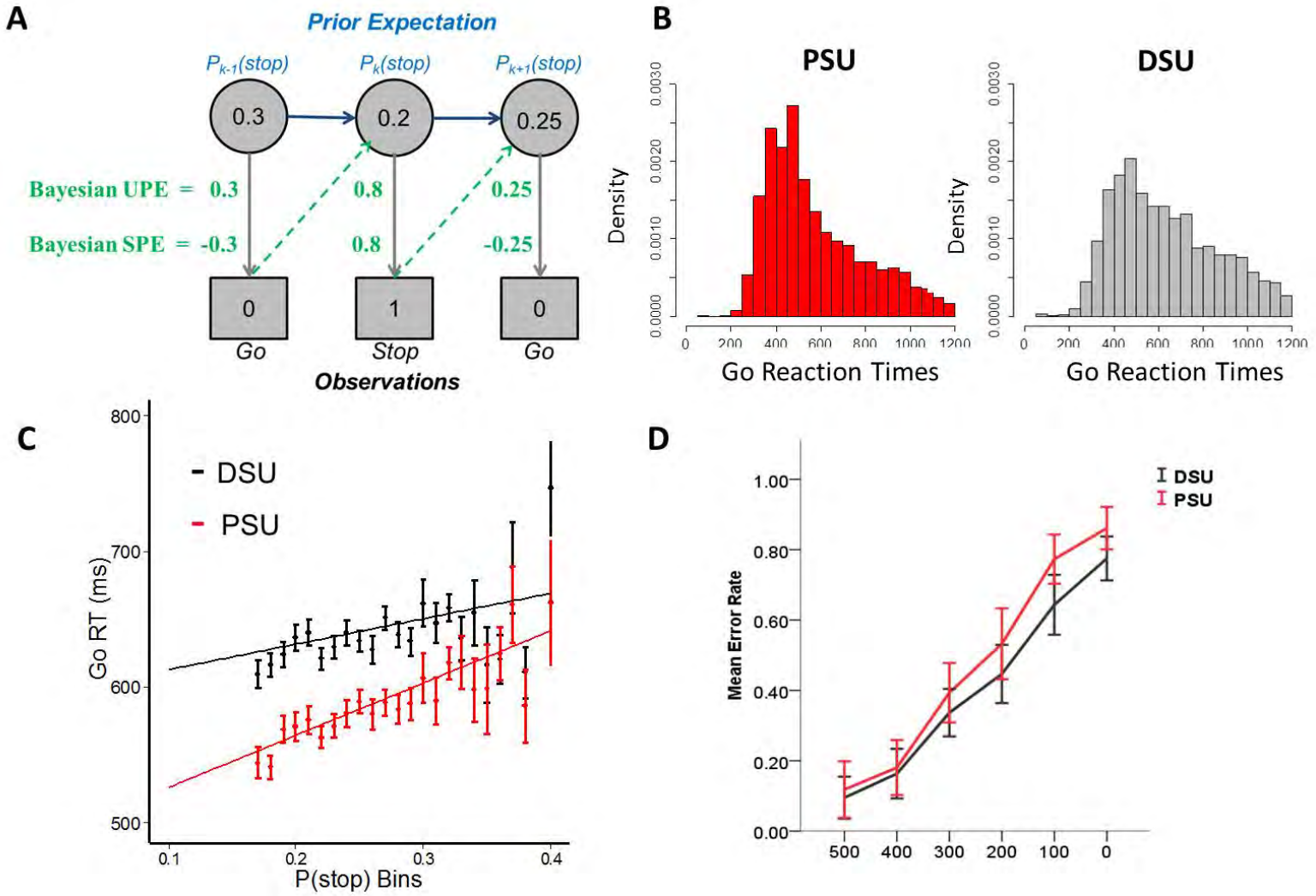
P(stop) – Dynamic Bayesian Model

- Assumptions about subjects' internal beliefs regarding task structure:
  - On each trial k, there is a hidden probability r_k of observing a stop signal (s_k =1 for stop trial) and probability 1-r_k of observing a go trial (s_k = 0).
  - r_k is the same as r_{k-1} with probability α, and re-sampled from a prior beta distribution $p_0(r)$ with probability 1-α.

Bayesian Prediction Error:

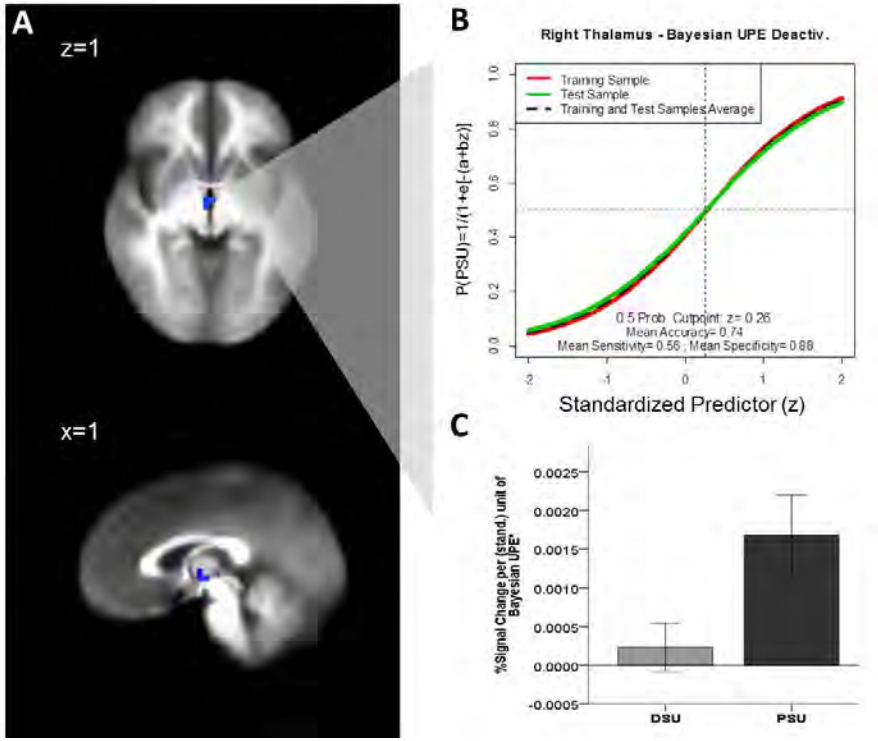


Dynamic Bayesian Belief Model



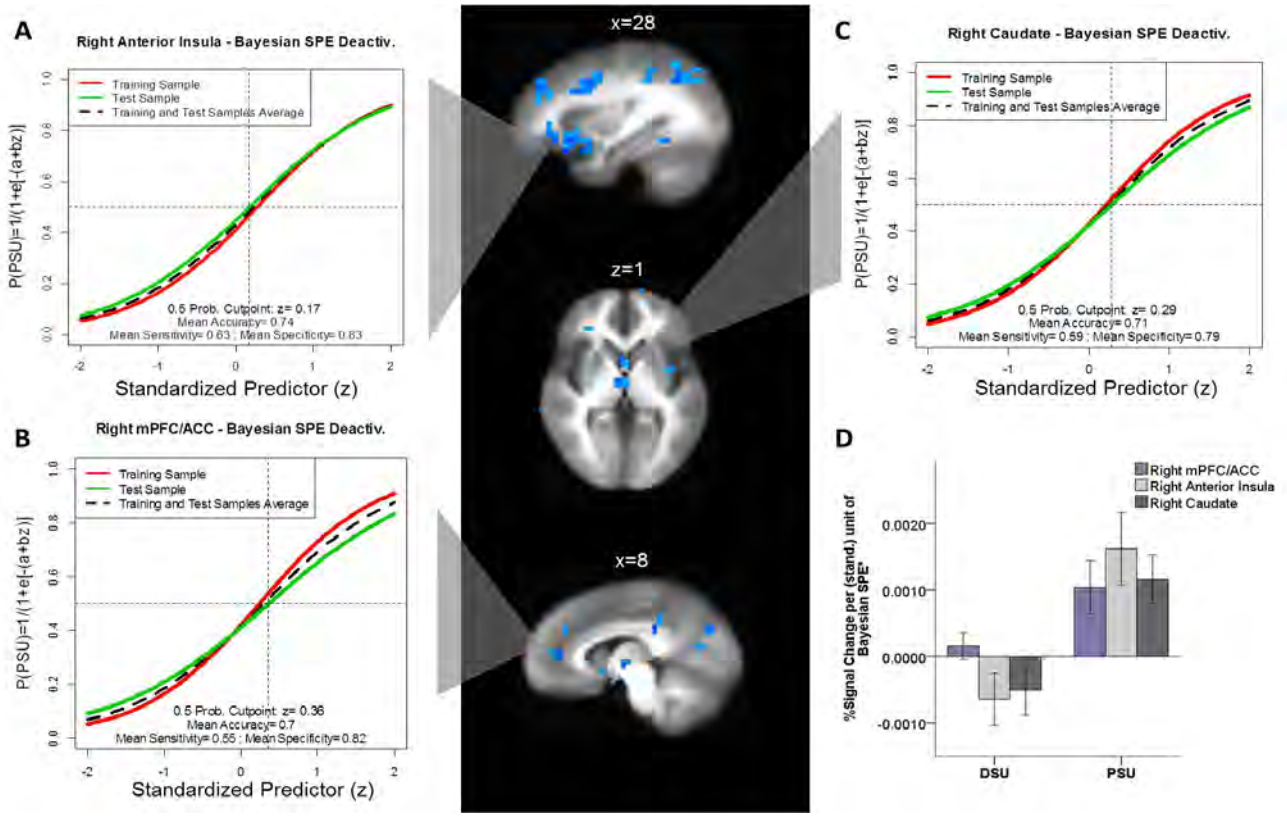
Unsigned Prediction Error (UPE):

- Thalamus:
  - For 1 standardized unit increase in UPE deactivation one was about three times as likely to develop a future stimulant use disorder:



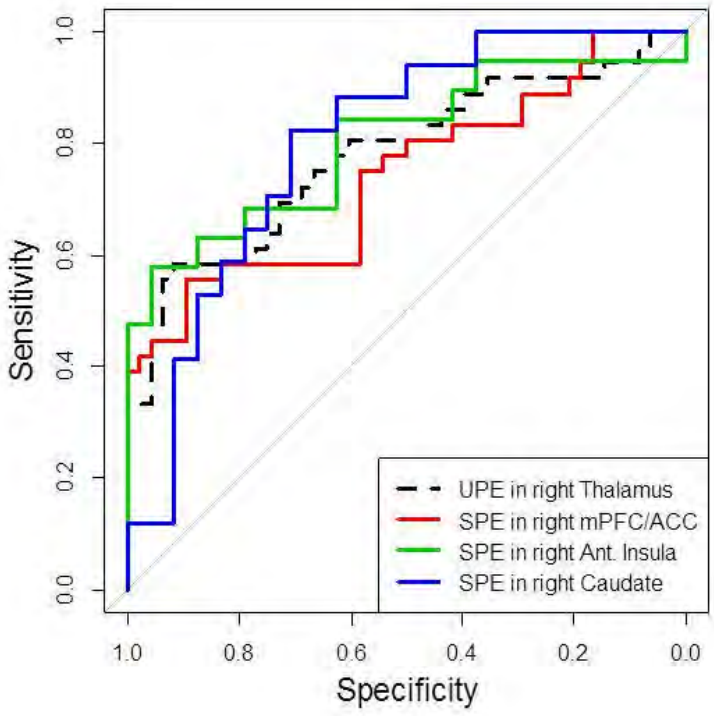
Signed Prediction Error (SPE):

- For 1 unit increase in SPE deactivation one was 2-3 times as likely to be in the PSU group in
  - mPFC/
  - anterior insula/IFG
  - and caudate



Prediction / Classification Characteristics:

- Individual computational model brain activation generate good single subject sensitivity:



Predictors	*B*	*SE*	*p* value	*exp(B)*	95% CI for *exp(B)*	Cut Point*
Right Thalamus (UPE)	1.24	.35	.001	3.45	1.74-6.86	.26
Right Medial PFC/ACC (SPE)	0.89	.41	.029	2.44	1.09-5.44	.36
Right Anterior Insula (SPE)	1.16	.45	.009	3.19	1.32-7.71	.17
Right Caudate (SPE)	1.10	.41	.007	3.02	1.35-6.74	.29

Machine Learning / Classifier (Pereira, Mitchell et al. 2009):

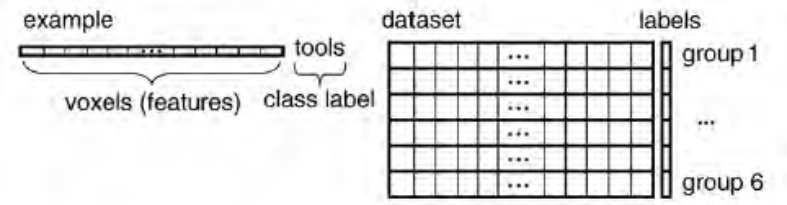


Fig. 1. An example where features are voxels arrayed as a row vector (left) and a dataset is matrix of such row vectors (right).

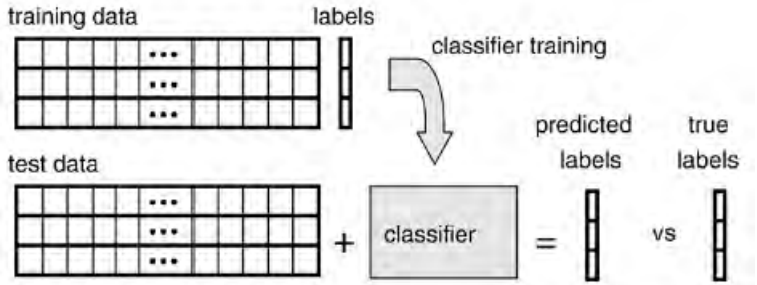


Fig. 2. A classifier is learned from the training set, examples whose labels it can see, and used to predict labels for a test set, examples whose labels it cannot see. The predicted labels are then compared to the true labels and the accuracy of the classifier–the fraction of examples where the prediction was correct–can be computed.

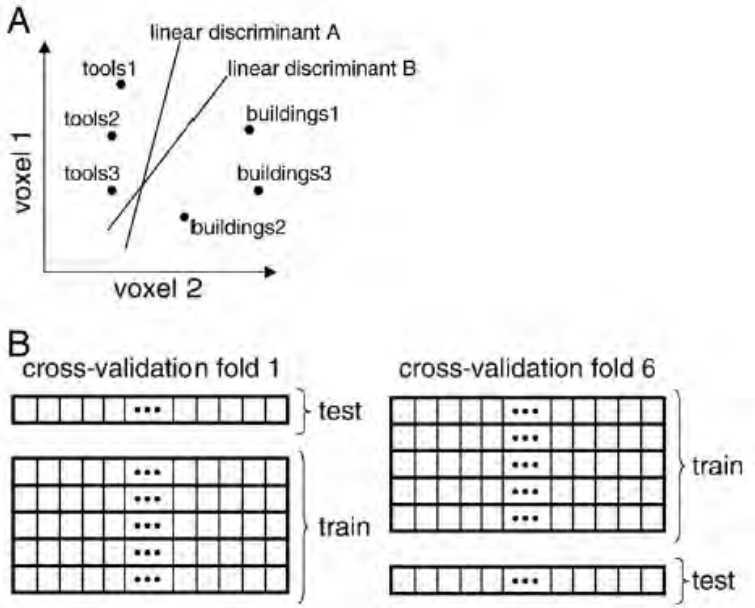
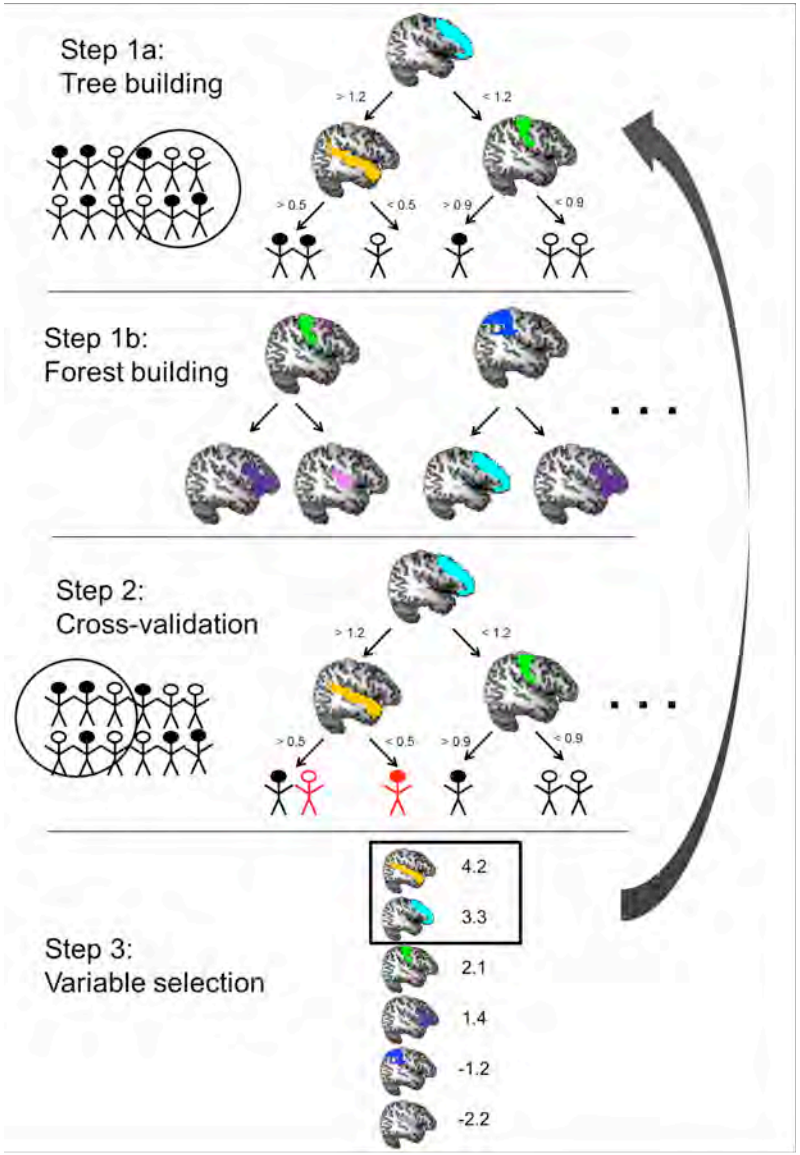


Fig. 3. (A) learning a linear classifier is equivalent to learning a line that separates examples in the two classes (vectors in a 2-voxel brain) as well as possible. (B) during cross-validation each of 6 groups of examples takes a turn as the test set while the rest serve as the training set.

LIBR
Laureate Institute for Brain Research

Random Forest Procedure

- Construct a large number of classification trees (n= 2000)
  - bootstrapped subsample of participants
  - randomly selected subset of the independent variables
- Data from participants that were not used to create Random Forest:
  - Evaluate how in aggregate RF classify "problem users" versus "desisters."
  - Evaluate variable importance: how much a particular variable contributes to classification accuracy.
- Based on variable important as in (Nicodemus, Callicott et al. 2010), a final, reduced model with 10 variables was generated.



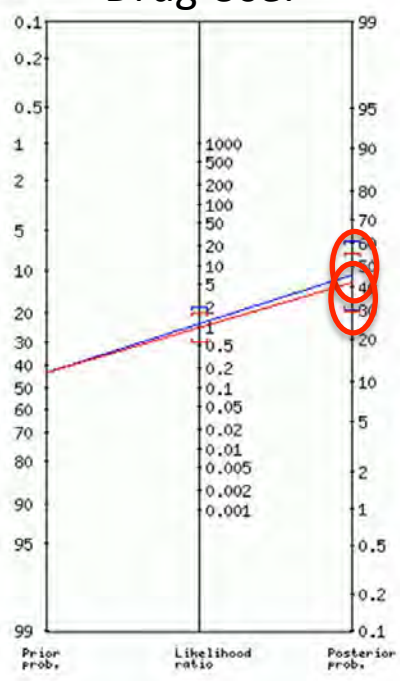
LIBR
Laureate Institute for Brain Research

Prediction Comparison

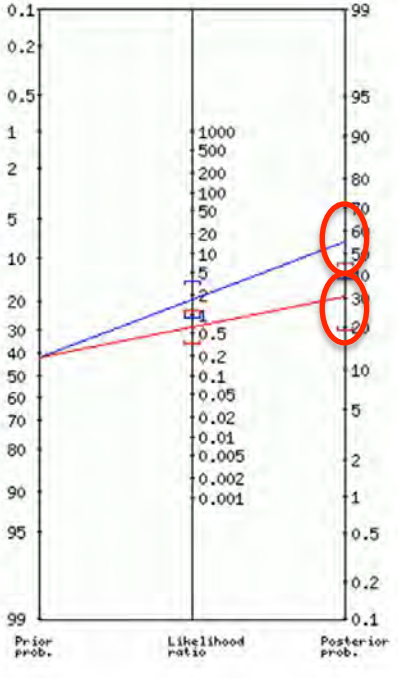
Model	Accuracy	Sensitivity	Specificity	Positive LR (95% CI)	Negative LR (95% CI)
Drug Use[a]	52%	48%	56%	1.08 (0.56, 2.08)	0.94 (0.54, 1.63)
Neural – SST Categorical	64%	59%	67%	1.76 (0.88, 3.52)	0.62 (0.33, 1.17)
Neural - SST Computational	74%*	62%	83%	3.51 (1.34, 9.21)	0.47 (0.26, 0.87)

A

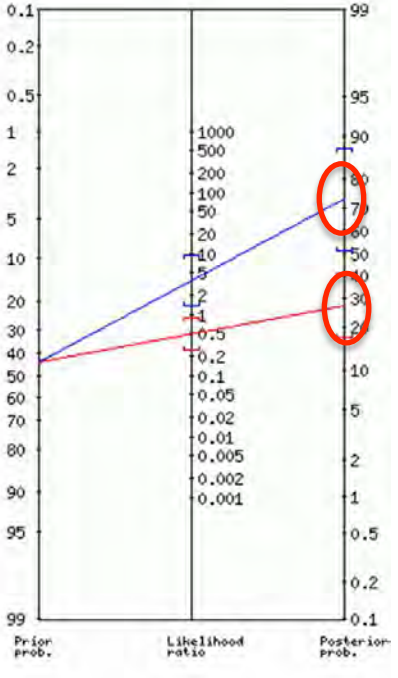
Drug Use:



fMRI only:



fMRI + Model:



General Conclusions:

- Problem users show a number of processing deficits:
  - Exaggerated positive interoceptive processing
  - Attenuated aversive interoceptive processing
  - Increase “surprise” to predicting inhibitory events

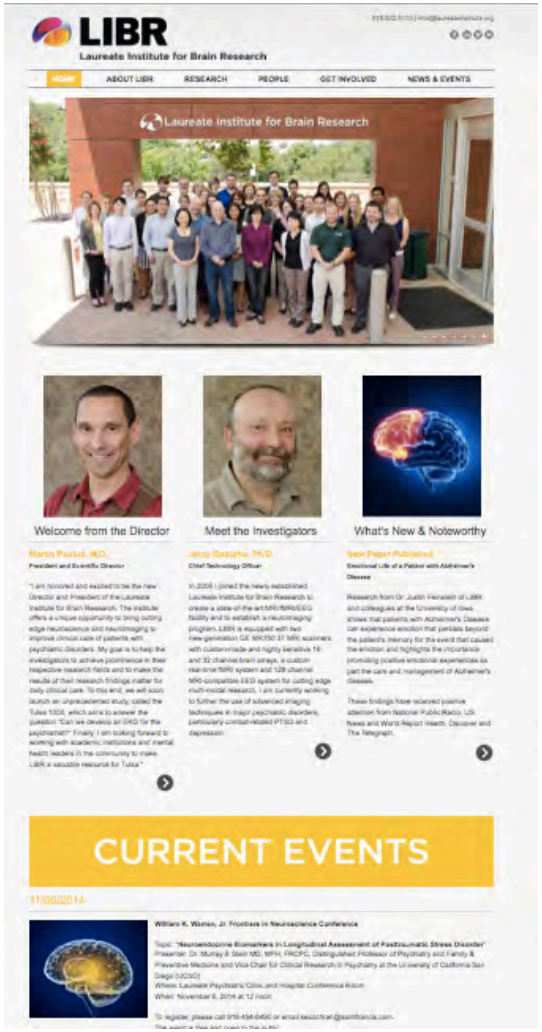
- **Problem users but not Desisters, when faced with inhibitory control, fail to build a good internal model of the environment.**

- Machine learning approaches help to improve robustness of prediction.
- Computational approaches modeling behavior may further enhance predictive accuracy.

Future Directions

- What are **important prediction targets** for different populations?
  - At risk individuals (substance use, mental illness)
  - Problem behaviors – harm to self and others
  - Insight and compliance
- What are **modifiable risk factors**?
  - E.g. does changing brain activation help to reduce risk?
- **Treatment implications**:
  - Can we use individual-level predictions to select specific interventions?
- What is the **economic utility** of individual level prediction?

Thank you!


LIBR
Laureate Institute for Brain Research
HOME
ABOUT LIBR
RESEARCH
PEOPLE
GET INVOLVED
NEWS & EVENTS
Laureate Institute for Brain Research
Welcome from the Director
Meet the Investigators
What's New & Noteworthy
CURRENT EVENTS