Conductome

Simposio: Inteligencia Artificial y Psicología

Christopher R. Stephens, Dagmara Wrzecionkowska, Romel Calero, Sofia Michaelian, Ixel Alonso

Proyecto Chilam: C3 – Centro de Ciencias de la Complejidad, Universidad Nacional Autónoma de México

AMEPSO 18-20/09/2024

Conductome

1. **Inteligencia Artificial ¿Por qué debería interesar a los psicólogos?** Chris Stephens, C3
2. **Plataforma Proyecto 42: Datos y la metodología atrás de los datos.** Dagmara Wrzecionkowska, C3
3. **Selección de las variables en modelos psicológicos: la búsqueda de explicabilidad y predictibilidad.** Sofía Ixchel Michaelian Martínez, C3 y Posgrado en Ciencia e Ingeniería de la UNAM
4. **Caso de uso 1: Obesidad: ¿qué tan importantes son los factores psicológicos?** Romel Calero, C3
5. **Caso de uso 2: Retraso de gratificación y conductas obesogénicas.** Adriana Ixel Alonso Orozco, C3 y Universidad Iberoamericana

Conductome

Inteligencia Artificial ¿por qué debería interesar a los psicólogos?

Christopher R. Stephens

C3 – Centro de Ciencias de la Complejidad y Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México

Simposio AMEPSO 2024

Conductome

Psychology

1. the study of the mind and behavior.
2.the supposed collection of behaviors, traits, attitudes, and so forth that characterize an individual or a group.

https://dictionary.apa.org/psychology

Goals: 1) to describe, explain, predict, and change or control behaviors
2) To describe, explain, predict and change or control groups

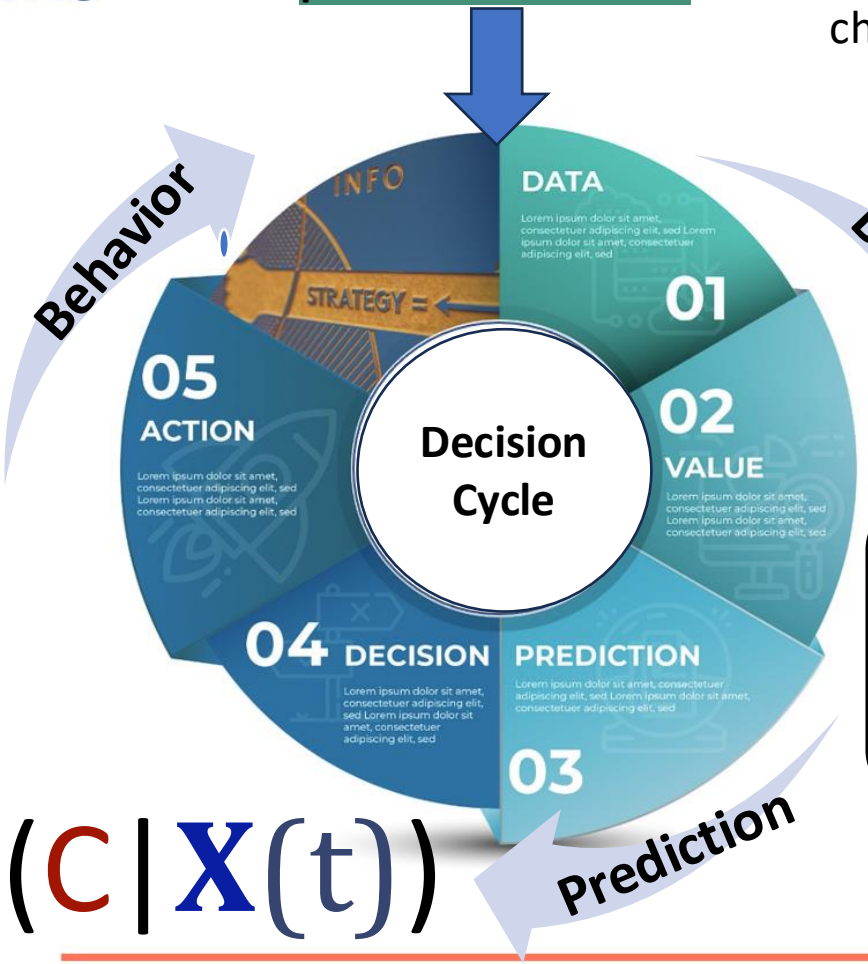
How can Artificial Intelligence/Machine Learning help achieve these goals?

Methodological versus philosophical differences with standard psychological approaches.
e.g., what's more important – explain or predict? BOTH!

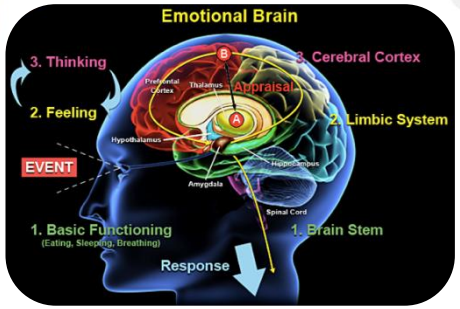
Conductome

Behavior is an emergent phenomena that is a result of multiple decisions and actions

Desired outcomes with performance metrics



Consequences of behavior: obesity, infections, climate change, poverty/inequality,…



Psychology

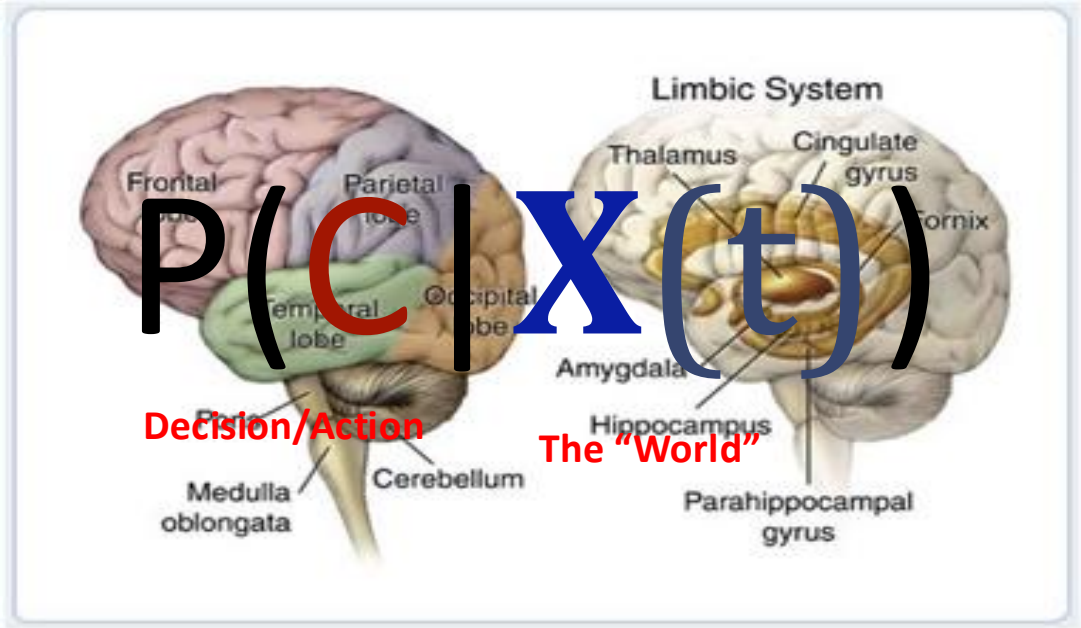
“Decision/action” C“+
“World” X +
“algorithm” P(|) +
“payoff/reward”

$$P(C|X(t))$$

How do we predict and explain behavior?
With psychological theories and psychological constructs

Psychological Theories

- Behaviorism
- Social Cognitive Theory
- Theory of Planned Behavior
- Integrated Model of Behavior
- Health Beliefs Model
- …



- Delayed gratification
- Locus of control
- Anxiety
- Depression
- Stress
- Auto-efficacy
- Power-distance
- Uncertainty avoidance
- Impulsivity
- …

Psychological Dimensions

How does my theory or my constructs help me to: to describe, explain, predict, and change or control behaviors or groups?

Conductome

- **Psychology has typically used a “top down” theory-based approach. Top-down approaches make many asumptions.**

- **Machine learning more naturally uses a “bottom up” data-based approach. Bottom-up approaches can be difficult to explain.**

Let’s make a comparison. Both should predict!

Top-down assumption type 1: the choice of predictors ("independent variables") and their relation to the predicted ("dependent variable")

Conductome

The role of these drivers is just to affect the "real" causes which are psychological constructs

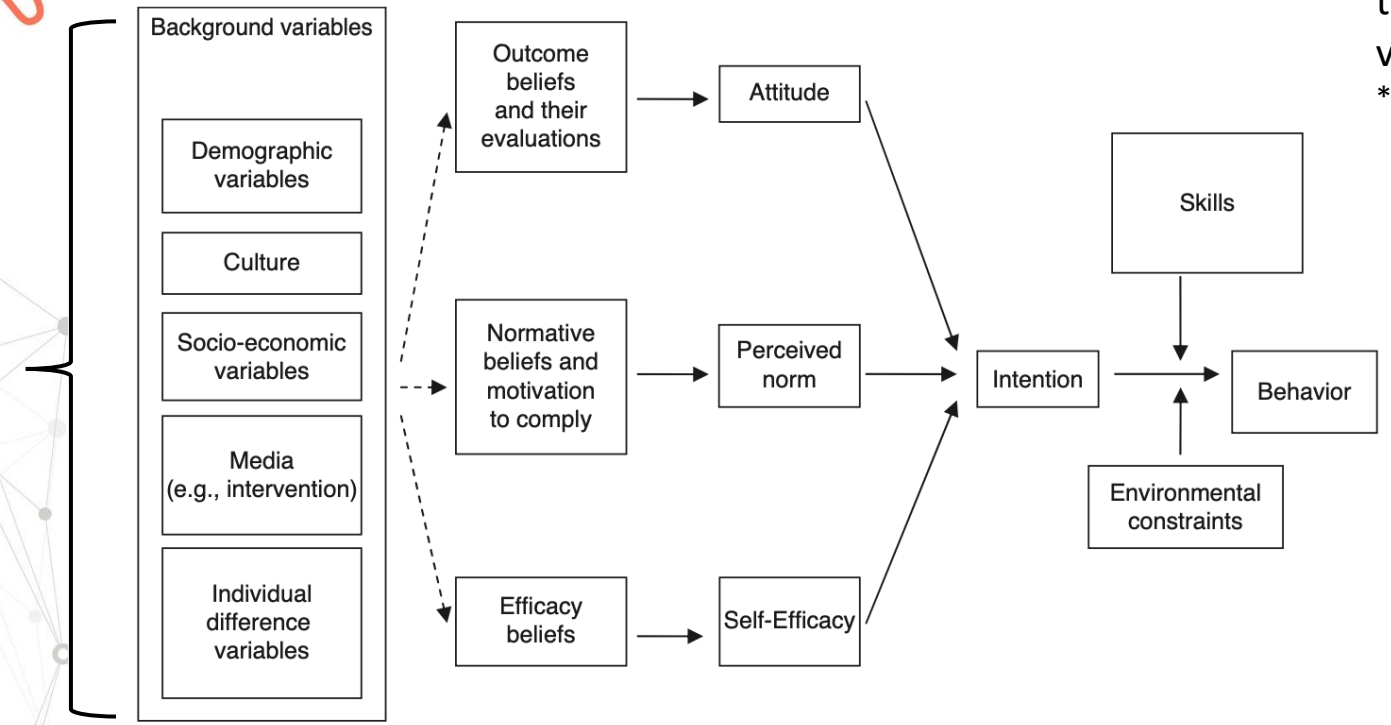


Figure 2.1 The Integrative Model of Behavioral Prediction

"Theory" is imposing a causal structure on the variables – many of which are latent variables in order to "**explain**" a behavior.

* In essence this is imposing a Bayesian prior.

This is a universal theory in that, in principle, it applies for ANY behavior. It is not an explicit predictive model until we fill the "boxes" with data and assume relations between them.

How well does the model **predict** the behavior?

Do we model the behavior C as a discrete variable – classification - or a continuous (or discretized continuous) variable - regression?

What will be our performance measure?

Top-down assumption type 2: Behavior can be predicted and explained with only a few predictors

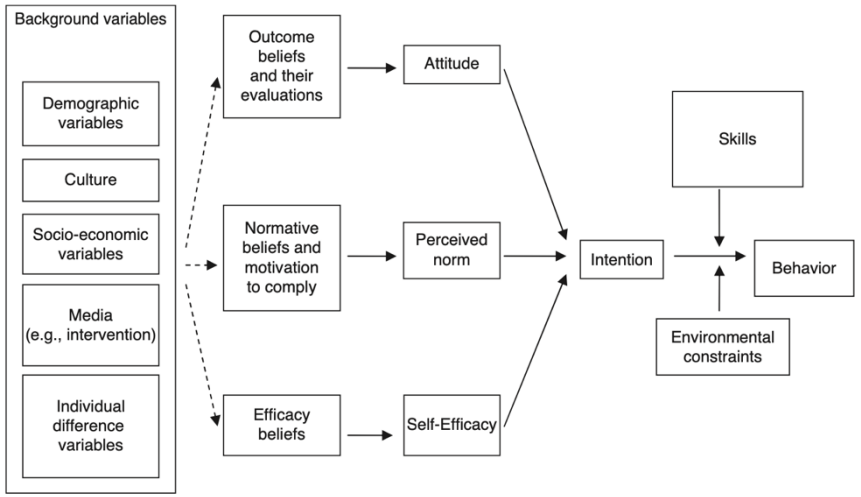


Figure 2.1 The Integrative Model of Behavioral Prediction

In this model, potentially, only 9 variables are needed, depending on how the different psychological constructs are built. For instance, Self-Efficacy could be represented by one scale associated with a summed total across a set of items on a Likert scale, thereby yielding just one predictor.

- How do we determine the quality of this model?
- We need to compare it to other models using objective metrics.
  - For any given chosen behavior - predictability is objective.
  - Explainability? Do we want an explanation that is good only for that behaviour, even if the explanation is different for another one, or do we want something that explains “universally”?
- To what other models can we/should we compare it?

We can first of all try and remove the assumptions (model restrictions) that are part of the model.

Top-down assumption type 3: Explaining is more important than predicting

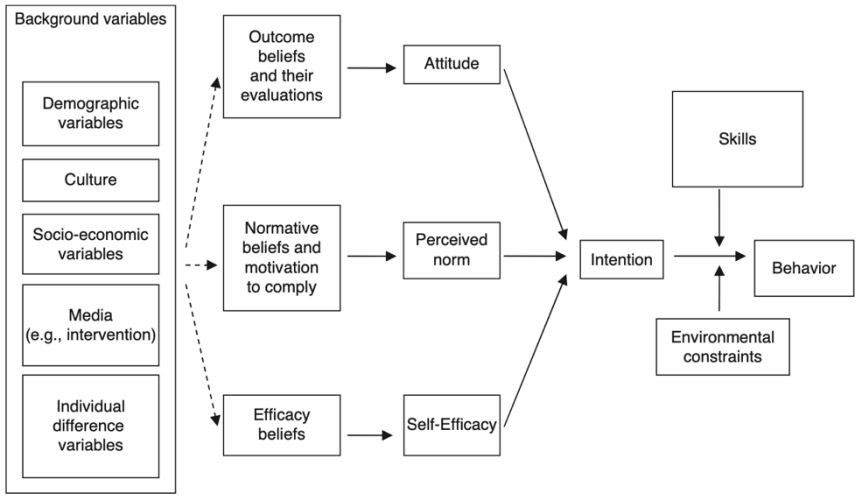


Figure 2.1 The Integrative Model of Behavioral Prediction

To test the predictive performance of the model we need to evaluate how it generalizes to **new data**. It is not good enough to find a good statistical fit just on the data on which the model was built!

This can be done in two ways:

i) Randomly divide your data into a training set for model build and a test set for model validation. There are multiple ways of varying sophistication to do this.

ii) Get new data – the problem there is that the new data should be statistically similar to the data you built the model on.

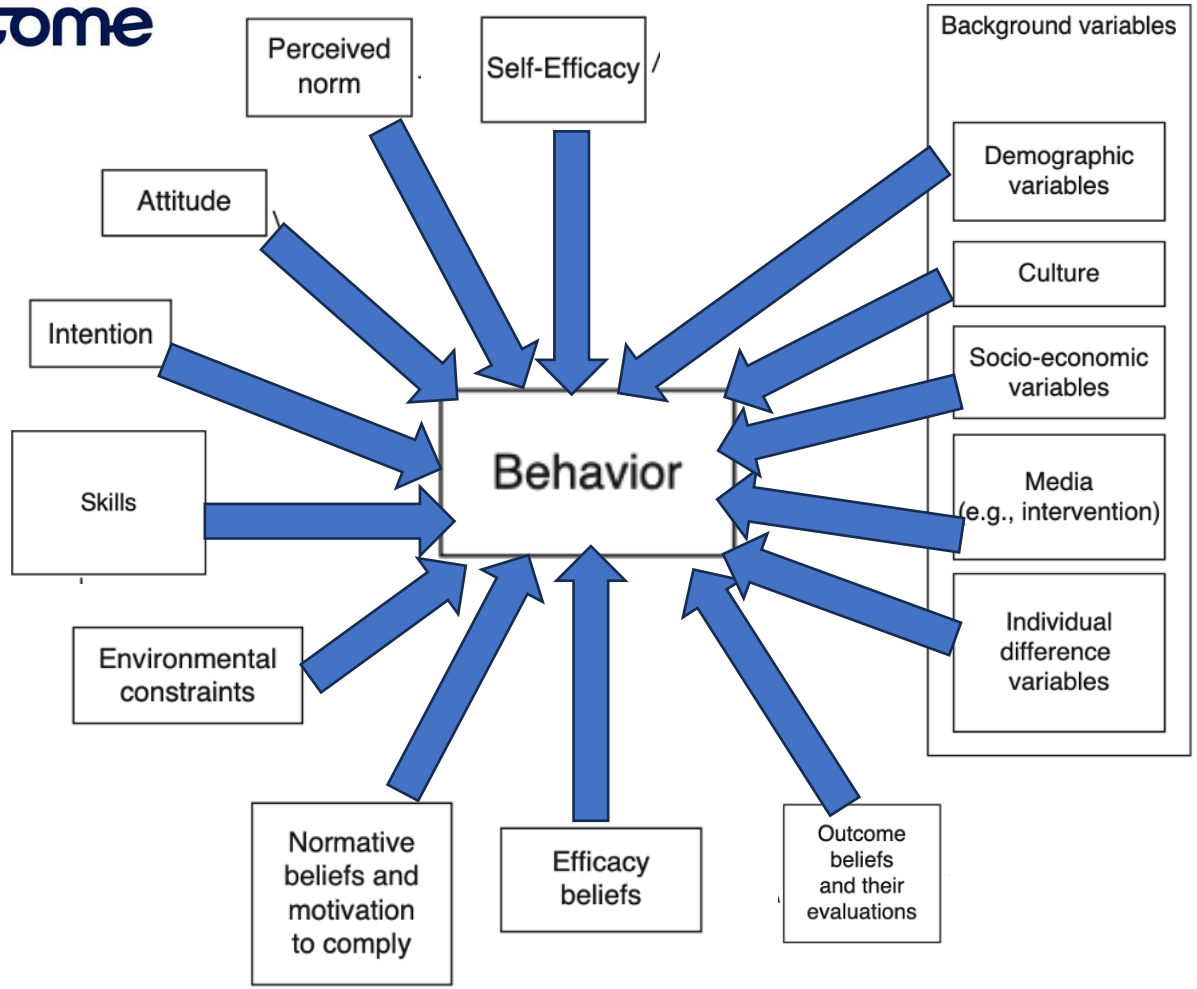
There are two broad classes of model – regression and classification – according to if the prediction a number or a class.

Regression – R^2, adjusted R^2, mean absolute error, mean squared error etc.
Classification: Confusion matrix, AUC ROC etc.

Eliminating top-down assumption type 1: Assume no causal structure between the “boxes”

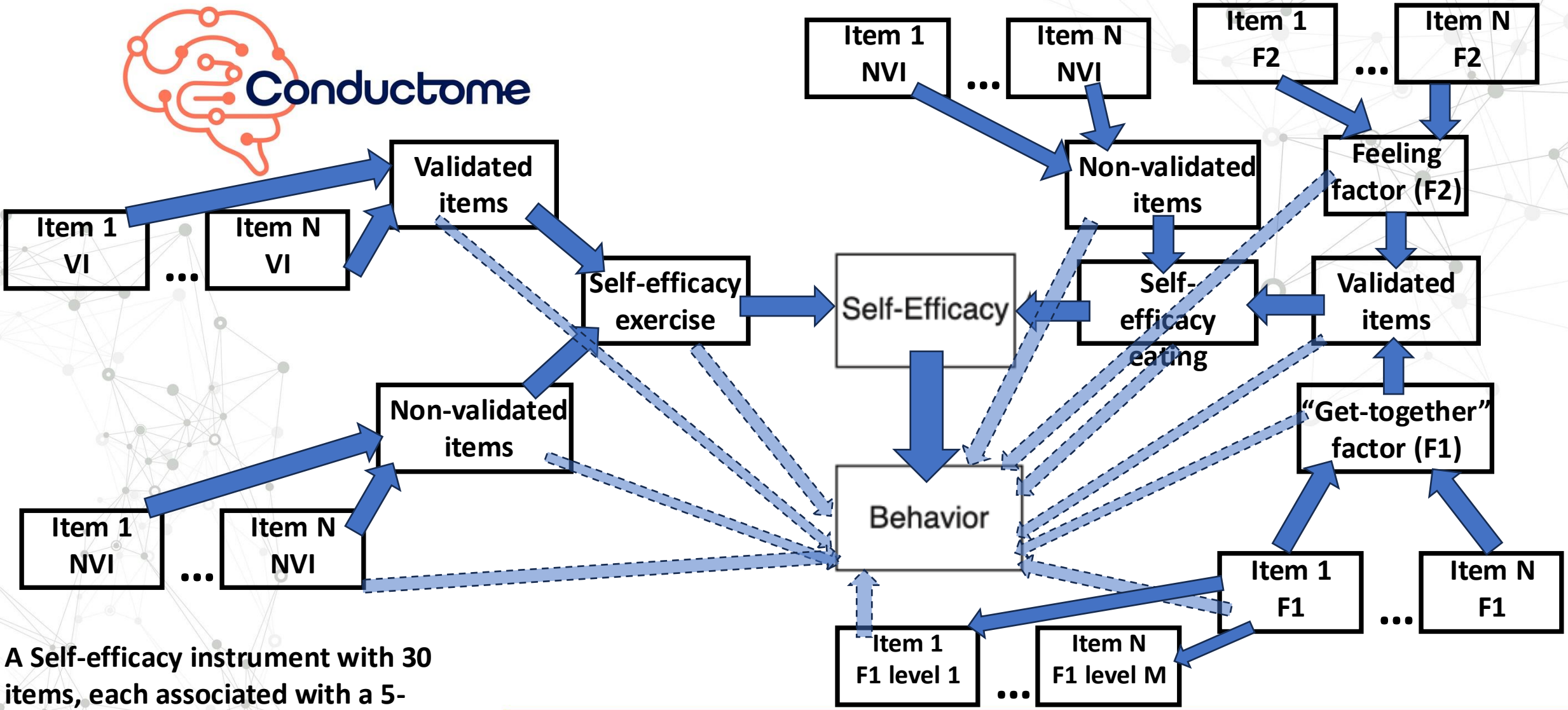
Conductome

There is less theoretical bias about potential causal structure at this level. In principle, every factor type/class is given equal a priori opportunity to predict or explain behavior.



Use a predictability performance measure to determine how well this model predicts

Eliminating top-down assumption type 2: Consider “all” possible predictors



A Self-efficacy instrument with 30 items, each associated with a 5-level Likert scale really has 150 underlying variables. Not counting the possible combinations!

Top-down assumption type 4: Properties of the constructs

Conductome

1. What is the logic to choose a subset? There are an exponential number of combinations.
   - Exploratory and/or confirmatory factor analysis is often used. This is based on linear relations between the items and is INDEPENDENT of the behavior to be predicted.
2. For any subset of N items with M values – scale, subscale etc. – we need to determine a function F(N,M) → S, a one-dimensional score that summarizes (reduces the dimensionality) of the scale.
   - E.g., the sum, or the average, of the values. But why not the square? Or square root? Why not use more than one?
3. We then need to put the results of 1. and 2. in a prediction model that has model bias, e.g., a linear regression is based on a linear relationship between the results of 1. and 2. and the dependent variable.
4. If we model at the level of the individual responses then no assumptions are made.

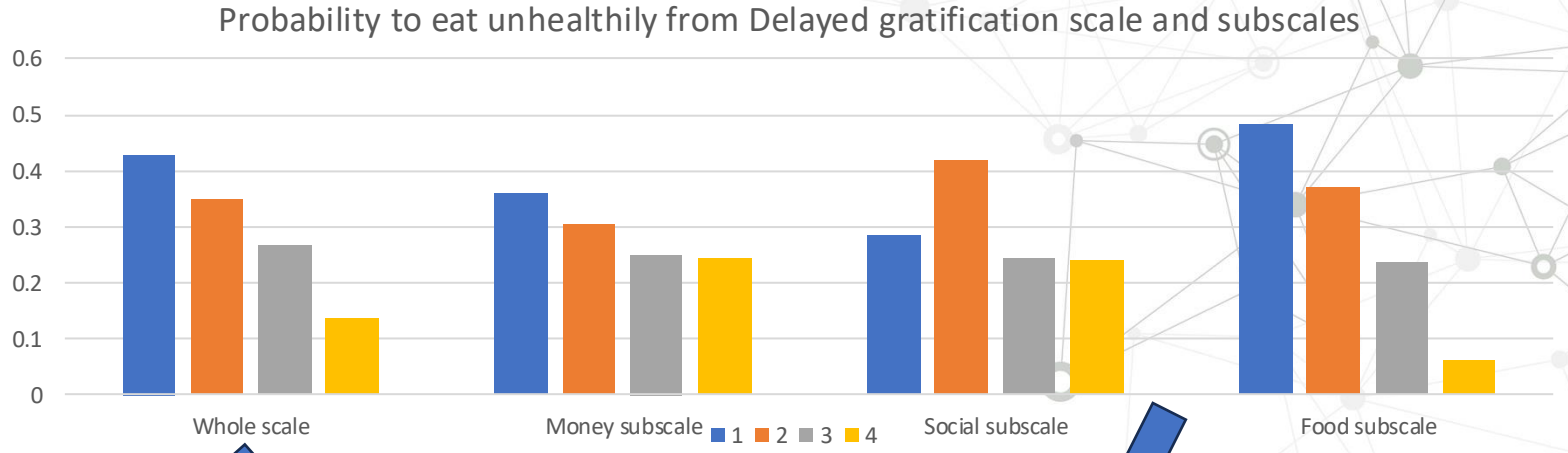
How do we determine which variables predict and/or explain better?
We can put them in a Bayesian machine learning model

Every ordinal variable is turned into a finite set of binomial variables. For example, the result of a sum of values from a construct of 30 items on a 5-point Likert scale (1, 2, 3, 4, 5) can be divided into a set of "score boxes" from a minimum (5) to a maximum (150) - e.g., 5-28, 29-52, 53-76, 77-100, 101-124, 125-150. We can then ask: which is more predictive – value 2 on the Likert scale of item 6 or whole score in the interval 101-124?

Eliminating top-down assumption type 4: Properties of the constructs… an example

Conductome

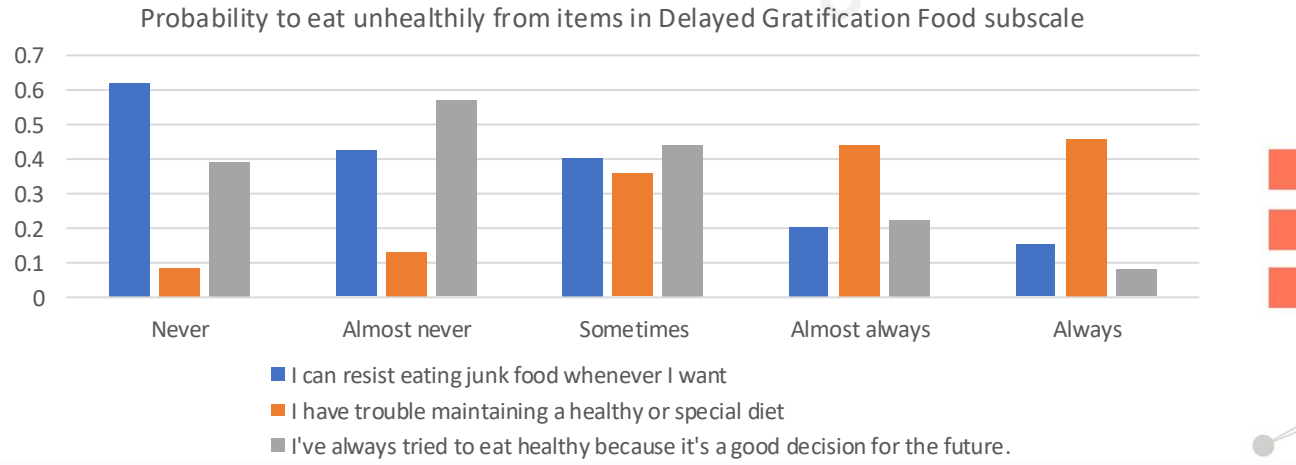
The behavior to be predicted is:
C = Eating unhealthily (self-reported).
The predictors come from a 45-item (5-point Likert responses) Delayed Gratification Scale = 45*5 = 225
Responses, 3 subscales (food, social and money) as well as the whole scale



Behaviour: Eating unhealthily

Predictability from Delayed Gratification

Eating behaviour



Eliminating top-down assumption type 3:
Explaining is more important than predicting

Question: What is more explanatory?

Behind the idea of “explanatory” is the idea of causality

- You eat unhealthily because you have low delayed gratification score
- You eat unhealthily because you have low delayed gratification score on the food/money/social subscales
- You eat unhealthily because: 1. you can never resist eating junk food whenever you want; 2. you always have trouble maintaining a special or healthy diet; 3. you’ve almost never tried to eat healthily thinking it was a good decision for the future

Question: What is more actionable?

Behind the idea of “actionability” is the idea of what intervention could possibly change the behavior?

Conductome

If the goal of psychology is to describe, explain, predict, and change or control behaviors, then Artificial Intelligence/Machine Learning has the potential to transform many areas of psychology:

i) By challenging, and offering alternatives, to many of the assumptions that underly psychological theories and constructs that can be barriers to predicting, explaining or changing behavior
ii) By emphasing a theory-agnostic, data-driven approach
iii) By offering tool sets that are more effective in modelling large, deep data sets than traditional statistical modelling tools
iv) By emphasizing the objective of prediction and offering a natural framework in which to measure it – though the type of machine learning algorithm used is crucial for preserving explainability
v) By offering a framework for comparing psychological factors to non-psychological factors in predicting, explaining and changing behavior

Next: Dagmara will show…, Romel will show…, Sofia will show…, and Ixel will show…

conductome.unam.mx
chilam.c3.unam.mx