# A comparison between propensity score matching, and weighting, in multiple treatment groups: A simulation study

Priyalatha Govindasamy, Ph.D
Department of Psychology & Counseling
(Sultan Idris Education University, Malaysia)

Antonio Olmos, Ph.D
Research Methods and Statistics
(University of Denver)

# Plan for today

What is the **problem?**

- Issues in the application of propensity score techniques with multiple groups

**How to address** the problem?

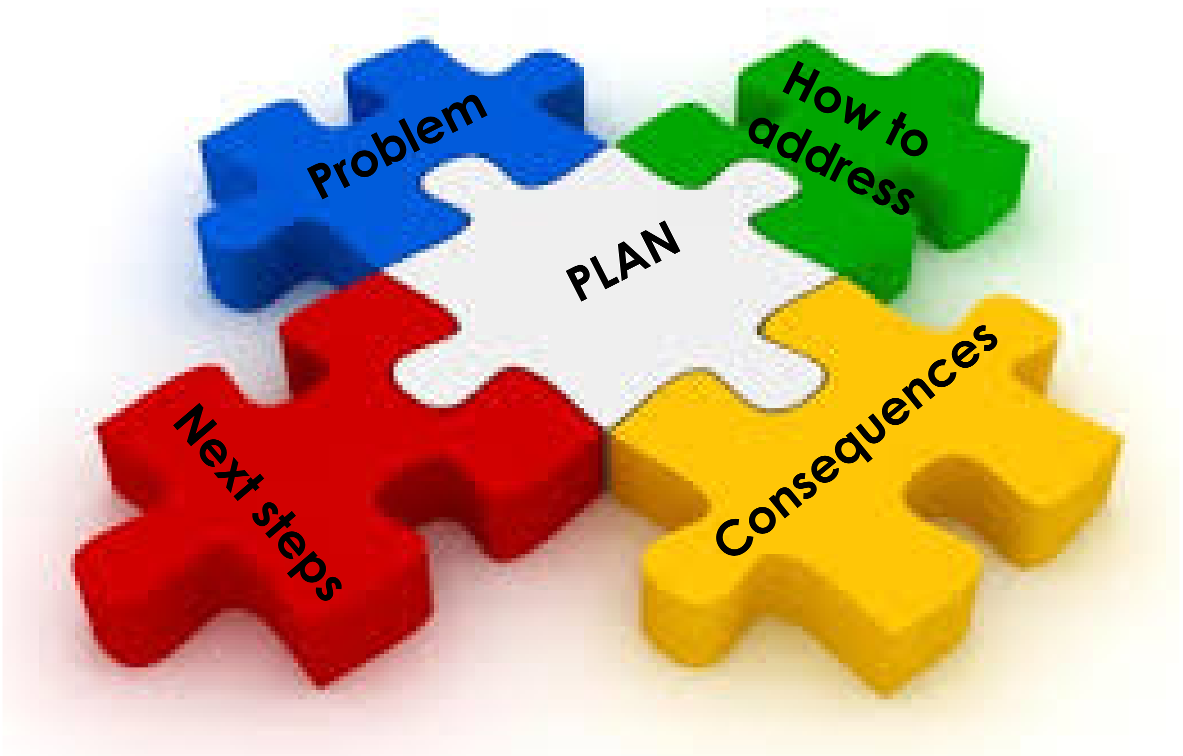
- Pose some questions addressing some application issues that can be common
- Suggest solutions to the questions
- Implementations of the solutions

What are the **consequences** of the solutions?

- Findings & Interpretations

**Next steps**?

- Directions for future research



# Before we start:
# Propensity score analysis approaches

**2. Stratification**
- assign treated and non-treated observations into strata with the same propensity score range

**1. Matching**
- find matches between treated and non-treated cases

**3. Weighting**
-weight to account for non-constant variability on the observed covariate between treated and non-treated groups

# What is the problem?

Given the differences in the implementation of propensity scores, how do we select an appropriate propensity score technique?

**In multiple treatment-groups settings**

# How to address the problem?

Compare the performance of propensity score analysis using **matching**, and **weighting** techniques in multiple treatment group settings under various data conditions:

a. Variable conditions:

1. Distributional characteristics of the treatment & outcome related variables (5 conditions)
2. Correlation between the treatment and outcome variable (3 conditions)

b. Sample size (3 conditions)

# Data conditions

**Condition 2:** Correlation between the variables (3 levels)

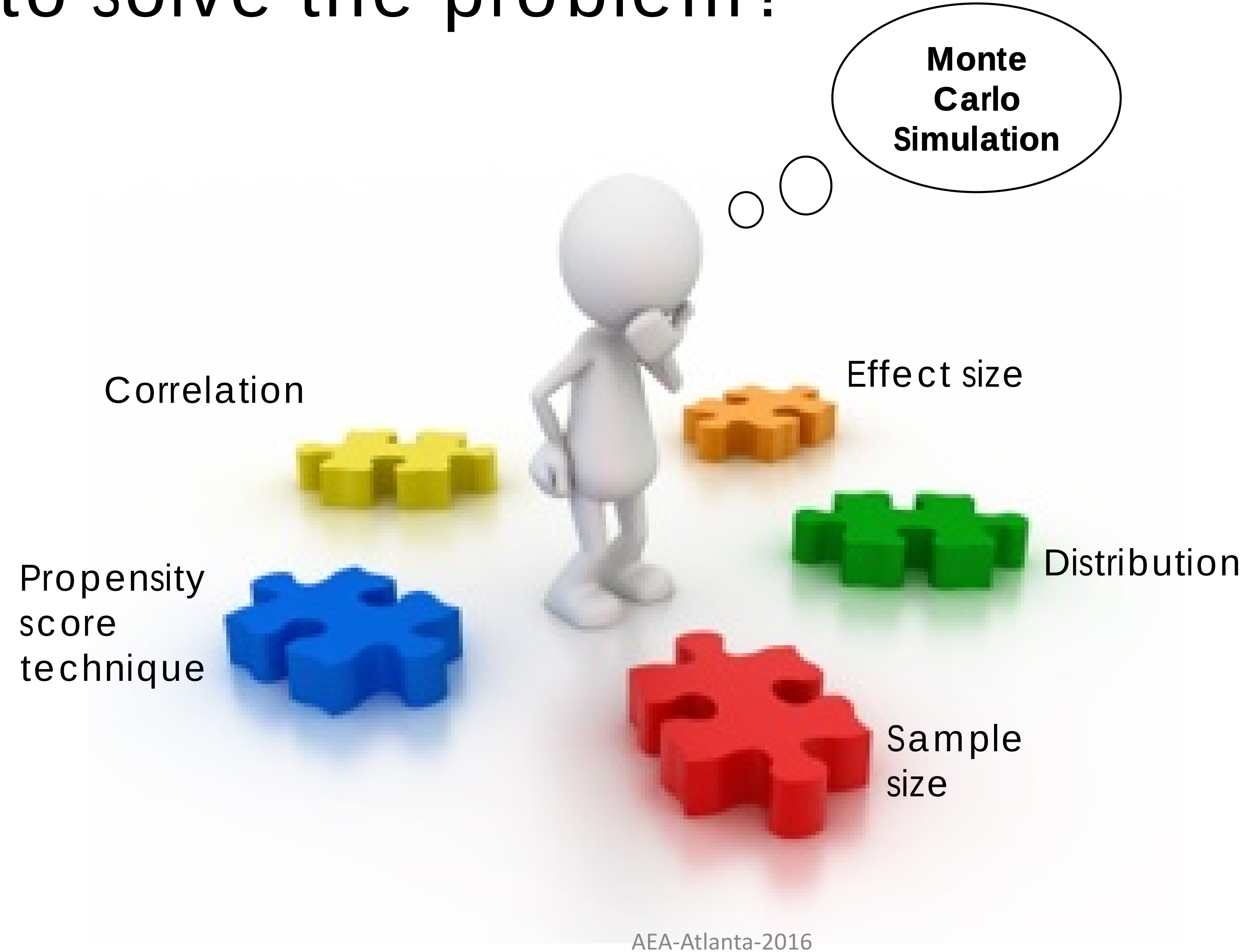
**Condition 3:** Distributional characteristics of the variables (5 levels)

**Condition 1:** sample size (3 levels)

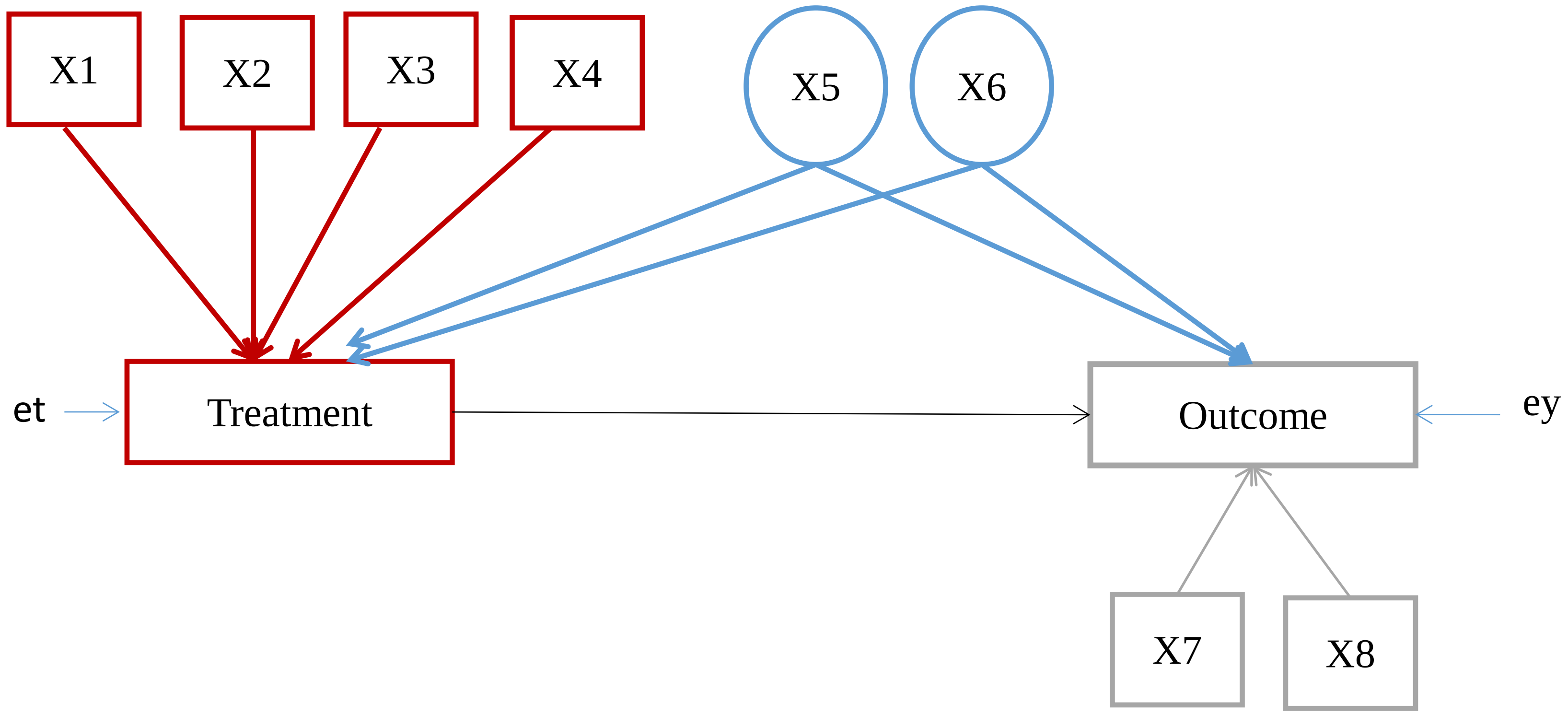
| Sample | Corr | Normal | Skewed | Mixed1 | Mixed2 | Mixed3 |
|---|---|---|---|---|---|---|
| n= 100 | 0.1 | Cond1 | Cond2 | Cond3 | Cond4 | Cond5 |
| | 0.3 | Cond6 | Cond7 | Cond8 | Cond9 | Cond10 |
| | 0.5 | Cond11 | Cond12 | Cond13 | Cond14 | Cond15 |
| n= 200 | 0.1 | Cond16 | Cond17 | Cond18 | Cond19 | Cond20 |
| | 0.3 | Cond21 | Cond22 | Cond23 | Cond24 | Cond25 |
| | 0.5 | Cond26 | Cond27 | Cond28 | Cond29 | Cond30 |
| n= 500 | 0.1 | Cond31 | Cond32 | Cond33 | Cond34 | Cond35 |
| | 0.3 | Cond36 | Cond37 | Cond38 | Cond39 | Cond40 |
| | 0.5 | Cond41 | Cond42 | Cond43 | Cond44 | Cond45 |

Notes: skewed = positively skewed, Mixed 1= Treatment variables are normal & outcome assigned variables are positively skewed, Mixed 2= Treatment variables are positively skewed & outcome variables are normal, Mixed 3 = one treatment and one outcome variable is positively skewed and the rest are all normal

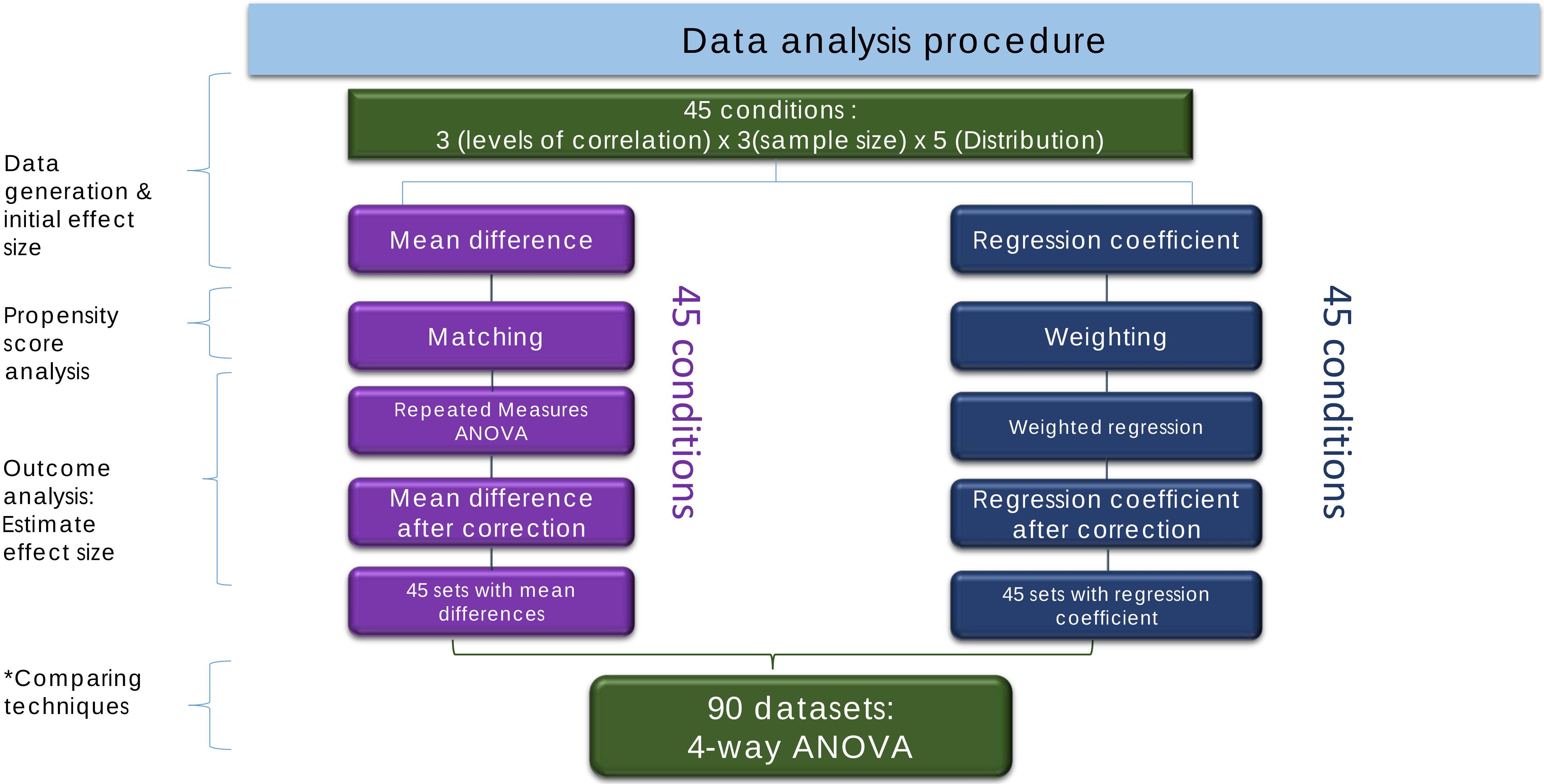
# How to solve the problem?



# Model depiction



# Procedure




Note: * Evidence for research questions

# How it was analyzed?


- Two **2x3x3x5** Analysis of Variance
  - **X1 vs Control**, **X2 vs Control**

- Used to evaluate the effects of
  1. **Propensity score technique** (2 levels: technique)
  2. **Sample size** (3 levels: Sample size)
  3. **Level of correlation** (3 levels : Level (degree of correlation)
  4. Distributional characteristics of the variables (5 levels: normallity/skewness)

  on the **effect size** (treatment effect)

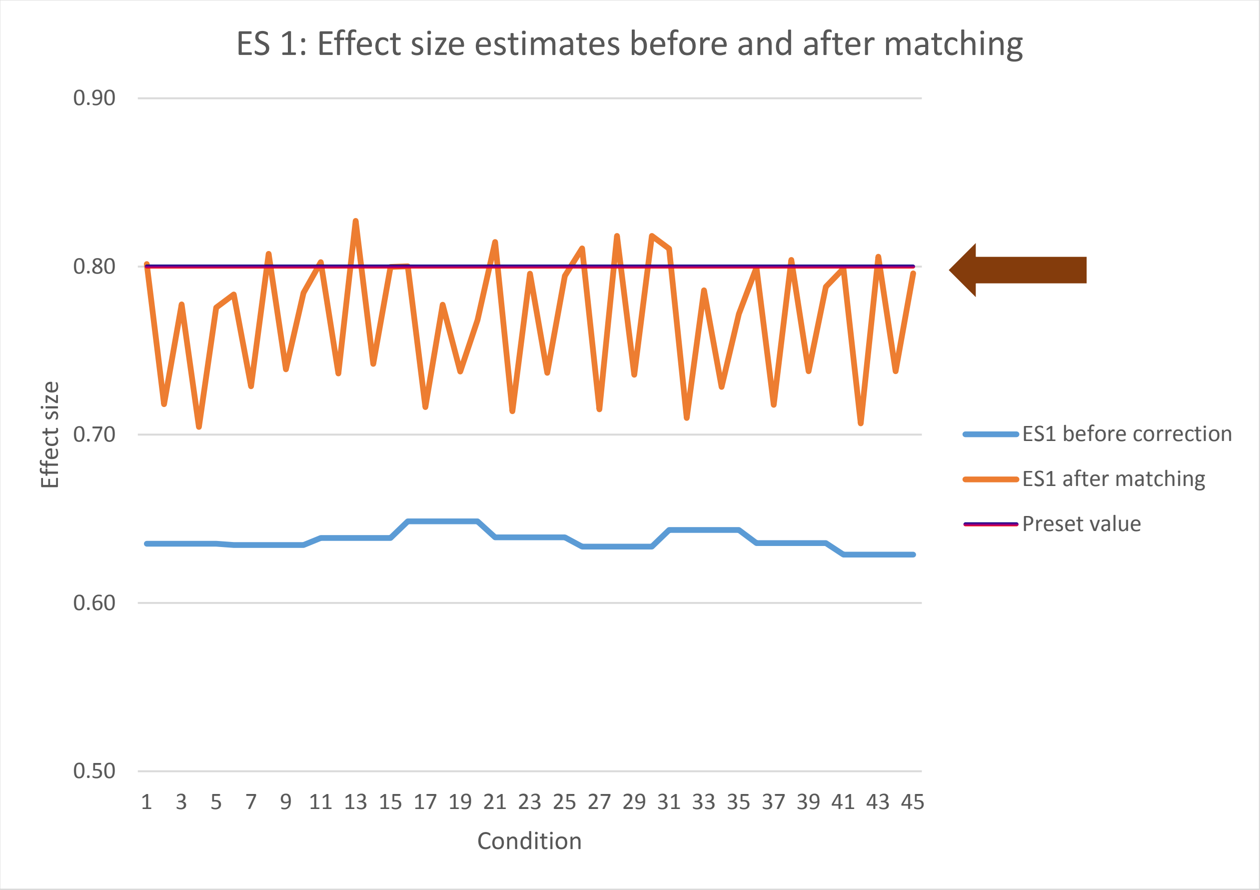
- **Effect size**
  = Mean outcome of treatment group– Mean outcome of control group

- **Interactions** and **main effects** of the four factors on the amount of bias were examined

# How was the performance?

## Propensity score matching

## ES1




After matching, the estimates were closer to the true estimate

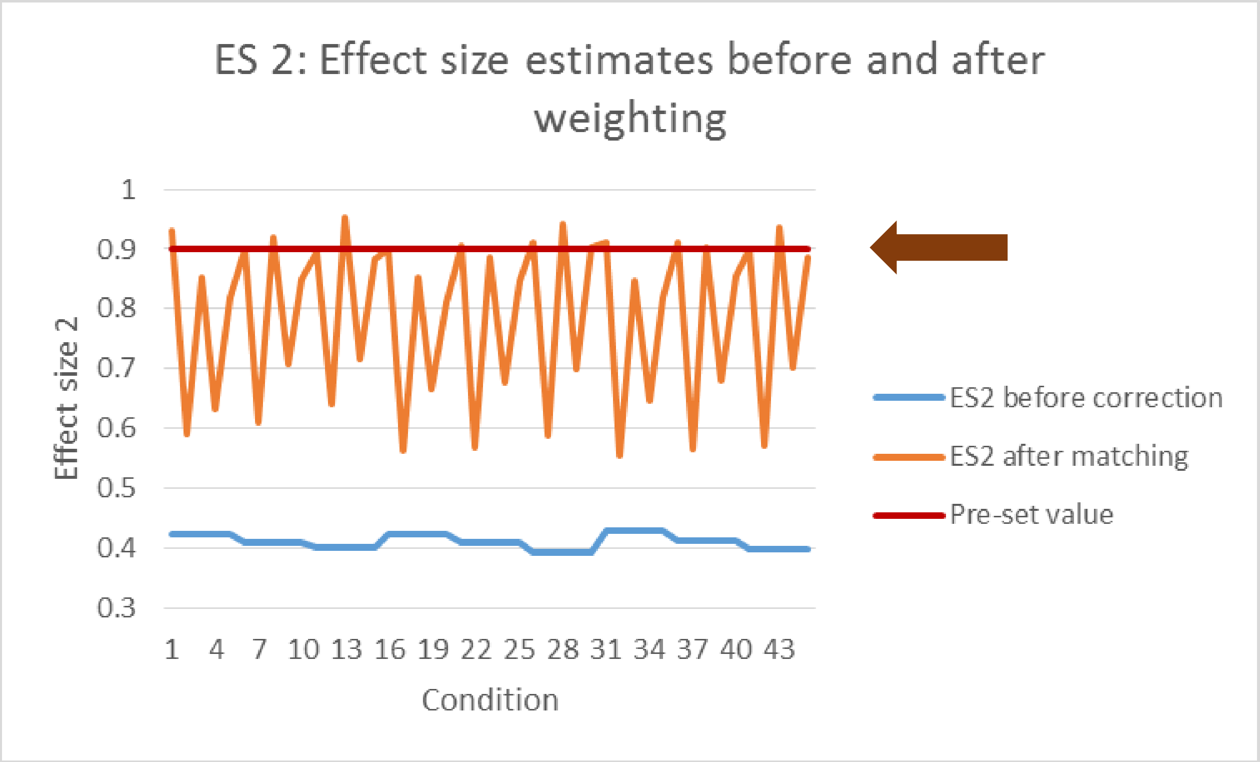
The estimates are initially biased (The estimated were below the true estimate)

Note: ES = Effect size

# How was the performance?

Propensity score matching

## ES2




After matching, the estimate were closer to the true estimate

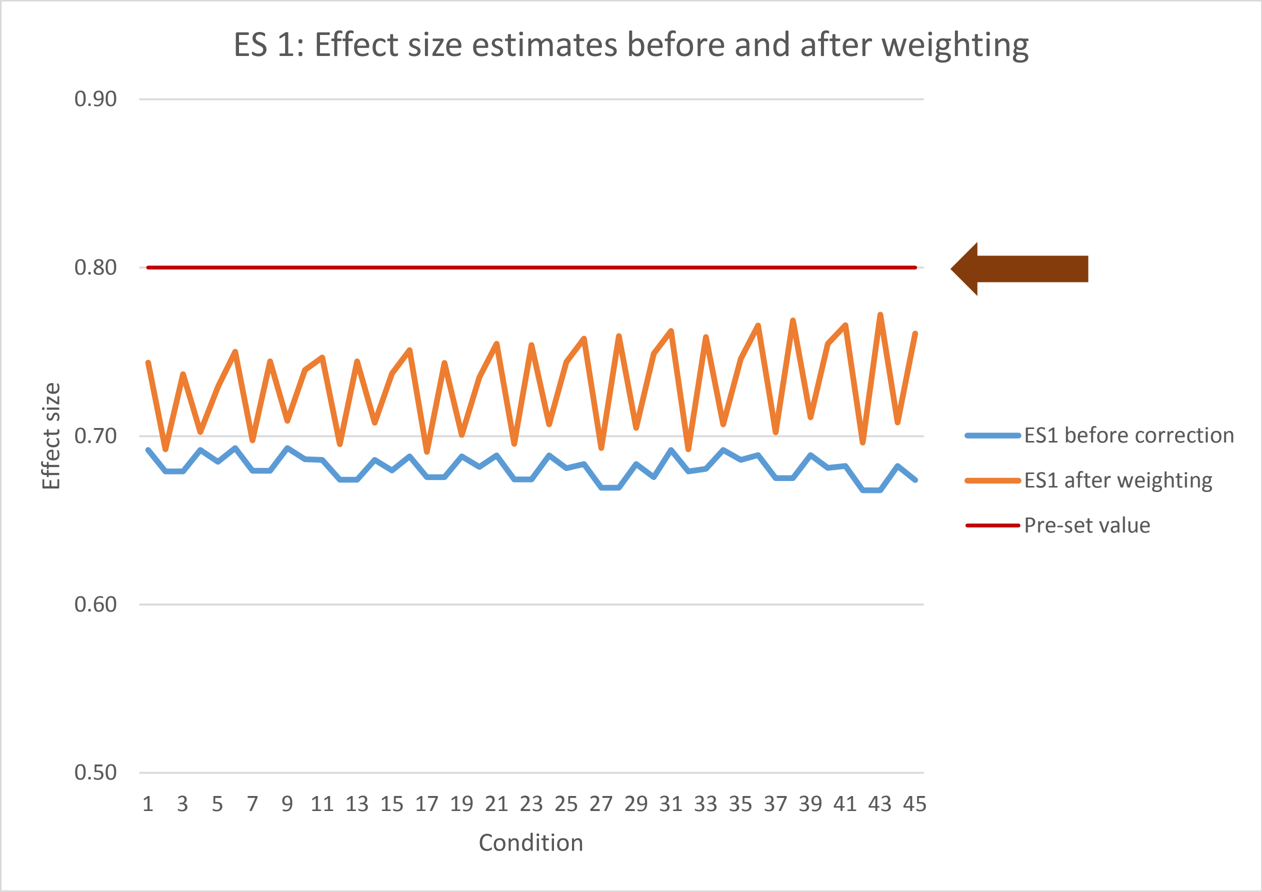
The estimates are initially biased (The estimated were below the true estimate)

Note: ES = Effect size

# How was the performance?

## Propensity score weighting

## ES1




After weighting, the estimate were closer to the true estimate

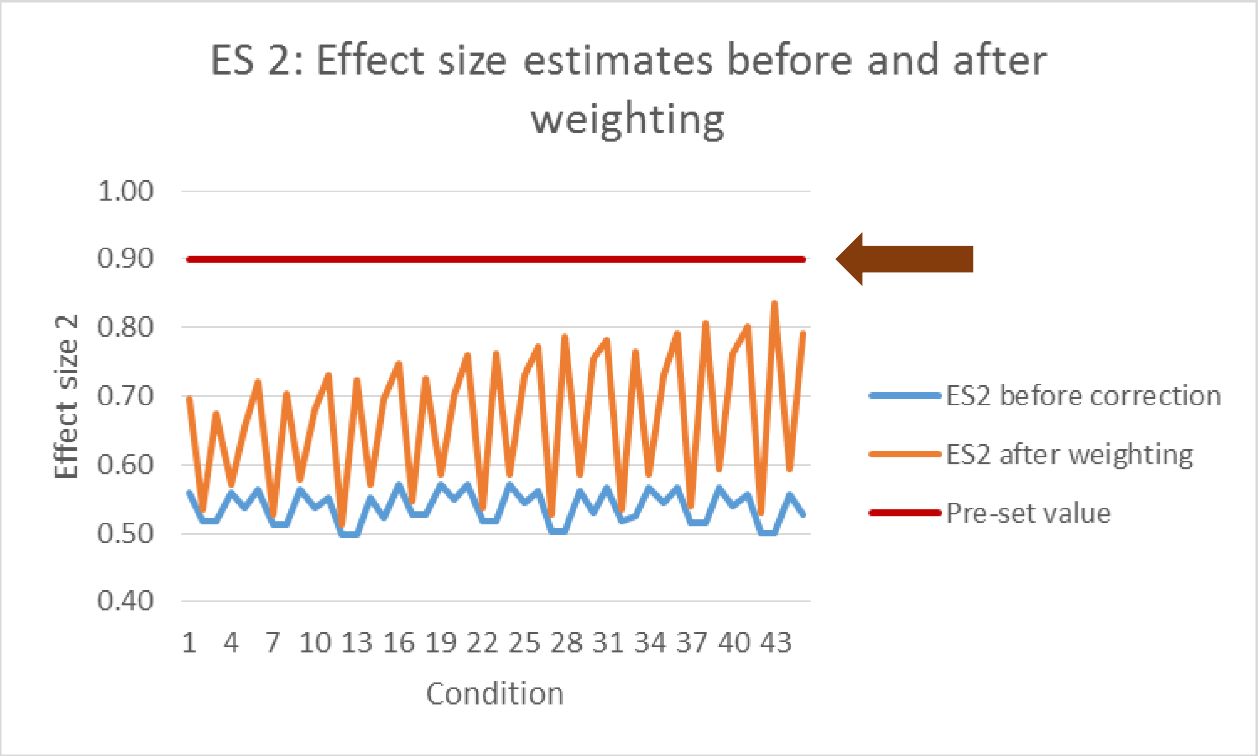
The estimates are initially biased (The estimated were below the true estimate)

Note: ES = Effect size

# How was the performance?

## Propensity score weighting

## ES2




After matching, the estimate were closer to the true estimate

The estimates are initially biased
(The estimated were below the true estimate)

Note: ES = Effect size

# Eta square results

Criteria: eta2 ≥ 0.1

| | Eta square |
|---|---|
| Technique | 0.0043 |
| Sample | 0.0001 |
| Correlation | 0.0003 |
| Normality | 0.0112 |
| TechniqueXSample | 0.0001 |
| TechniqueXCorrelation | 0.0001 |
| SampleXCorrelation | 0.0000 |
| TechniqueXNormality | 0.0003 |
| SampleXNormality | 0.0001 |
| CorrelationXNormality | 0.0002 |
| TechniqueXSampleXCorrelation | 0.0000 |
| TechniqueXSampleXNormality | 0.0001 |
| TechniqueXCorrelationXNormality | 0.0001 |
| SampleXCorrelationXNormality | 0.0001 |
| TechniqueXSampleXCorrelationXNormality | 0.0001 |

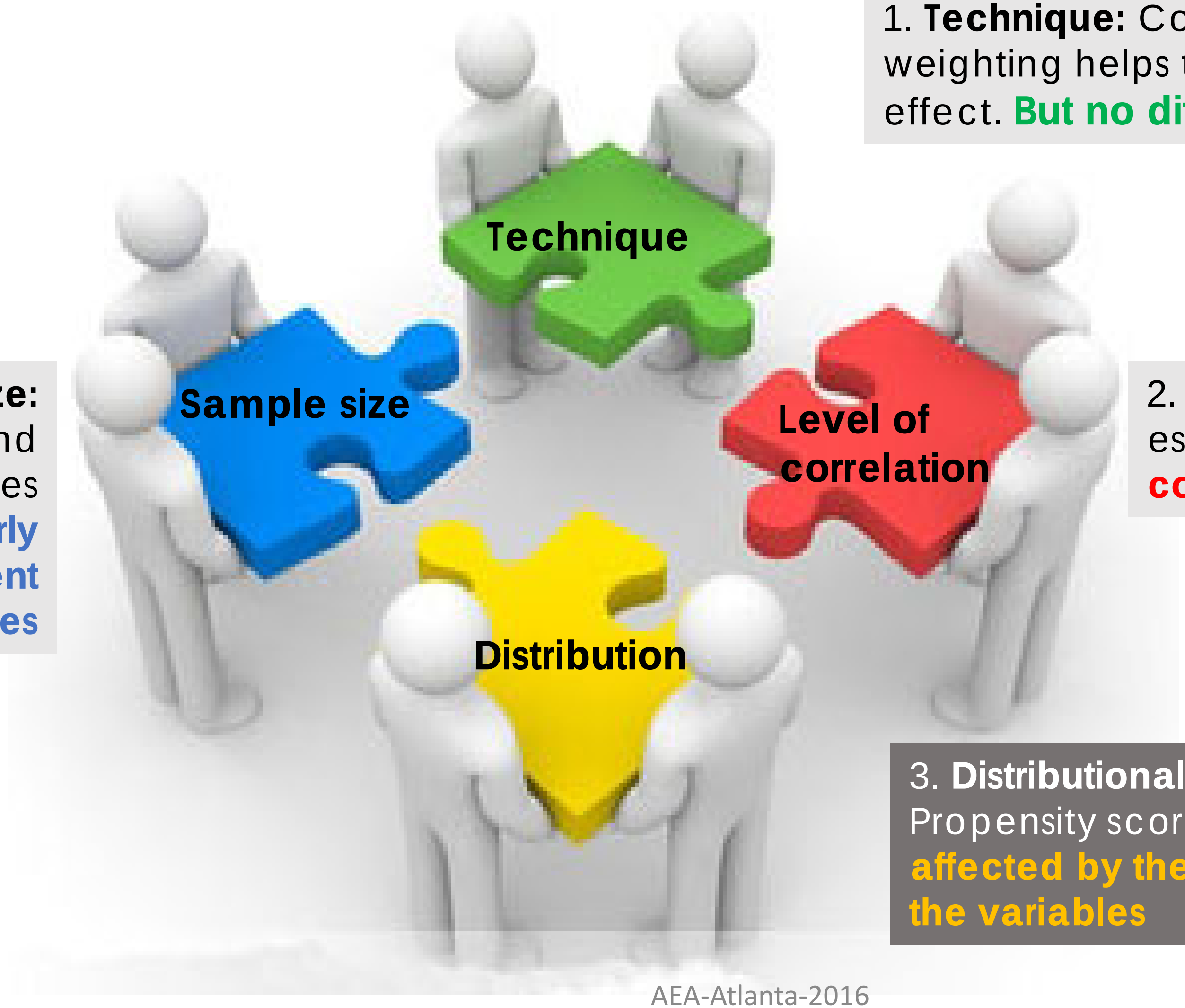
# What does the finding means?

1. **Technique:** Corrections using matching and weighting helps to estimates **TRUE** treatment effect. **But no differences between techniques**

2. **Level of correlation:** Treatment effect estimates are **not affected by the level of correlation between the variables.**

3. **Distributional characteristics of the variable:** Propensity score analysis techniques are **not affected by the distributional characteristics of the variables**

**4. Sample size:** Matching and weighting techniques **performed similarly across different sample sizes**

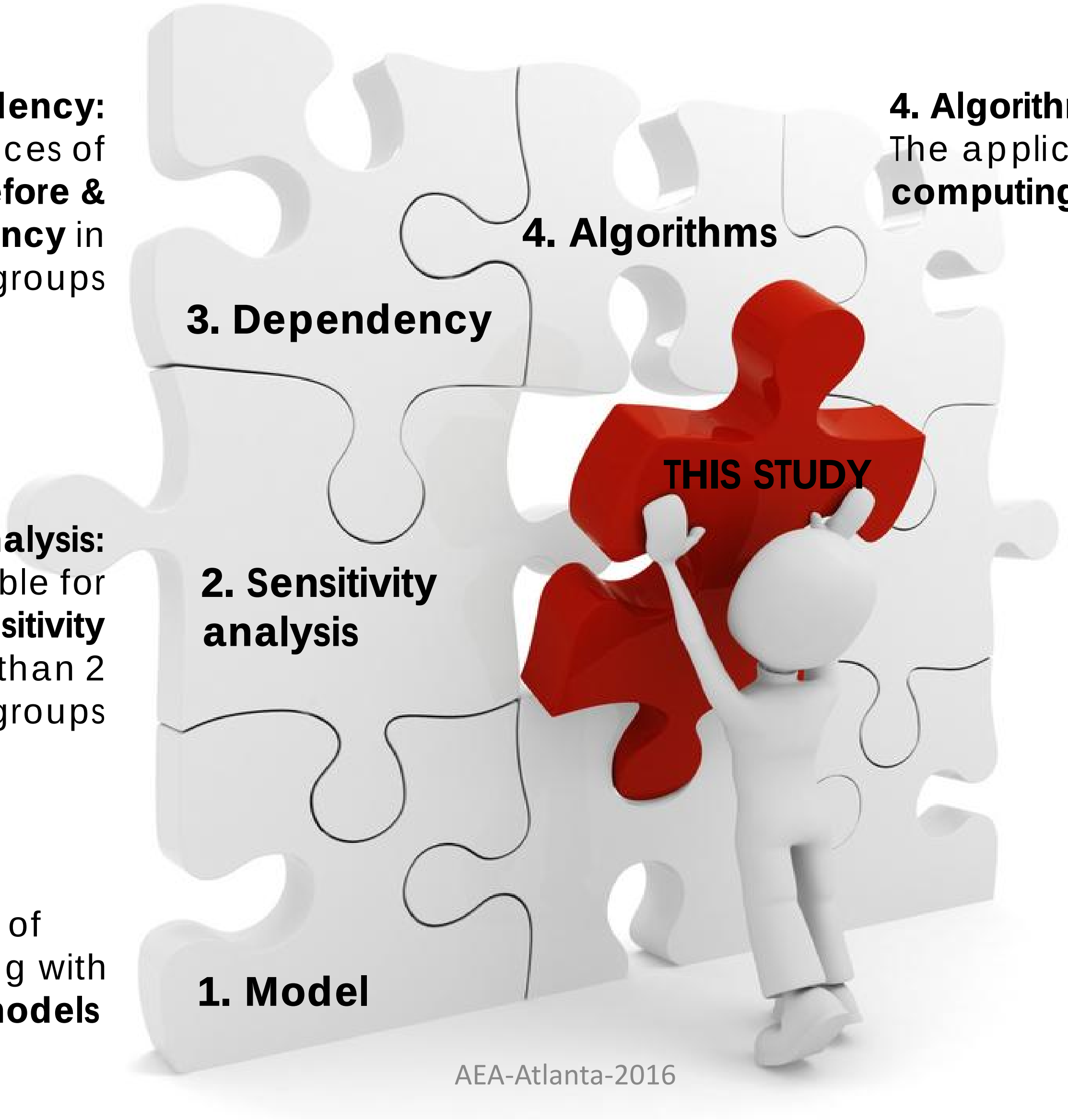


# Next steps

**3. Dependency:**
Determine the consequences of estimating **treatment effect before & after** correcting for **dependency** in multiple treatment groups

**2. Sensitivity analysis:**
No clear direction is available for conducting and assessing **sensitivity analysis** for **hidden bias** in more than 2 groups

**1. Model:** Determine the sensitivity of matching, stratification & weighting with **poorly defined propensity score models**




**4. Algorithms:**
The application of different **algorithms** in **computing propensity scores**

**THIS STUDY:** an initial step in increasing our understanding the application of propensity scores with more than 2 groups.

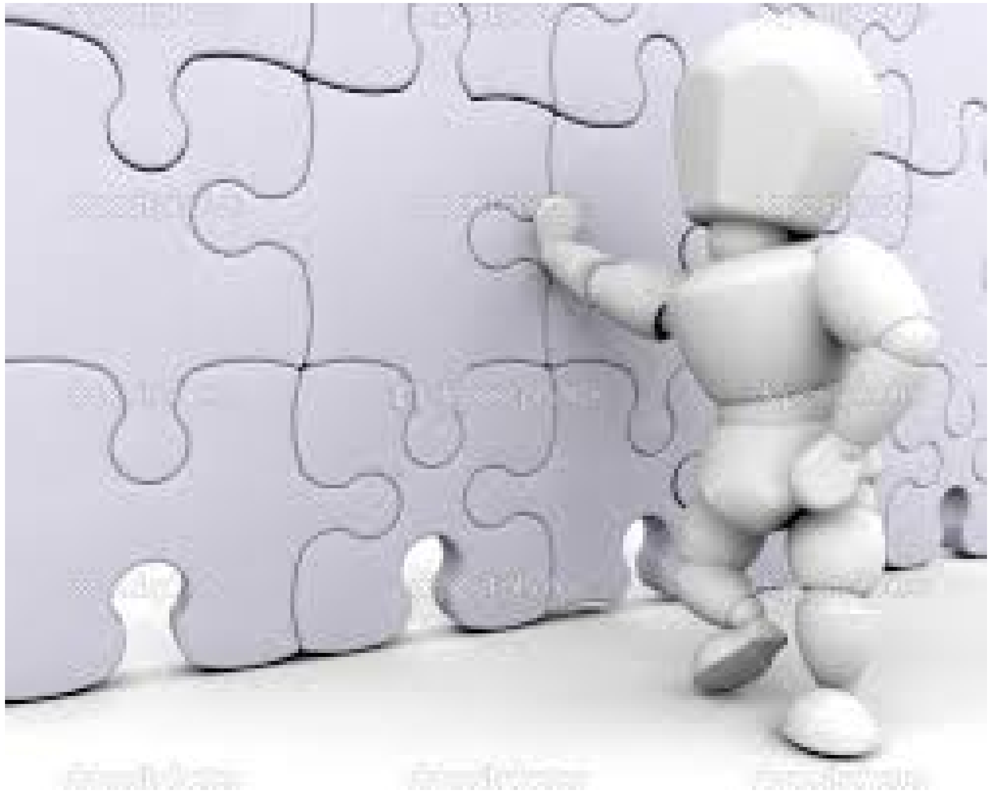
# Thank you

Priyalatha Govindasamy

gpriyalatha@fppm.upsi.edu.my

Antonio Olmos

antonio.olmos@du.edu



http://portfolio.du.edu/antonio/olmos