Coding Guidelines for REDLOC Curated Forms and Instruments
Version: 2023-03-27

After instruments have been suggested by researchers, reviewed by REDLOC librarians, and recommended for inclusion by REDLOC members, the forms must then be translated into REDCap format. At this time, REDLOC members are involved in this process on a rotating basis. In order to assist in form translation and encourage consistency throughout the instruments, REDLOC has developed the following guidelines for library forms developers. This document is intended to be used for the development of new instruments. If the instrument has already been developed, this document will apply only in circumstances where the coder/committee has deemed it necessary to make changes to the submitted instrument. The scoring or coding instructions provided by the author take precedence over this guidance.

Instrument Name
- If possible, use the full instrument name, abbreviated instrument name and version number. If the name is too long, be as descriptive as possible while still staying within the size limitation. For example, the full instrument name is *Infant Behavior Questionnaire Revised - (IBQ-R).* But *Pelvic Organ Prolapse/Urinary Incontinence Sexual Function Questionnaire (PISQ)* is *PISQ.*

Variable Names
- Variable/field names should begin with a brief representation of the instrument. For example, the field label "patient" in the "Agitated Behavior Scale" is coded as variable name "abs_patient".

- Variable/field names for numbered questions should be coded with instrument abbreviation and the question number, such as "abs_1".

- For not-numbered questionnaires, variables/fields which are coded 0 = No and 1 = Yes should include a verb as part of the variable name. For example, a question on the Nacho Craving Index that asks, "Are you currently craving nachos?" which is coded 0 = No and 1 = Yes, would be named "nci_is_craving". Similarly, questions which ask "Have you ever had a cancer diagnosis" would be coded with "had_cancer" as part of the name or "Are you currently pregnant" would include "is_pregnant" as part of the variable name. If sex assigned at birth is coded 0 = Male, 1 = Female, the variable would include "is_female" as part of the name.

- Variable/field names for instruments that have been translated will contain a two-letter abbreviation between the instrument indicator and the actual variable name, such as "abs_es_1" or "abs_es_patient". A listing of the language abbreviations can be found here: https://www.sitepoint.com/iso-2-letter-language-codes/. Instruments in English should **not** include the language code.

Scored Version of an Instrument

- If an instrument/questionnaire can be "scored" in REDCap, please produce a scored and not-scored version. For the not-scored version, use the instrument name and variable name guidance described above. For the scored version, add "Scored" to the name of the instrument and add _sc to the variable names. In other words, variables should be the same for both instruments except for adding "_sc" to differentiate between scored and not-scored variables (i.e., instrument_1 and instrument_1_sc).

- When submitting an instrument to the library, if you are not submitting both a scored and not-scored version, please let us know why a scored version was not created. Typically, this will be because the instrument does not have scoring instructions or the library committee pre-determined that the scoring is too difficult/confusing to code properly in REDCap.

Coding Values for Categorical Fields

- If variable codes are provided in the original instrument documentation, use as described. Example: BRFSS 2009 Section 20: HIV-AIDS Question 20.1 responses –

  - 1, Yes

  - 2, No

  - 7, Don't know / Not sure

  - 9, Refused

- If variable codes are not provided, here are some suggestions:

  - Binary responses should be coded 0/1.

    - No=0, Yes=1

    - False=0, True=1

  - Response sets with more than two meaningful options (i.e., three or more options beyond don't know/refused/etc.) should begin with 1. For example: 1, small |2, medium |3, large.

  - Responses that represent an unknown type of code value (e.g., don't know, refused to answer or missing) should be coded with a number that is obviously out of the typical range for expected responses. For example:

    - Has the patient tested his/her blood sugar levels every day this week?

      ☐ Yes
      ☐ No

☐ Don't know
☐ Refused
☐ Missing

might be coded as:

1, Yes
0, No
7777, Don't know
8888, Refused
9999, Missing

Reverse Coding

- If questions have reverse scoring, code the selections according to the value needed for scoring unless otherwise specified. Here is an example, where #4 is coded normally and #5 is reversed:

  - 4. Made me need someone to reposition me.

    - 0, Never true for me

    - 1, Rarely true for me

    - 2, Sometimes true for me

    - 3, Often true for me

    - 4, Very often true for me

  - 5. Helped me keep my muscles exercised.

    - 4, Never true for me

    - 3, Rarely true for me

    - 2, Sometimes true for me

    - 1, Often true for me

    - 0, Very often true for me

- A note *must* be included in the instrument header that says the equivalent of "Reverse coding has already been applied for scoring purposes."

- Mixing normal and reverse coding in a matrix is impossible.

Text Fields

- If questions have text field responses, use field validation whenever possible (i.e., date, phone, email, integer, number, or other available field types). For example, in a study that runs until

01/18/2025: (field label) *Today's Date* – (possible response) *06/24/24* (Validation? – *Date(M-D-Y)*)

- If number or integer validation is applicable, define the minimum/maximum range as much as possible. Example, for data collected in a study on children aged 6-12: (field label) *Age of child:* (possible response) *12* (Validation? *Integer* /Minimum: *6* – Maximum: *12*)

Form Layout
- Include instructions in the header when available. Example for PedMIDAS instrument: *Headache Disability: The following questions try to assess how much the headaches are affecting day-to-day activity. Your answers should be based on the last three months. There are no "right" or "wrong" answers so please put down your best guess.*

- Structure the REDCap form to resemble the original instrument as closely as possible. For the response area, take advantage of *Custom Alignment* options such as Left/Horizontal (LH) (applied below); Left/Vertical (LV); Right Vertical (RV); Right/Horizontal (RH).

Field Labels
- Typos and grammatical errors may be corrected, as long as the intent is not affected. Example: "Do you worry a lot about **you** health?" is okay to change to "Do you worry a lot about **your** health?"

- Slight wording changes may be implemented if needed because of the document format change from paper to electronic. Example: "**Circle** the correct response" becomes "**Choose** the correct response".

- If the paper version of an instrument includes question numbers and the REDCap instrument **does not** use branching logic, include the question number in the field label.

- If the paper version of a questionnaire includes question numbers and they are included in a REDCap instrument that **does** include branching logic, the numbers can be confusing for respondents because question numbers will be missing. So, do not include the question number

in the field label on instruments with branching logic unless the respondent needs the question numbers to correctly fill out the form or if the question numbers are needed for scoring/evaluation. For example, include question numbers if a questionnaire includes questions like "Above, on questions 14 and 17 you reported using drugs to help you stay awake, were those drugs prescribed to you?".

- Simple HTML may be implemented in field labels to make it more closely resemble the original instrument. It may also be used to clarify or highlight information, as long as the intent is not affected. Examples: <u>text</u> for underline, <i>text</i> for italicize, <div><span style='color:red;'>text</div> to color red and highlight the interviewer instructions.

Field Type

- For data entry forms, dropdown fields are preferred to radio button fields, but use your judgment as to what is most suitable.

- For surveys, use radio buttons for short lists and dropdowns for longer lists. Again, use your judgment as to what is most appropriate for the form.

- Use the Descriptive Text field type to attach scoring information or stats package coding, if available. See **Interpersonal Reactivity Index** instrument.

- Mark *Identifiers*. To be HIPAA-compliant, there are at least 18 pieces of information that must be marked as *Identifiers* in a REDCap data dictionary.

    1. Names
    2. Fax numbers
    3. Phone numbers
    4. E-mail addresses
    5. Account numbers
    6. Social Security numbers
    7. Medical Record numbers
    8. Health Plan numbers
    9. Certificate/license numbers
    10. URLs
    11. IP addresses
    12. Vehicle identifiers
    13. Device IDs
    14. Biometric IDs
    15. Full face/identifying photos
    16. Other unique identifying numbers, characteristics, or codes
    17. Postal addresses (geographic subdivisions smaller than state)
    18. Date precision beyond year

- Mark initials as *Identifiers*. According to the National Institutes of Health, (https://privacyruleandresearch.nih.gov/research_repositories.asp) because an individual's name is an identifier and initials are derived from the individual's name, initials are considered identifiers under the Privacy Rule. Thus, in addition to full names, fields that contain initials must be marked as identifiers.

Branching Logic

- Use branching logic unless the document layout dictates otherwise. Example: BRFSS has questions that would branch based on answers from different forms, but since the library provides each form individually, the branching logic would not work/be applicable.