International Journal of Community Medicine and Public Health
Dharaiya DK. Int J Community Med Public Health. 2026 Sep;13(9):5405-5411
https://www.ijcmph.com
pISSN 2394-6032 | eISSN 2394-6040

DOI: https://dx.doi.org/10.18203/2394-6040.ijcmph20263233

Review Article

Designing data collection tools for psychiatric research: a practical guide for postgraduate psychiatry trainees

Darshankumar K. Dharaiya*

Department of psychiatry, Gujarat Institute of Mental Health, Hospital for Mental Health, Ahmedabad, Gujarat, India

Received: 21 May 2026
Revised: 20 August 2026
Accepted: 24 August 2026

***Correspondence:**
Dr. Darshankumar K Dharaiya,
E-mail: dharaiya.darshan@gmail.com

ABSTRACT

Data collection tools are used to document information gathered from participants or clinical records. Their structure and content vary depending on the study type, instruments used, and variables of interest. A poorly designed tool introduces systematic errors that no statistical method can correct. Since data collection is an irreversible step, and the quality of all downstream analyses depends on what is collected and how it is documented, careful attention to tool design is essential. This guide provides a structured, practically oriented framework for designing paper-based data collection tools in psychiatric research, covering study type, format, language, ordering, coding, and data entry compatibility.

Keywords: Survey, Data collection, Data collection tool, Validity, Reliability, Research ethics

INTRODUCTION

Most psychiatric research rests on information collected from participants. Common variables include physical parameters, cognitive task results, and a wide range of psychometric instruments measuring constructs such as emotion, personality, and attachment. Different study designs and research questions require tools capable of capturing appropriate data for that specific design and question.

Data collection is an irreversible step-errors in tool design cannot be corrected once recruitment has begun without compromising the dataset. Adequate time must therefore be devoted to designing a tool suited to the specific study design before data collection begins. The data collection tool should be finalised before ethics submission, as any substantive modification after approval requires a formal amendment application. Common consequences of poor design are summarised in Table 1.

MATCHING THE TOOL TO THE STUDY DESIGN

The structure of a data collection tool should reflect the demands of the study design. Table 2 outlines key considerations by study type. In general, cross sectional study design is the most commonly used in psychiatric and-psychology-research.

Table 1: Consequences of poorly designed data collection tool

Risk Factors	Consequence
Poor visual appearance	Difficulty in completing the tool; reduced compliance
Poor item sequencing	Response bias, participant dropout
Inconsistent coding decisions	Messy, unanalysable data
Language mixing within a tool	Invalid cross-participant comparisons
Altered scale item order	Invalidates psychometric properties
Retrospective data entry design	Systematic entry errors, lost data
No pilot testing	Ambiguities discovered only after recruitment begins
Excessive length or number of scales	Data dredging, poor completion rates, inconsistent or random responding

*This guide does not cover instrument development, translation methodology, or psychometric validation, or designing online data collection tools, each of which constitutes a study in its own right. This is an opinion and practical suggestion, based on personal experiences, which addresses the practical decisions involved in assembling a tool from existing validated instruments and locally designed components.

Table 2: Matching the tool to the study design

Study Types	Key Tool Design Considerations
Cross-sectional	Single administration; prioritise brevity and logical sequencing; manage participant burden carefully.
Longitudinal	Label baseline and repeat sections clearly; maintain strict consistency across time points
RCT	Separate outcome assessment from allocation information; design arm-specific and time-point-specific versions
Qualitative	Open-ended prompts; generous response space required; clear visual separation between items needed
Scale translation/validation	Requires expert rating grids, relevance scales, and commentary fields - governed by separate methodological standards.[1]
Content/face validation	Requires separate structured feedback sections; not a standard data collection tool.[2,3]
Performance-based studies	The data entry sheet may differ from the actual task materials. For example, a participant performs the Stroop test while the researcher records results on a separate scoring sheet, which becomes the functional data collection tool
Systematic Review and meta-analysis- and other review articles	Extraction of data of interest from published resource-journal articles and other grey literature. Specific software is available.

*RCT-Randomised controlled trail

ACCURACY WHEN COPYING SCALE ITEMS

Research tools are frequently assembled by copying (or manually typing) scale items from published papers, scanned originals, translated documents, or prior studies. This process may introduce errors: items may be inadvertently truncated, words may be dropped or altered, response anchors may shift, and item numbers may be displaced - all without the researcher noticing, because the resulting text might still read as plausible. An item that has lost a negation, gained a word, or had its anchor labels swapped is no longer the validated item; it is a modified item with unknown psychometric properties. To guard against this, each scale in the assembled tool should be compared word-for-word against the validated source document before finalisation, not merely read through for general sense. Where the original instrument exists in a published or official format, that version should be the reference, not a version copied from another study's appendix, which may itself contain errors.

TOOL LENGTH AND PARTICIPANT BURDEN

There is no universally accepted upper limit on the number of scales or items a research tool may contain, but the guiding principle is that every scale must be independently justified by the research question. Including scales because they are commonly used, freely available, or tangentially interesting - sometimes called data dredging - inflates participant burden without

proportionate scientific return and increases the risk of incomplete responses, random responding, and early dropout. In a narrative review addressing the optimum length of a questionnaire, 25–30 questions and a completion time of 30 minutes were suggested.[4] A self-administered tool taking longer than 45 to 60 minutes is likely to cause significant fatigue in most outpatient and community samples; inpatient or acutely unwell populations may tolerate considerably less. Estimated completion time should be calculated during pilot testing rather than assumed.[5]

MODE OF ADMINISTRATION (INTERVIEWER-ADMINISTERED VERSUS SELF-ADMINISTERED FORMATS)

The choice of administration format is a design decision that should be made deliberately, documented in the protocol, and reflected consistently in the tool layout (Table 3).

Table 3: Mode of administration.

Considerations	Guidance
Self-administered	Appropriate when participants have adequate literacy, cognitive capacity, and no significant language barrier; reduces researcher influence; feasible at large scale, online data collection
Interviewer-administered	Preferred when participants have limited literacy, cognitive impairment, or significant psychiatric symptomatology; required when the scale demands clinical judgement to score (e.g. HDRS, PANSS, CGI)
Format switching	Administering a scale in a format other than that for which it was validated may raise question of validation
Mixed tools	When both formats appear in the same tool, clear visual and instructional demarcation between sections is essential

*HDRS – Hamilton Depression Rating Scale, PANSS – Positive and Negative Syndrome Scale, CGI – Clinical Global Impression

Table 4: Recommended section sequence for paper-based psychiatric research tools

Positions	Section	Notes
1	**Information sheet**	Plain language; covers study purpose, procedures, confidentiality, and contact information
2	**Consent form**	Same language as information sheet; witness signature if participant has low literacy
3	**Sociodemographic data**	Minimum and analytically relevant variables only
4	**Scales - easy or neutral first**	Non-threatening, cognitive, or general constructs
5	**Scales - sensitive or emotionally loaded last**	Trauma, adversity, intrusive cognitions (e.g., ACE-IQ, MCQ-30)
6	**Open-ended or qualitative items**	Should be kept at last; requires greatest engagement and willingness to disclose

*ACE-IQ – Adverse Childhood Experiences International Questionnaire, MCQ-30 – Metacognitions Questionnaire-30

FORMATION OF ADMINISTRATIONS (ONLINE VS. PAPER-BASED)

This guide focuses on paper-based tools. Online data collection requires separate consideration, including issues of format validation and psychometric equivalence. Reader is directed to other source for designing online data collection tool.[6] For paper-based data collection tool, participant may require almost no resources, and does not require digital literacy. For in-patient or out-patient, with limited digital literacy, and when participant might need support in competing the data collection tool, paper-based tool might be preferred. Manual data entry from paper-based form may introduce errors. For large community-based survey, when digital literacy may not be an issue, online data collection tools are widely used. Limited skip logic may be available based on online data collection platform, which can be used by indicating arrows or instruction in paper-based forms.

LANGUAGE

In general, one participant should receive the tool in one language only. Mixing languages within a single participant's tool should not be done unless explicitly justified in the methodology - for example, in a language validation study. Two separate language versions of the tool should be prepared rather than placing two languages side by side within a single document. It is important to note that different language versions are considered data-equivalent only when formal cross-cultural validation has been demonstrated for each version.[7]

PHYSICAL FORMAT

The physical appearance of the tool affects participant compliance and data quality. Generally, A4 portrait paper may be used commonly, but for complex grid, visual analogue scale, branched logic, A3, or booklet format too can be adapted. Clear and easy to read font of appropriate size should be used; decorative fonts should be avoided and professional appearance should be maintained). To improve readability, adequate white space should be kept in between the questions and scales. For qualitative research adequate space for documentation of responses should be provided.

STRUCTURAL ORGANISATION OF THE TOOL

The following sequence is generally recommended for paper-based psychiatric research tools. (Table 4) Modifications may be necessary depending on the study design, but the information sheet and informed consent form should almost always appear first, given their ethical and legal requirements. The general ordering principle - easy before difficult, neutral before emotionally activating - should be followed. Within the sociodemographic section, sensitive items such as religion, caste, and financial details can be presented last.

SOCIODEMOGRAPHIC DATA

Sociodemographic data should be limited to variables that are analytically relevant to the study. Blanket data collection - gathering variables simply because they are routinely collected - should be avoided. Where possible, data should be collected in continuous or ordinal format rather than categorical, as continuous data permit a wider range of analytic approaches. For example, age in years, years of formal education, total duration of illness, and duration of untreated illness in months are all preferable to their categorical equivalents.[8,9] For items that may be ambiguous, brief explanatory notes should be provided within the form - for instance, clarifying the definition of occupation as used in the Kuppuswamy scale, or explaining how residential status should be interpreted.

Table 5: Scale ordering principle

Principless	Example Applications
Easy before difficult	Demographic-adjacent scales before complex self-reflection scales
General before specific	Global wellbeing before disorder-specific symptom scales
Emotionally neutral before emotionally loaded	PHQ-9 before ACE-IQ
Cognitively engaging before emotionally activating	Attention or memory scales before trauma inventories

*ACE-IQ- Adverse Childhood Experiences International Questionnaire (WHO).

Table 6: Critical rules for use of validated scales

Rules	Rationales
Internal item order of scale should not be altered	Item sequence is part of the validated instrument; alteration changes psychometric properties
Response anchors or format should not be altered	Changing from Likert to VAS, or modifying anchor labels, invalidates the instrument
Only a part of a scale should not be used without justification	Partial scale use has direct implications for reliability and validity; requires explicit rationale.
Subscales from different instruments should not be combined without justification	Creates an unvalidated composite measure.
Mix language versions of the same scale across participants should not be applied	Different language versions are data-equivalent only if cross-cultural validation has been established.
Selection and use of each scale should be justified independently	Using multiple scales for the same construct - for example, PHQ-9 alongside HDRS and MADRS - requires explicit scientific rationale.
Reverse scoring should be applied as specified in the scoring manual	Several instruments include negatively worded items that must be reverse-scored before summing; omitting this step may inflates or deflates total scores and invalidates all downstream analyses.

*VAS – Visual Analogue Scale, PHQ-9 – Patient Health Questionnaire-9, HDRS – Hamilton Depression Rating Scale, MADRS – Montgomery-Åsberg Depression Rating Scale

SCALES AND STRUCTURED INSTRUMENTS

Scale sequencing within the tool should follow the same principles applied to the overall tool structure. (Table 5).

CRITICAL RULES FOR VALIDATED SCALES

Validated scales must be used as developed - altering item order, response format, language, or mode of administration without revalidation compromises psychometric properties. Trainees should also use scales validated for the specific language, population, and clinical condition relevant to their study, rather than defaulting to scales validated in a different population simply because they share the same language (Table 6).[7]

INSTRUCTIONS

Instructions should be provided at two levels - global and section-specific - as outlined in Table 7. Instructions must be visually distinct from item text so that participants do not confuse them with response items.

CODING AND DATA ENTRY COMPATIBILITY

The data entry spreadsheet and coding scheme should be designed simultaneously with the data collection tool, before recruitment begins. Attempting to construct a coding system retrospectively reliably produces confusion, inconsistency, and avoidable errors.[9] A consistent missing data convention must be established before entry begins and applied across the entire dataset.[10] (Table 8)

SPECIAL CONSIDERATIONS IN PSYCHIATRIC POPULATIONS

Data collection in psychiatric populations requires particular care. Where sensitive or distressing content is included, contact details for available helplines or support services should be provided within the tool as part of standard ethical requirements (Table 9) [11-13]

PILOT TESTING

Pilot testing the assembled tool is recommended before recruitment begins. The pilot should involve both knowledgeable colleagues and, wherever possible, participants from the target population - colleagues can identify structural and instructional ambiguities, while target population members may reveal problems of comprehension, language, emotional reaction, and completion time that colleagues may not anticipate. A sample of 10 to 30 is commonly cited, though fewer may suffice if the tool is straightforward. All revisions made following the pilot should be documented and reported in the methods section as part of the methodological record (Table 10).[5]

Table 7: Instructions within the data collection tools

Levels	Contents	Placements
Global	Study purpose, estimated time, general response format, (what to do if stuck-for online data collection tool)	Before sociodemographic section. Can be divided between participant information sheet and rest of the data collection tool
Section-specific	Format change alerts, rating guidance, interviewer vs. self-completion distinction	Immediately before each scale
Visual distinction	Bold, boxed, or larger font	Clearly separated from item text

Table 8: Data entry design principles

Principles	Details
Numeric coding for all responses	Avoid free text in the entry sheet wherever possible
Preserve original scale scoring range	If a scale uses 0–3, do not recode to 1-4 in the entry sheet; this creates confusion for all involved, including collaborating statisticians
Linear item sequence	Items entered together should appear in an unambiguous linear sequence in the tool
Branching logic marking	Indicate skip rules using arrows, brackets, or shaded boxes in the tool; document skip logic explicitly in the entry protocol
Data dictionary	Variable names, labels, codes, and missing value conventions should be completed before data entry begins
Response coding conventions	Use 98 = not applicable, 99 = not answered/missing as a recommended convention; note that some instruments use predefined codes that should not be overridden. For binary variables, maintain consistent 0/1 coding throughout where instrument scoring permits. In R, user-defined missing values must be explicitly converted to NA before analysis; SPSS handles these natively via the MISSING VALUES declaration

Table 9: Special considerations in psychiatric populations

Challenges	Considerations	Practical Responses
Cognitive impairment	Scales requiring sustained attention, abstract reasoning, or intact insight - such as metacognition or emotion regulation measures - may produce unreliable responses during acute illness	Consider clinician-administered formats; review appropriateness of scale for current clinical state
Emotionally distressing	Items addressing trauma, suicidality, or psychotic experiences can cause significant distress	Place these scales last; ensure a debrief and referral pathway is available
Limited literacy	Common in inpatient and rural outpatient populations	Use interviewer-administered format; adapt the tool layout accordingly
Language preference	Mismatch between tool language and participant's primary language invalidates data	Use strict single-language tools; verify language preference before administration

Table 10: Goals of pilot testing

Goal	Details
Item wording	Identify ambiguous or poorly worded items
Instructions	Identify confusing or missing instructions
Response format	Identify problematic or unclear response formats
Completion time	Measure rather than estimate; researchers routinely underestimate
Participant burden	Identify unanticipated fatigue or distress

THE DATA DICTIONARY

A data dictionary is a reference document that defines every variable in the dataset before data entry begins. At minimum it should record, for each variable: the variable name as it appears in the spreadsheet, a plain-language label describing what it represents, the scale or section it belongs to, the permissible range of values, the meaning of each numeric code, and the codes used for missing or non-applicable responses. It should also record any scoring or recoding rules applied after entry. The data dictionary allows a collaborating statistician, a co-investigator, or the researcher themselves at later date to understand the dataset without ambiguity. It should be maintained by the principal investigator, updated whenever a variable definition or coding decision changes, and stored alongside the dataset. A well-maintained data dictionary also serves as an audit trail for methodological decisions made during the study.

DATA SECURITY AND CONFIDENTIALITY

Participant data must be protected from the pint of collection onwards, which is standard ethical requirement. 11-13 participants should be identified using id numbers only, with a separate secured linkage log maintained apart from the data. For anonymous studies, no identifying information should appear on the tool itself, and this should be stated explicitly in the information sheet. Physical copies should be stored in a locked cabinet restricted to the research team, while digital data should be password-protected and access restricted to those directly involved in the study.

CONCLUSION

The design of a data collection tool is a methodological decision with direct consequences for data quality, participant experience, analytic validity, and the scientific contribution of the study. Postgraduate trainees who invest time in careful tool design - attending to sequence, format, coding, and pilot testing - are far less likely to encounter incomplete datasets, ambiguous variables, or unanalysable responses after recruitment has closed. The principles outlined in this guide apply across the range of study designs encountered in psychiatric research and provide a foundation for developing this skill systematically across a research career.

Funding: No funding sources
Conflict of interest: None declared
Ethical approval: Not required

REFERENCES

1. Ohrbach R. Guidelines for establishing cultural equivalency of instruments. Available from: https://www.researchgate.net/publication/265229959_Guidelines_for_Establishing_Cultural_Equivalency_of_Instruments/citations. Accessed on 20 March 2026.

2. Desai S, Patel N. ABC of face validity for questionnaire. Int J Pharm Sci Rev Res. 2020;65(1):164-8.
3. Yusoff MSB. ABC of content validation and content validity index calculation. Education in Medicine Journal. 2019;11(2):49-54.
4. Sharma H. How short or long should be a questionnaire for any research? Researchers dilemma in deciding the appropriate questionnaire length. Saudi J Anaesth. 2022;16(1):65-68.
5. Dharaiya D. A practical primer on pilot study for postgraduate psychiatry students. Ann Indian Psychiatry. 2026.
6. Dharaiya D. Online data collection in psychiatric research: a practical guide to platform selection, form design, and ethical conduct for postgraduate psychiatry students. Ind J Priv Psychiatry. 2026.
7. Andrade C. Rating patients in different languages: reliability and validity. Indian J Psychol Med. 2024;46(3):268-9.
8. Andrade C. A student's guide to the classification and operationalization of variables in the conceptualization and design of a clinical study: Part 2. Indian J Psychol Med. 2021;43(3):265-8.
9. Andrade C. The young researcher's guide to creating data mastersheets. Indian J Psychol Med. 2022;44(6):618-9.
10. Williams R. Coding missing data. University of Notre Dame Department of Sociology. Available from: https://academicweb.nd.edu/~rwilliam/stats3/MDCoding.pdf. Accessed on 20 March 2026.
11. Ethical principles and practices for research involving human participants with mental illness. Psychiatr Serv. 2006;57(4):552-7.
12. Indian Council of Medical Research. National Ethical Guidelines for Biomedical and Health Research Involving Human Participants [Internet]. New Delhi: ICMR; 2017 [cited 2026 April 21]. Available from: https://ethics.ncdirindia.org/asset/pdf/ICMR_National_Ethical_Guidelines.pdf. Accessed on 20 March 2026.
13. Grover S, Sarkar S, Sahoo S. Ethics in psychiatric research. Indian J Psychiatry. 2022;64(Suppl 1):S68-80.

Cite this article as: Dharaiya DK. Designing data collection tools for psychiatric research: a practical guide for postgraduate psychiatry trainees. Int J Community Med Public Health 2026;13:5405-11.