USING LOG FILES TO IDENTIFY
SEQUENTIAL PATTERNS IN PIAAC
PROBLEM SOLVING ENVIRONMENTS BY
US ADULTS’ EMPLOYMENT STATUS

DISCUSSION

Marco Paccagnella, OECD

Washington, December 6th 2018
Summary of the paper

• Three empirical exercises:
  1. Identify consistent action sequences in different subgroups
  2. Cluster respondents on the basis of action sequences
  3. Check consistency of action sequences across different items

• Three papers?

• Useful characterization in terms of background characteristics of groups of respondents undertaking similar strategies
1. Higher-performing groups more likely to use purposeful actions, lower-performing groups more likely to use redundant/useless actions. Peculiar behaviour of young adults

2. Three clusters, correlated with effort and rate of success

3. About 2/3 of sample showed inconsistent behaviour. Very diverse group
1. Identification of consistent actions in different groups

• Wording: why “employment status”?

• Kitchen-sink approach: what’s the logic?
  – Statistical significance vs. predictive power

• Categories not mutually exclusive
  – Are they all policy-relevant?
1. Identification of consistent actions in different groups

• Why don’t simply group respondents according to PSTRE proficiency/success rate in Uo2?

• Good discussion of the link between actions and underlying cognitive processes. More of that would be even better
2. Identify and characterize clusters of respondents based on their actions

- Part of the paper I liked the most
- Careful how you interpret presence of skilled respondents in each cluster
  - Clustering algorithm is “blind”
  - Only unigrams are used
  - How order of actions is accounted for?
3. Consistency

• Are the items really comparable?
  – Show first that actions are similarly related to outcomes. Actually, some evidence this is not the case

• First step should be to look at characteristics of same groups in the two items
  – You work directly with consistency groups
3. Consistency

• What if you exclude No. actions and time from previous cluster analysis?
• Why four consistency groups and just one inconsistent group?