Package 'CEDMr'

May 7, 2026

Title Capability-Ecological Developmental Model (CEDM) Analysis

Version 0.1.0

Description Implements the Capability-Ecological Developmental Model (CEDM) for longitudinal and multilevel data. The package supports estimation and interpretation of models examining how socioeconomic status (SES), health indicators, and contextual factors jointly relate to academic outcomes. Functionality includes: (1) classification of ecological capability regimes (amplifying, neutral, compensatory); (2) estimation of moderated multilevel models with higher-order interaction terms; (3) causal mediation analysis using doubly robust estimation; (4) random-effects within-between (REWB) decomposition; (5) nonlinear moderation using restricted cubic splines; (6) clustering of longitudinal health trajectories; and (7) sensitivity analysis using the impact threshold for a confounding variable (ITCV) and robustness-to-replacement (RIR) measures. The package is designed for use with general longitudinal multilevel datasets.

License MIT + file LICENSE

URL `https://github.com/causalfragility-lab/CEDMr`

BugReports `https://github.com/causalfragility-lab/CEDMr/issues`

Encoding UTF-8

Language en-US

RoxygenNote 7.3.3

Depends R (>= 4.1.0)

Imports dplyr (>= 1.1.0), ggplot2 (>= 3.4.0), lme4 (>= 1.1-31), magrittr (>= 2.0.0), mediation (>= 4.5.0), rlang (>= 1.1.0), rms (>= 6.7-0), stats, tidyr (>= 1.3.0)

Suggests cluster (>= 2.1.4), konfound (>= 0.4.0), mice (>= 3.16.0), testthat (>= 3.0.0)

Config/testthat/edition 3

NeedsCompilation no

Author Subir Hait [aut, cre] (ORCID: <`https://orcid.org/0009-0004-9871-9677`>)

Maintainer Subir Hait <haitsubi@msu.edu>

Repository CRAN

Date/Publication 2026-04-08 14:30:02 UTC

Contents

`cedm_full_pipeline` *Run the Full CEDM Analysis Pipeline*

Description

A convenience wrapper that runs all major CEDM analytical steps in sequence: regime classification, production function estimation, causal mediation, REWB decomposition, spline moderation, trajectory clustering, and sensitivity analysis. Designed for quick application to new datasets (e.g., any study with longitudinal SES, health, and outcome data).

Usage

```
cedm_full_pipeline(
  data,
  outcome_var,
  ses_var,
  health_var,
  id_var,
  time_var,
  opportunity_var = NULL,
  cluster_var = NULL,
  covariates = NULL,
  n_boot = 500,
  k_trajectories = 3,
  run_mediation = TRUE,
  run_spline = TRUE,
```

```
  run_trajectory = TRUE,
  seed = 123,
  verbose = TRUE
)
```

Arguments

`data`	A data.frame in LONG format.
`outcome_var`	Character string: dependent variable.
`ses_var`	Character string: SES predictor.
`health_var`	Character string: health variable (e.g., BMI).
`id_var`	Character string: person-level ID (for REWB and trajectory).
`time_var`	Character string: wave/time variable.
`opportunity_var`	Character string (optional): school/neighborhood opportunity index for regime classification.
`cluster_var`	Character string (optional): school/cluster ID for random intercepts.
`covariates`	Character vector of additional covariate names.
`n_boot`	Integer: bootstrap replications for mediation. Default 500.
`k_trajectories`	Integer: number of health trajectory clusters. Default 3.
`run_mediation`	Logical: whether to run mediation (can be slow). Default TRUE.
`run_spline`	Logical: whether to run spline moderation. Default TRUE.
`run_trajectory`	Logical: whether to run trajectory clustering. Default TRUE.
`seed`	Integer: random seed. Default 123.
`verbose`	Logical: print progress messages. Default TRUE.

Value

A named list of class `"cedm_pipeline"` containing outputs from each analysis step:

- `data`: data with regime classification added.
- `regime_table`: frequency table of capability regimes.
- `model_base`: CEDM base production model (SES x health).
- `model_regime`: CEDM full regime model (SES x health x regime).
- `mediation`: causal mediation results (if run).
- `rewb`: within-between decomposition results.
- `spline`: nonlinear spline moderation results (if run).
- `trajectory`: health trajectory clustering (if run).
- `sensitivity`: sensitivity analysis for key terms.
- `plots`: named list of ggplot2 objects.

Examples

```
sim <- cedm_simulate(n = 3000, n_waves = 5, seed = 42)
pipeline <- cedm_full_pipeline(
  data        = sim,
  outcome_var = "achievement",
  ses_var     = "SES",
  health_var  = "health",
  id_var      = "id",
  time_var    = "wave",
  n_boot      = 200
)
summary(pipeline)
pipeline$plots$regimes
pipeline$plots$interaction
```

`cedm_mediation` *CEDM Causal Mediation Analysis*

Description

Tests whether a health indicator (e.g., BMI) mediates the relationship between SES and academic achievement using the counterfactual framework of Imai et al. (2010). Implements the doubly robust design used in Hait (2026) with nonparametric bootstrap confidence intervals and built-in sensitivity analysis. Consistent with CEDM Proposition 1: mediation effects are expected to be small and unstable; moderation (see `cedm_production`) is expected to dominate.

Usage

```
cedm_mediation(
  data,
  outcome_var,
  ses_var,
  health_var,
  covariates = NULL,
  n_boot = 1000,
  conf_level = 0.95,
  sensitivity = TRUE,
  seed = 123
)
```

Arguments

`data`	A data.frame.
`outcome_var`	Character string: dependent variable (academic achievement).
`ses_var`	Character string: treatment/exposure variable (SES).

`health_var`	Character string: mediator variable (e.g., BMI).
`covariates`	Character vector of covariate names for both models.
`n_boot`	Integer: number of bootstrap replications. Default 1000.
`conf_level`	Numeric: confidence level. Default 0.95.
`sensitivity`	Logical: if TRUE, runs sensitivity analysis via `medsens()` to assess robustness to unmeasured confounding. Default TRUE.
`seed`	Integer: random seed for reproducibility. Default 123.

Details

The mediation model follows the two-equation structure used in the ECLS-K analysis (Hait, 2026):

$$BMI = \alpha + \gamma \times SES + \delta^\top \times Covariates + \varepsilon_1$$

$$Math = \beta + \tau' \times SES + \zeta \times BMI + \theta^\top \times Covariates + \varepsilon_2$$

ACME = $\gamma \times \zeta$; ADE = τ'.

Per CEDM Proposition 1, a small ACME relative to the total SES effect (proportion mediated < 5\ result, indicating that health does not function as a causal conduit but as a conditional conversion moderator.

Value

A list of class `"cedm_mediation"` with elements:

- `mediation_result`: output from `mediation::mediate()`.
- `summary`: summary of mediation results.
- `acme`: Average Causal Mediation Effect.
- `ade`: Average Direct Effect.
- `total_effect`: Total causal effect of SES.
- `prop_mediated`: Proportion of total effect mediated through health.
- `sensitivity`: sensitivity analysis results (if requested).
- `interpretation`: automated CEDM-based interpretation string.

References

Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, 15(4), 309-334.

Hait, S. (2026). Socioeconomic Status, Health, and Academic Achievement: A Capability-Ecological Developmental Model.

Examples

```
set.seed(42)
df <- data.frame(
  math = rnorm(500, 500, 100),
  ses  = rnorm(500),
  bmi  = rnorm(500, 25, 5),
  sex  = sample(0:1, 500, replace = TRUE)
)
result <- cedm_mediation(df, outcome_var = "math", ses_var = "ses",
                         health_var = "bmi", covariates = "sex",
                         n_boot = 200)
print(result)
```

`cedm_production` *Fit the CEDM Capability-Conversion Production Function*

Description

Estimates the CEDM production function relating SES, health (e.g., BMI), and ecological capability regime to academic achievement. Supports the base two-way (SES x Health) model and the full three-way (SES x Health x Regime) model described in the theory paper. Uses multilevel (mixed-effects) or single-level OLS depending on the presence of cluster variables.

Usage

```
cedm_production(
  data,
  outcome_var,
  ses_var,
  health_var,
  regime_var = NULL,
  covariates = NULL,
  cluster_var = NULL,
  model = c("base", "regime", "additive"),
  center = TRUE,
  weights = NULL
)
```

Arguments

`data`	A data.frame.
`outcome_var`	Character string: dependent variable (e.g., math score).
`ses_var`	Character string: SES predictor.
`health_var`	Character string: health indicator (e.g., BMI).

`regime_var`	Character string: ecological regime variable (typically `"cedm_regime"` produced by `classify_regime()`). If NULL, the two-way SES x health model is estimated.
`covariates`	Character vector of additional covariate names.
`cluster_var`	Character string: grouping/cluster variable (e.g., school ID) for random intercepts. If NULL, a standard OLS model is fitted.
`model`	One of `"base"` (SES + health + SES:health), `"regime"` (full three-way with regime), or `"additive"` (no interaction, main effects only).
`center`	Logical. If TRUE (default), SES and health are mean-centered before fitting.
`weights`	Character string naming a survey weight variable, or NULL.

Details

Base model (Proposition 1 + 4):

$$A_i = \beta_0 + \beta_1 SES_i + \beta_2 H_i + \beta_3 (SES_i \times H_i) + \varepsilon$$

A significant, negative β_3 indicates that health constraints steepen the SES-achievement gradient for low-SES children (amplifying effect).

Full regime model (Proposition 2):

$$A_i = \beta_0 + \beta_1 SES_i + \beta_2 H_i + \beta_3 (SES_i \times H_i) + \beta_4 (SES_i \times H_i \times C_j) + \varepsilon$$

The three-way interaction β_4 is the empirical index of ecological capability regime type.

Value

A list of class `"cedm_production"` with elements:

- `model`: the fitted model object (`lm` or `lmerMod`).
- `summary`: model summary.
- `formula`: the formula used.
- `centered`: logical, whether centering was applied.
- `coefficients`: tidy coefficient table.

References

Hait, S. (2025). Socioeconomic Status, Health, and Academic Achievement: A Capability-Ecological Developmental Model.

Examples

```
set.seed(42)
df <- data.frame(
  math  = rnorm(300, 500, 100),
  ses   = rnorm(300),
  bmi   = rnorm(300, 25, 5),
  school = sample(1:30, 300, replace = TRUE)
```

```
)
df <- classify_regime(df, ses_var = "ses")
result <- cedm_production(df, outcome_var = "math", ses_var = "ses",
                          health_var = "bmi", regime_var = "cedm_regime",
                          cluster_var = "school")
print(result)
```

`cedm_rewb` *Random-Effects Within-Between (REWB) Decomposition*

Description

Decomposes a health indicator (e.g., BMI) into between-person (stable, chronic) and within-person (transient, time-varying) components to test CEDM Proposition 3 (Developmental Recursion). This mirrors the REWB design used in Hait (2025) for the ECLS-K:2011 data.

Usage

```
cedm_rewb(
  data,
  outcome_var,
  health_var,
  id_var,
  time_var,
  ses_var = NULL,
  covariates = NULL,
  cluster_var = NULL
)
```

Arguments

`data`	A data.frame in LONG format.
`outcome_var`	Character string: dependent variable (e.g., math score).
`health_var`	Character string: health variable to decompose (e.g., BMI).
`id_var`	Character string: person-level ID variable.
`time_var`	Character string: time/wave indicator variable.
`ses_var`	Character string: SES predictor (included as covariate).
`covariates`	Character vector of additional covariate names.
`cluster_var`	Character string: school or higher-level cluster variable for additional random intercept. If NULL, only person-level random intercepts are modelled.

Details

Per the CEDM, between-child health differences (chronic status) are expected to have significant associations with achievement, while within-child fluctuations are expected to be non-significant – indicating that health functions as a stable developmental risk factor rather than a time-varying dynamic predictor.

The between component is each person's mean of the health variable across waves. The within component is the wave-specific deviation from that mean:

$$H_{between,i} = \bar{H}_i$$

$$H_{within,it} = H_{it} - \bar{H}_i$$

The model then includes both components as separate predictors:

$$Y_{it} = \beta_0 + \beta_1 SES_i + \beta_2 H_{between,i} + \beta_3 H_{within,it} + (1|id) + \varepsilon$$

Value

A list of class `"cedm_rewb"` with:

- `model`: fitted `lmerMod` object.
- `summary`: model summary.
- `between_effect`: coefficient for between-person health variable.
- `within_effect`: coefficient for within-person health variable.
- `data_decomposed`: original data with added `_between` and `_within` columns.
- `interpretation`: CEDM interpretation string.

References

Curran, P. J., Howard, A. L., Bainter, S. A., Lane, S. T., & McGinley, J. S. (2014). The separation of between-person and within-person components of individual change over time. *Journal of Consulting and Clinical Psychology*.

Hait, S. (2025). Socioeconomic Status, Health, and Academic Achievement: A Capability-Ecological Developmental Model.

Examples

```
set.seed(42)
df_long <- data.frame(
  id     = rep(1:200, each = 5),
  wave   = rep(1:5, times = 200),
  math   = rnorm(1000, 500, 100),
  bmi    = rnorm(1000, 25, 5),
  ses    = rep(rnorm(200), each = 5),
  school = rep(sample(1:20, 200, replace = TRUE), each = 5)
)
result <- cedm_rewb(df_long, outcome_var = "math", health_var = "bmi",
                    id_var = "id", time_var = "wave", ses_var = "ses",
                    cluster_var = "school")
print(result)
```

`cedm_sensitivity` *CEDM Sensitivity Analysis: Frank's ITCV and Robustness-to-Replacement (RIR)*

Description

Computes sensitivity metrics for CEDM regression effects using Frank's (2000) Impact Threshold for a Confounding Variable (ITCV) and the Robustness of Inference to Replacement (RIR) index. Within the CEDM, these metrics serve not merely as statistical diagnostics but as indicators of CAPABILITY STABILITY (Proposition 5): large ITCV/RIR = stable conversion processes; small ITCV/RIR = fragile conversion processes, especially in amplifying regimes.

Usage

```
cedm_sensitivity(
  model,
  term = NULL,
  n_obs = NULL,
  alpha = 0.05,
  benchmark = NULL
)
```

Arguments

`model`	A fitted `lm` or `lmerMod` object, OR the output of `cedm_production()`.
`term`	Character string naming the coefficient of interest (e.g., "ses_c", "ses_c:bmi_c"). If NULL, all terms except the intercept are evaluated.
`n_obs`	Integer: number of observations used in the model. Required when `model` is a `lmerMod` and `nobs()` may not be available.
`alpha`	Numeric: significance threshold. Default 0.05.
`benchmark`	Character string: label for the benchmark effect size (e.g., "SES main effect"). Used in the interpretation.

Details

The ITCV is the minimum correlation an omitted confounder would need with both the treatment and the outcome to nullify the observed effect:

$$ITCV = \frac{t^2 - t_{crit}^2}{t^2 \cdot (n-q) + t_{crit}^2}$$

where t is the observed t-statistic, t_{crit} is the critical t-value, n is sample size, and q is the number of parameters.

The RIR is the number (or percentage) of observations that would need to be replaced with cases showing no effect to nullify the inference.

Per CEDM Proposition 5:

- High ITCV/RIR -> Capability stability: conversion processes are robust to structural disruption.
- Low ITCV/RIR -> Capability fragility: conversion processes are easily disrupted, especially in amplifying ecological regimes.

Value

A data.frame of class `"cedm_sensitivity"` with one row per evaluated term, containing:

- `term`: predictor name.
- `estimate`: regression coefficient.
- `std_error`: standard error.
- `t_value`: t-statistic.
- `p_value`: p-value.
- `itcv`: Impact Threshold for a Confounding Variable.
- `rir_pct`: Robustness of Inference to Replacement (percent).
- `rir_n`: RIR in number of observations.
- `stability`: Qualitative label ("extremely robust", "robust", "moderately robust", "sensitive", "very sensitive").
- `cedm_interpretation`: CEDM-specific capability stability label.

References

Frank, K. A. (2000). Impact of a confounding variable on a regression coefficient. *Sociological Methods & Research*, 29(2), 147-194.

Frank, K. A., Maroulis, S. J., Duong, M. Q., & Kelcey, B. M. (2013). What would it take to change an inference? Using Rubin's causal model to interpret the robustness of causal inferences. *Educational Evaluation and Policy Analysis*, 35(4), 437-460.

Hait, S. (2026). Socioeconomic Status, Health, and Academic Achievement: A Capability-Ecological Developmental Model.

Examples

```
set.seed(42)
df <- data.frame(
  math = rnorm(500, 500, 100),
  ses  = rnorm(500),
  bmi  = rnorm(500, 25, 5)
)
fit  <- lm(math ~ ses * bmi, data = df)
sens <- cedm_sensitivity(fit, term = "ses")
print(sens)
```

`cedm_simulate` *Simulate Data Under CEDM Ecological Capability Regimes*

Description

Generates synthetic longitudinal data from the CEDM toy data-generating process, replicating Appendix A of Hait (2025). This simulation encodes three ecological capability regimes and confirms the CEDM's core prediction: weak mediation (small indirect effect) alongside strong and sign-varying SES x health moderation across regimes.

Usage

```
cedm_simulate(
  n = 3000,
  alpha1 = 0.15,
  beta1 = 0.5,
  beta2 = 0.1,
  beta3_neutral = 0,
  beta3_amplifying = 0.6,
  beta3_compensatory = -0.3,
  alpha_reg_amplifying = 0.3,
  alpha_reg_compensatory = -0.3,
  n_waves = 1,
  seed = 123
)
```

Arguments

`n`	Integer: total sample size. Default 3000 (1000 per regime).
`alpha1`	Numeric: SES -> health effect (small = weak mediation). Default 0.15.
`beta1`	Numeric: main SES -> achievement effect. Default 0.50.
`beta2`	Numeric: main health -> achievement effect. Default 0.10.
`beta3_neutral`	Numeric: SES x health interaction in neutral regime. Default 0.00.
`beta3_amplifying`	Numeric: SES x health interaction in amplifying regime. Default 0.60 (positive = health amplifies low-SES disadvantage when SES is negatively coded, or use negative values depending on your parameterization).
`beta3_compensatory`	Numeric: SES x health interaction in compensatory regime. Default -0.30.
`alpha_reg_amplifying`	Numeric: regime-specific intercept shift for health in amplifying regime. Default 0.30.
`alpha_reg_compensatory`	Numeric: regime-specific intercept shift for health in compensatory regime. Default -0.30.

`n_waves`	Integer: number of longitudinal waves to generate. Default 1 (cross-sectional). Set > 1 for longitudinal data.
`seed`	Integer: random seed. Default 123.

Details

The data-generating process is:

$$Health_i = \alpha_0 + \alpha_1 SES_i + \alpha_{regime} + \varepsilon_M$$

$$Achievement_i = \beta_0 + \beta_1 SES_i + \beta_2 Health_i + \beta_{3,regime} SES_i \times Health_i + \varepsilon_Y$$

where regime-specific parameters encode amplifying, neutral, and compensatory ecological contexts.

Value

A data.frame with columns: `id`, `SES`, `health`, `achievement`, `regime`, and (if `n_waves > 1`) `wave`.

References

Hait, S. (2025). Socioeconomic Status, Health, and Academic Achievement: A Capability-Ecological Developmental Model. Appendix A.

Examples

```
sim_data <- cedm_simulate(n = 3000, seed = 42)
table(sim_data$regime)
head(sim_data)

# Run full CEDM analysis on simulated data
sim_data <- classify_regime(sim_data, ses_var = "SES",
                            opportunity_var = NULL, ses_tertiles = TRUE)
prod <- cedm_production(sim_data, outcome_var = "achievement",
                        ses_var = "SES", health_var = "health",
                        regime_var = "cedm_regime", model = "regime")
print(prod)
```

`cedm_spline_moderation` *Nonlinear Moderation via Restricted Cubic Splines (CEDM Proposition 1 & 2)*

Description

Fits a multilevel nonlinear moderation model using restricted cubic splines (RCS) to capture threshold and nonlinear effects of health on the SES-achievement relationship. Implements the spline-based approach used in Hait (2026) for detecting nonlinearities in the BMI-achievement link that are invisible in linear models.

Usage

```
cedm_spline_moderation(
  data,
  outcome_var,
  ses_var,
  health_var,
  df = 5,
  covariates = NULL,
  cluster_var = NULL,
  interaction = TRUE,
  plot = TRUE
)
```

Arguments

`data`	A data.frame.
`outcome_var`	Character string: dependent variable.
`ses_var`	Character string: SES predictor.
`health_var`	Character string: health variable to spline-transform (e.g., BMI).
`df`	Integer: degrees of freedom for the restricted cubic spline. Default 5.
`covariates`	Character vector of covariate names.
`cluster_var`	Character string: cluster variable for random intercepts. If NULL, OLS is used.
`interaction`	Logical: if TRUE (default), include SES x spline(health) interaction terms to model nonlinear moderation.
`plot`	Logical: if TRUE (default), generate a marginal effects plot.

Details

Restricted cubic splines allow the health-achievement relationship to be nonlinear and threshold-based – exactly the pattern predicted by the CEDM for amplifying contexts, where health constraints accelerate sharply at the upper end of the health-risk distribution.

Value

A list of class `"cedm_spline"` with:

- `model`: fitted model object.
- `summary`: model summary.
- `anova_test`: ANOVA test for overall nonlinearity.
- `plot`: ggplot2 object showing predicted achievement vs health by SES level (if `plot = TRUE`).

References

Harrell, F. E. (2015). *Regression Modeling Strategies*. Springer.

Hait, S. (2026). Socioeconomic Status, Health, and Academic Achievement: A Capability-Ecological Developmental Model.

Examples

```
set.seed(42)
df <- data.frame(
  math   = rnorm(400, 500, 100),
  ses    = rnorm(400),
  bmi    = rnorm(400, 25, 5),
  school = sample(1:40, 400, replace = TRUE)
)
result <- cedm_spline_moderation(df, outcome_var = "math",
                                  ses_var = "ses", health_var = "bmi",
                                  cluster_var = "school")
print(result)
```

`cedm_trajectory` *Longitudinal Health Trajectory Clustering (CEDM Proposition 3)*

Description

Identifies developmental health phenotypes (trajectory classes) using k-means or hierarchical clustering on person-level longitudinal health profiles. Replicates the BMI trajectory analysis from Hait (2025) which identified stable-average, persistently-low, and high-rising BMI classes. Supports CEDM Proposition 3 (Developmental Recursion) by identifying children on cumulative health trajectories.

Usage

```
cedm_trajectory(
  data,
  health_var,
  id_var,
  time_var,
  k = 3,
  method = c("kmeans", "hierarchical"),
  outcome_var = NULL,
  ses_var = NULL,
  seed = 123,
  plot = TRUE
)
```

Arguments

`data`	A data.frame in LONG format.
`health_var`	Character string: health variable (e.g., BMI).
`id_var`	Character string: person-level ID variable.
`time_var`	Character string: wave/time variable.

`k`	Integer: number of trajectory clusters. Default 3 (reflecting the CEDM's three-class structure: stable-average, low, high-rising).
`method`	Character: clustering method, one of `"kmeans"` (default) or `"hierarchical"`.
`outcome_var`	Character string (optional): if provided, mean outcome is computed by cluster for interpretation.
`ses_var`	Character string (optional): if provided, mean SES is computed by cluster.
`seed`	Integer: random seed. Default 123.
`plot`	Logical: if TRUE (default), generate trajectory plot.

Value

A list of class `"cedm_trajectory"` with:

- `cluster_assignment`: data.frame with id and cluster label.
- `cluster_profiles`: mean health by wave and cluster.
- `cluster_summary`: mean outcome and SES by cluster (if provided).
- `plot`: ggplot2 trajectory plot (if `plot = TRUE`).
- `k`: number of clusters.

Examples

```
set.seed(42)
df <- data.frame(
  id   = rep(1:200, each = 5),
  wave = rep(1:5, times = 200),
  bmi  = c(rnorm(200 * 5, 25, 3)),
  math = rnorm(200 * 5, 500, 100),
  ses  = rep(rnorm(200), each = 5)
)
result <- cedm_trajectory(df, health_var = "bmi", id_var = "id",
                          time_var = "wave", outcome_var = "math",
                          ses_var = "ses")
print(result)
```

`classify_regime` *Classify Ecological Capability Regimes*

Description

Assigns each observation to one of three CEDM ecological capability regimes (amplifying, neutral, compensatory) based on individual-level SES and school/context-level opportunity index. Implements the formal operationalization from Proposition 2 of the CEDM (Hait, 2025).

Usage

```
classify_regime(
  data,
  ses_var,
  opportunity_var = NULL,
  ses_cutpoint = NULL,
  opportunity_cutpoint = NULL,
  method = c("hard", "continuous"),
  ses_tertiles = FALSE
)
```

Arguments

`data`	A data.frame containing the variables specified below.
`ses_var`	Character string naming the SES variable (numeric).
`opportunity_var`	Character string naming the ecological opportunity variable (e.g., school resources index, neighborhood opportunity score). If NULL, classification is based on SES alone using tertile cutpoints.
`ses_cutpoint`	Numeric cutpoint for SES. Defaults to the sample median.
`opportunity_cutpoint`	Numeric cutpoint for the opportunity variable. Defaults to the sample median.
`method`	One of `"hard"` (discrete three-category assignment) or `"continuous"` (returns a continuous capability index). Default is `"hard"`.
`ses_tertiles`	Logical. If TRUE and `opportunity_var` is NULL, classify regimes using SES tertiles only. Default FALSE.

Details

The CEDM defines three ecological capability regimes (Hait, 2025):

- **Amplifying**: Low SES AND low ecological opportunity. Health constraints intensify the academic penalty of low SES.
- **Neutral**: Moderate SES and/or moderate opportunity. Weak SES x health moderation.
- **Compensatory**: High SES AND high ecological opportunity. Ecological supports buffer health constraints.

Formally: R_ij = Amplifying if SES_i < c_SES and O_j < c_O; Compensatory if SES_i >= c_SES and O_j >= c_O; Neutral otherwise.

Value

The original data.frame with an added `cedm_regime` column (factor: "amplifying", "neutral", "compensatory") and, when `method = "continuous"`, an additional `cedm_capability_index` column.

References

Hait, S. (2025). Socioeconomic Status, Health, and Academic Achievement: A Capability-Ecological Developmental Model. OSF Preprints.

Examples

```
set.seed(42)
df <- data.frame(
  ses   = rnorm(500),
  opp   = rnorm(500),
  bmi   = rnorm(500, 25, 5),
  math  = rnorm(500, 100, 15)
)
df <- classify_regime(df, ses_var = "ses", opportunity_var = "opp")
table(df$cedm_regime)
```

`plot_cedm_interaction` *Plot CEDM Interaction: SES x Health by Regime*

Description

Creates a paneled interaction plot showing predicted achievement as a function of health at different SES levels, separately for each ecological capability regime. This is the core visualization for CEDM Proposition 2.

Usage

```
plot_cedm_interaction(
  cedm_prod_result,
  data,
  ses_var,
  health_var,
  regime_var,
  outcome_var,
  n_points = 50
)
```

Arguments

`cedm_prod_result`	Output from `cedm_production()` with `model = "regime"`.
`data`	The original data.frame used to fit the model.
`ses_var`	Character string: SES variable name.
`health_var`	Character string: health variable name.
`regime_var`	Character string: regime variable name.
`outcome_var`	Character string: outcome variable name.
`n_points`	Integer: number of health values for prediction grid. Default 50.

Value

A ggplot2 object.

`plot_regimes` *Plot Ecological Capability Regimes*

Description

Generates a scatter plot of SES vs. a health variable, colored by CEDM ecological capability regime, with optional outcome overlaid as point size.

Usage

```
plot_regimes(
  data,
  ses_var,
  health_var,
  regime_var = "cedm_regime",
  outcome_var = NULL,
  alpha_pt = 0.5,
  title = "CEDM Ecological Capability Regimes"
)
```

Arguments

`data`	A data.frame with regime classification (output of `classify_regime()`).
`ses_var`	Character string: SES variable name.
`health_var`	Character string: health variable name.
`regime_var`	Character string: regime variable name. Default `"cedm_regime"`.
`outcome_var`	Character string (optional): if provided, points are sized by this variable.
`alpha_pt`	Numeric: point transparency. Default 0.5.
`title`	Character string: plot title.

Value

A ggplot2 object.

Examples

```
set.seed(42)
df <- data.frame(ses = rnorm(300), bmi = rnorm(300, 25, 5),
                 math = rnorm(300, 500, 100), opp = rnorm(300))
df <- classify_regime(df, ses_var = "ses", opportunity_var = "opp")
plot_regimes(df, ses_var = "ses", health_var = "bmi", outcome_var = "math")
```

Index