Initial Codes for SES Quotes

Three dimensions:

1. How SES theoretically defined
   * Categories:
     + “None given” = didn’t provide definition
     + Remainder will be inferred from the data (no categories to start with)
2. How SES operationalized
   * Categories:
     + **Note**: These refer to how *SES* was operationalized, and NOT to how indicators of SES were (e.g., using PCA or a sum score for family assets does *not* count as dimension reduction, but using PCA or a sum score for family assets and income does count as dimension reduction)
     + “None given” = didn’t say
     + “Did not study SES” = particular study didn’t study SES
     + “Single indicator” = used only one indicator
     + “Multiple single indicators” = used more than one single indicator
     + “Dimension reduction” = used process to reduce multiple single indicators into fewer variables (e.g., from K=3 to K=2 or 1, where K is the number of variables referred to as SES)
       - + “Formative” = used a dimension reduction technique wherein the indicators define the latent variable (e.g., PCA, WLS)
         * “PCA” = principal components analysis
         * “PLS” = partial least squares
       - + “Reflective: = used a dimension reduction technique wherein the latent variable defines (gives rise to) the indicators (e.g., CFA, EFA)
         * “CFA” = confirmatory factor analysis
         * “EFA” = exploratory factor analysis
       - + “Other composite” = used a modeling strategy that doesn’t fall neatly into either of the above (e.g., equal weights could be either reflective or formative, Hollingshead, Duncan SEI)
         * “Standardized mean” = all indicators standardized and average of the standardized indicators used as measure
         * “Quintiles” = split continuous composite into five groups
         * “Trichotomized” = split continuous composite into three groups
         * “Normalized mean” = all indicators changed to same scale and average of normalized indicators used as measure
         * “Logarithm” = logarithm of indicators used
         * “Sum” = sum of indicators
         * “Mean” = mean of indicators
         * “Standardized sum” = all indicators standardized and sum of the standardized indicators used as measure
         * “Dichotomized” = split continuous composite into two groups
     + Remainder will be inferred from the data (no categories to start with)
     + Modifications to indicators:
       - “Dichotomized” = made an otherwise non-binary indicator binary
     + “Mixed” = mix of single indicator approaches and dimension reduction approaches

Protocol for writing codes:

* No abbreviations or conjunctions other than SES (i.e., write “did not”, not “don’t”)
* Capitalize the first letter of each code
* Use spaces to separate words
* Elaborating operationalizations:
  + In addition to recording the type of operationalization (e.g., “Multiple single indicators”), need to record content of operationalization (e.g., income, education)
  + Separate type from content using an underscore (i.e., “Single indicator\_income”)
  + If writing a list of content, separate elements of list using hyphens (i.e., “Multiple single indicators\_income-education-occupational prestige”)
    - Within an element of a list, separate words using spaces (e.g., “occupational prestige”, not “occupationalPrestige” or something else)
  + Modifications are listed in brackets ([]) after the indicator to which they apply and separated by a space from the indicator (e.g., “income [dichotomized]”)
* Elaborating “Dimension reduction”:
  + Place type of dimension reduction in brackets ([]) next to after “Dimension reduction” (e.g., “Dimension reduction [Reflective]\_income-education”)
  + Specific method of dimension reduction should follow general type, separated by a period (e.g., “Dimension reduction [Reflective.CFA]\_income-education-employment”)
  + Elaborating “Other composite”:
    - * For non-standard approaches (e.g., not using name-brand indices like Hollingshead or Duncan), indicate the modeling strategy used after “Other composite”, separated by a period (e.g., “Dimension reduction [Other composite.mean]\_income-education”)
      * Each period represents a separate elaboration approach
* Elaborating “Mixed”:
  + If dimension reduction involves multiple smaller indicators, give a label to the latent variable (e.g., “subjective SES composite”) and indicate indicators in brackets, spearated by hyphens (e.g., “subjective SES composite [subjective MacArthur ladder-perceived financial difficulties]”)
  + If method for combination given (e.g., mean, sum, etc.), write it after the indicators, separated by a period (e.g., “objective SES composite [income-education.Standardized mean]

1. Reason for operationalization
   * Categories:
     + **Note:** These refer to *both* how indicators were chosen and how a modeling procedure was chosen.
     + **Note:** These refer to how *SES* was operationalized, and NOT to how indicators of SES were (e.g., using PCA or a sum score for family assets because it is recommended by other researchers does *not* count as “Prior work”, but using PCA or a sum score for family assets and income to measure SES because other researchers recommend it does count as “Prior work”)
     + “None given” = didn’t say
     + “Did not study SES” = particular study didn’t study SES
     + Remainder will be inferred from the data (no categories to start with)
       - But some ideas from Kachmar et al. (2019):
         * “Evidence based” = using prior knowledge of relation between particular operationalization and particular outcome or related outcomes

e.g., “A major review has suggested its [two factor Hollingshead] consistency in pre- dicting psychotherapy outcomes”

e.g., “Meta- analyses show that the same factors traditionally appear to predict poor treatment outcome and high attrition from trials (Lundahl et al., 2006; Reyno & McGrath, 2006). Following this literature, we examined two sets of potential moderators of outcome, assessed at age 2. Family risk factors included mother a single parent, mother a teen at the birth of her first child, low maternal educational level, and maternal alcohol or drug problem.”

e.g., “In addition, this SES index is more linked to obesity than family income”

* + - * + “Prior work” = prior research used same operationalization, but no stated relevance of operationalization to current study

e.g., "Hollingshead’s operationalization of socioeconomic status using four factors is the most widely used and validated measure of socioeconomic status in psychological research”

* + - * + “Undetailed reason” = stated a rationale without justifying details or criterion [I don’t know what this means]
    - From data (indicators):
      * “Parallel forms” = Operationalization of SES correlates with other possible operationalizations of SES
        + e.g., “The measure [Hollingshead’s index] correlates highly with other indices of social status”
      * “Content validity” = Operationalization of SES intended to capture multiple different aspects of SES
        + e.g., “In this study, using all three indicators serves the purpose of attempting to address discrepancies in the literature on SES and treatment outcome, because most studies have used only one of these three indicators, with varying results"
      * “Data availability” = Operationalization chosen, or not chosen, because of high, or low, availability for many participants
        + e.g., “An occupational measure of SES [at 22 years] was not used at that age because only about half of the subjects had completed their vocational studies by that age and had moved into working life”
      * “Population appropriate” = Operationalization of SES particularly good for proxying SES within a particular population
        + e.g., “Our focus on family structural assets as a culturally ap-propriate measure of disadvantage may be a better proxy of socioeconomic risks that undermine maternal opportunities to de-velop verbal skills than of acute stressors known to affect maternalfluid intelligence and executive function skills.”
        + e.g., “In the Finnish schooling system, completing high school is the most prominent divisor as regards later educational attainment, and a prerequisite for practically all university studies.”
        + e.g., “This is also considered best for use with members of racial/ethnic minority groups, who do not receive the same financial gains for equivalent years of education as do Whites”
      * “Dataset provided” = Operationalization of SES was predetermined and given in the dataset used in the study
        + e.g., “The ECLS-K data set provides a continuous SES variable,which was utilized for our analyses.”
    - From data (modeling procedures):
      * “Sample size” = Operationalization [modeling procedure] of SES particularly appropriate for the sample size of the study
        + e.g., “Partial least squares modeling isideal for examining all of the possible associations between thevariables of interest as it can be used with small data sets and isrecommended for exploratory purposes, as it does not have thestrict measurement and distributional requirements of other mod-eling procedure”
      * “Non-interchangeable” = Operationalization [modeling procedure] of SES designed to address non-equivalence of indicators of SES
        + e.g., “Though related, indicators of SES are notinterchangeable, as each may affect disease risk in different ways(Braveman et al., 2005), and therefore may have different inter-vention implications.”
      * “Reduce measurement error” = Operationalization of SES intended to reduce measurement error that impacts alternative operationalizations [modeling procedures]
        + e.g., “SEM allows the simultaneous test of a hypothesized structure of relationships among variables of interest as well an inclusive test of the measurement model of the latent or unobserved variables (i.e., SES, mental illness). When this is not done, measurement error in the composite variables is ignored.”

Protocol for writing codes:

* No abbreviations or conjunctions other than SES (i.e., write “did not”, not “don’t”)
* Capitalize the first letter of each code
* Use spaces to separate words
* Structure rules:
  + Each code contains two halves, separated by an underscore (“\_”):
    - Half 1 = Reason for indicators
    - Half 2 = Reason for modeling procedure
  + Within each half, different reasons are separated by a hyphen (“-”)
    - e.g., “Prior work-Parallel forms\_None given” indicates that a paper cited prior work using the same indicators and that the indicators correlate with other operationalizations of SES but provided no reason for the modeling procedure used
    - e.g., “None given\_None given” indicates that a paper gave no reasons for its choice of indicators or modeling procedure