

# The College-to-Work Transition Study (C2W)

Focused on young adults transitioning from college to the work force, the College-to-Work Transition Study (C2W) involved a multi-year data collection effort where seniors in their final quarter/semester before graduation from universities located in the Pacific Northwest, Midwest, Southeast, and Northeast parts of the USA were invited to participate.

## Code in this repository

This repository contains code to process data collected during the study and investigate hypotheses relating to the time varying change in alcohol use between the *early onboarding phase* and the *late onboarding phase*, moments in time corresponding to when young adults first engage in full time work and at twelve months immediately afterwards, respectively. Specifically, the following were used as outcomes (i.e., dependent variables):

- **Rutgers Alcohol Problem Index (RAPI):** RAPI was assessed with a 19-item measure drawn from a broader 23-item measure developed by White & LaBouvie (White & LaBouvie, 1989). Each item asked participants to rate on a scale of 1 (never) to 5 (More than 10 times) the frequency of a given event happening in the past month while they were drinking or because of their alcohol use. Items included, ‘Got into fights, acted bad, or did mean things’, ‘Felt that you needed more alcohol than you used to have in order to get the same effect’, and ‘Felt that you had a problem with alcohol’. Responses for each item were rescaled to range from 0 (never) to 4 (More than 10 times) and then summed to obtain an overall score.
- **Number of days within the past month with any heavy drinking (HED):** HED was assessed with a one-item measure. Participants were asked, ‘How often in the past month did you drink more than 4 (if you are a woman) or 5 (if you are a man) standard drinks in a single day?’ Responses for each item ranged from 0 (= None in the past month) to 30 (= 30 days in the past month).

## 1. Constructing datasets used in analyses

### Code within the ‘construct-dataset’ folder

File	Brief Description
data-pipeline.R	Documents sequence in which the scripts within the ‘construct-dataset’ folder were run.
construct-variables-baseline.R	Construct variables using raw data collected at baseline (i.e., the ‘baseline dataset’); raw data is in tabular format.
construct-variables-screener.R	Construct variables raw data collected at screening (i.e., the ‘screener dataset’); raw data is in tabular format.
construct-variables-wnw.R	Construct variables using raw data collected after baseline (i.e., the ‘wnw dataset’); raw data is in tabular format.
reshape-data.R	Identify rows within the wnw dataset to associate with the first time an individual reports to be full-time employed and twelve months after that time. In addition, apply exclusion criteria (e.g., exclude individuals who are not full time employed at both of these timepoints).

### Code within the ‘bootstrap’ folder

File	Brief Description
boot-mediation-02a.R	Obtain bootstrap samples to be used for testing Hypothesis H1c
boot-mediation-02b.R	Obtain bootstrap samples to be used for testing Hypothesis H2
boot-mediation-02c.R	Obtain bootstrap samples to be used for testing Hypothesis H3

## 2. Testing hypotheses

### Code within the ‘analysis’ folder

File	Brief Description
analysis-01-RAPI.Rmd	Investigate Hypothesis H1a, H1b, H4, H5 with RAPI as alcohol use outcome. Results are displayed in analysis-01-RAPI.pdf
analysis-02a-RAPI.Rmd	Investigate Hypothesis H1c with RAPI as alcohol use outcome. Results are displayed in analysis-02a-RAPI.pdf
analysis-02bc-RAPI.Rmd	Investigate Hypothesis H2, H3 with RAPI as alcohol use outcome. Results are displayed in analysis-02bc-RAPI.pdf
analysis-01-HED.Rmd	Investigate Hypothesis H1a, H1b, H4, H5 with HED as alcohol use outcome. Results are displayed in analysis-01-HED.pdf
analysis-02a-HED.Rmd	Investigate Hypothesis H1c with HED as alcohol use outcome. Results are displayed in analysis-02a-HED.pdf
analysis-02bc-HED.Rmd	Investigate Hypothesis H2, H3 with HED as alcohol use outcome. Results are displayed in analysis-02bc-HED.pdf
analysis-utils.R	A collection of functions for obtaining summary statistics, estimating moderated effects, and formatting output.

### Intermediate output within the ‘bootstrap’ folder

File	Brief Description
list_models_rutgers_02a.RData	Output of boot-mediation-2a.R
list_models_rutgers_02b.RData	Output of boot-mediation-2b.R
list_models_rutgers_02c.RData	Output of boot-mediation-2c.R
list_models_HED_02a.RData	Output of boot-mediation-2a.R
list_models_HED_02b.RData	Output of boot-mediation-2b.R
list_models_HED_02c.RData	Output of boot-mediation-2c.R

### Code within the ‘explore-dataset’ folder

File	Brief Description
display-possible-range-of-values.Rmd	Displays the range of possible range of values variables can take on. Results are displayed in display-possible-range-of-values.pdf
display-actual-range-of-values.Rmd	Displays the range of actual range of values variables can take on. Results are displayed in display-actual-range-of-values.pdf
display-correlation-matrix.Rmd	Displays the correlation among variables . Results are displayed in display-correlation-matrix.pdf; to enable results to fit in one page, variable names are not displayed. The numbers 1, 2, ..., 23 labeling the rows and columns in this file correspond to the variables listed in order from the 1st, 2nd, ..., 23rd row in the file display-actual-range-of-values.pdf
test-homogeneity.Rmd	Conduct Chi-squared tests (for binary and count variables) or Kolmogorov-Smirnov tests (for continuous variables) to test for homogeneity across two samples; the two samples being compared are 1. Participants having missing data in the outcome (i.e., RAPI/HED), versus 2. Participants having data in the outcome. Results are displayed in test-homogeneity.pdf

### 3. Record of R packages utilized

File	Brief Description
renv.lock	Records the collection of R packages and specific versions utilized.