The Arizona Border Study

An Extension of the
Arizona National Human Exposure Assessment Survey (NHEXAS)Study
Sponsored by the Environmental Health Workgroup of the Border XXI Program

Quality Systems and Implementation Plan for Human Exposure Assessment

The University of Arizona
Tucson, Arizona 85721

Cooperative Agreement CR 824719

Standard Operating Procedure **SOP-UA-T-5.0**

Title: General Training of Student Data Assistants

Source: The University of Arizona

U.S. Environmental Protection Agency
Office of Research and Development
Human Exposure & Atmospheric Sciences Division
Exposure & Dose Research Branch

ESG 7.14.97

~~Environmental Protection Agency~~ ESG 7.14.97
~~Contract Number: CR821560~~ ESG 7.14.97

~~NHEXAS Arizona Project~~

Title: **STUDENT DATA ASSISTANT TRAINING PLAN**

Document No. UA-T-5.0

APPROVALS

☒ Full SOP ☐ Working SOP #pages 7

On Site Principal Investigator:

Issue Date: 5/16/94

Project QA Director:

Revision No. 5.0

Independent Reviewer:

Revision No:
Revision Date:
Revision Made:

On Site PI:

Project QA Director:

Independent Reviewer:

Revision No:
Revision Date:
Revision Made:

On Site PI:

Project QA Director:

Independent Reviewer:

Distributed To:	Revision No.					
	1	2	3	4	5	6

Form TP-1

GENERAL TRAINING OF STUDENT DATA ASSISTANTS

1.0 Purpose and Applicability

This SOP describes the training sequence for incoming Student Data Assistants (Students). The procedure is designed to provide them with an overview of the study in terms of study goals, structure and project data needs. This overview familiarizes the students with project components and personnel at all levels and the training enables them to produce consistently cleaned data.

This SOP applies to Student Data Assistants hired for the NHEXAS, ~~Arizona~~ Border and Other Health & Environment Projects.

ESG 1.14.97

2.0 Definitions

2.1 BACKUP = (v.) The process of creating a duplicate of a file, directory, or drive to protect against data loss during a hardware or software failure. (n.) The duplicate copy created during this process.

2.2 BUCKET = A plastic container with a buckle top. One bucket is assigned to each household to be visited. HHID and visit number are listed on the outside of the container. The container holds all paperwork to be completed in a by field staff or household residents through a series of household visits.

2.3 DATA, ELECTRONIC = Data stored on some type of magnetic or optical medium (for example: floppy disk, hard disk, Bernoulli, tape)

2.4 DATA, VERIFIED = Electronic data that was re-entered into the same table and database into which it was originally entered, and compared against the original entered values.

2.5 DATA CLEANING = The process of locating and correcting data processing errors. They can be individual level errors in the electronic and physical data, or they can be system level errors in the data collection, packaging, coding, entry, and cleaning procedures themselves. This process is also referred to as "data validation."

2.6 DATA PROCESSING BATCH (DP BATCH): A collection of household packets or physical data forms reviewed for quality assurance and ready for data entry. Each DP batch is assigned a unique numeric or alphanumeric code that is written on all forms in the DP batch and is entered into the database corresponding to that form.

2.7 FORM, PHYSICAL = The paper or "hard copy" version of a data form. This is also referred to as a "physical data form."

2.8 HOUSEHOLD = The residence occupied by study subjects.

2.9 HOUSEHOLD IDENTIFICATION NUMBER (HHID) = A unique identification number assigned to a study household containing subjects. An associated letter (or extension) defines any divisions of single households into multiple units.

2.10 KEYPUNCH = DATA ENTRY = The place in the Arizona Respiratory Science Center where data is coded, entered and verified.

2.11 NHEXAS = **N**ational **H**uman **EX**posure **A**ssessment **S**urvey, a research project conducted in Arizona by the University of Arizona/Battelle/Illinois Institute of Technology consortium.

2.12 PACKET: A sturdy, envelope-like container that can be fully closed and is large enough to hold the physical data form(s) generated by a study household, laboratory, research site, or data processing batch. One type of packet is used for one type of physical data forms (eg., manila envelopes would be used for all lab forms processed at the HRP site). Packets are either color coded, labeled according to their contents, or both.

2.13 PACKET PREP = def (1) personnel in this part of the project prepare material for use in the field. This includes (a) pre-printing labels (b) assembling all data forms and questionnaires to be completed (c) assigning equipment for field use; def (2) location in the building where packet assembly is performed.

2.14 TEAM LEADER = Leader of a field team responsible for subject contact, scheduling field visits, collection of data, and completion of all field forms and questionnaires.

2.15 QA = QUALITY ASSURANCE = all those planned and systematic actions necessary for ensuring the validity, integrity, preservation and retrievability of the data.

2.16 QC = QUALITY CONTROL = those quality assurance actions providing a means to control and measure the characteristics of an item, process, or the establishment of requirements.

2.17 RESPONDENT = SUBJECT = A person in the study population of the NHEXAS Arizona Project.

2.18 STUDENT = Student Data Assistant.

2.19 STUDY = NHEXAS Arizona Project.

2.20 SUBJECT = RESPONDENT (see above).

2.21 TRACKING = the method of determining where a data batch is in the data system and what procedures were performed on the data.

3.0 References

None

4.0 Discussion

None

5.0 Responsibilities

5.1 Project Data Coordinator
- (a) Hires, fires, supervises, disciplines and praises the student throughout his or her employment with the study.
- (b) Schedules student training and introduces student to appropriate personnel.
- (c) Negotiates the student's work schedule (no less than 15 hrs/week).
- (d) Instructs Students on the specifics of data validation and backup.
- (e) Instructs Students on form cross checking.
- (f) Instructs Students on entry of laboratory data.
- (g) Instructs Students on QA check techniques.
- (h) Trains Students in supplemental procedures as the need arises.
- (i) Instructs Students on the locations of data supplies and resources and provides him or her with needed materials.

5.2 Project Data Manager
- (a) Provides an overview of the Local Area Network (LAN).
- (b) Provides the student with a "username," "password," instructions on changing the assigned "password," workspace on the network and a personal "Backup" Bernoulli.
- (c) Responds to hardware or software problems encountered by students working on the student computers.

5.3 Departmental Administrative Assistant completes all paperwork for hire and directs student in the steps required by personnel.

5.4 Office Manager (or delegate) shows student required office procedures:
- (a) Sign in / sign out board.
- (b) Sign in and sign out book recording hours worked.
- (c) Appropriate technique for completion of time cards.
- (d) Notebook of student work requirements and memos.

5.5 On-Site Principal Investigator or Project Manager (or delegate) will cover the study goals, the types of data collected and a project overview.

5.6 Project Field Coordinator schedules student for field visits illustrating all different field collection procedures.

5.7 Project Laboratory Supervisor schedules observation time for the Student with a technician performing each of the lab procedures within a survey.

5.8 Team Leader describes the field procedure, explains the field use of all forms, questionnaires and answers any questions regarding field procedures.

5.9 Packet Prep describes how buckets are assembled for the field.

5.10 Keypunch shows students how to code field forms used in the study and how data are entered and verified.

6.0 Materials and Reagents

6.1 Materials

(a) Computers linked to the Local Area Network as described in SOP #UA-D-1.0
(b) A packet with all study forms
(c) Black and purple pens
(d) Access to all coding and cleaning protocols
(e) A personal copy SOP #UA-D-16.0 regarding data cleaning

6.2 Reagents

None

7.0 Standard Operating Procedure

7.1 Preparations

7.1.1 Site Selection Criteria

None

7.2 Standards and Blanks

None

7.3 Procedure Description

The training described in this SOP takes 4-5 weeks to accomplish. This is the minimum, first stage training procedure. Subsequent training takes place through one-on-one instruction by the Project Data Coordinator, the Project Data Manager or other trained personnel (depending on the task).

The training sequence is as follows (steps are dated and initialed on the "Student Data Assistant Training Record" Figure 1).

7.3.1 The Student is hired and told to report to the Project Data Coordinator at the Study Office at a given time.
7.3.2 The Student is introduced to other personnel and sent to the Administrative Assistant who will complete the paperwork required for hiring (1-3 hours duration).
7.3.3 The Student returns to the Study Office and is instructed on office procedures related to the sign-in/sign-out board, sign-in and sign- out book recording hours worked, appropriate technique for completion of time cards and the notebook containing student work policies and procedures including requirements and memos (1 hour).
7.3.4 Presentation/Discussion of the overall project goals and approaches by the Principal Investigator, Project Manager or informed delegate. At this time the student is instructed to complete the paperwork described by Packet Prep (1.5 hour).
7.3.5 Packet Prep & a Team Leader discuss the assembly of data forms and their use in the field with the student. Packet prep provides the student with a complete set of forms. The Team Leader instructs

the student on completion of the forms and keeping a daily diary for one weeks. The Student is told to report to the Project Field Coordinator.

7.3.6 The Project Field Coordinator schedules the Student for field visits so the student can observe all field procedures at least once. [Field visits will take place over a 1 to 2 week period (total time 6 hours)]. The student is sent to the Project Laboratory Supervisor.

7.3.7 The Project Laboratory Supervisor schedules observation time for the Student with the technician(s) performing lab procedures. All lab procedures within a survey will be observed. [Lab observations will be interspersed with the Field visits over a 1 to 2 week period (total time 6 hours)]. The Student is sent to the Project Data Manager.

7.3.8 The Project Data Manager familiarizes the student with the LAN and computers the student is able to use. The student is assigned a username and password. The student is instructed to immediately change the password so that he/she is the only one to know it. The student is instructed not to give his/her password to any other person. The Student is sent to the Project Data Coordinator.

7.3.9 The Project Data Coordinator starts the Student on the Rbase tutorial. [6.1 (h) and 6.1 (i) will be interspersed with the Lab and Field visits over the first and second week of employment (total time 3+ hours). This concludes the first two weeks].

NOTE: The Student now has a fairly complete view of what the data is and its source.

7.3.10 The Project Data Coordinator, Project Field Coordinator and Keypunch arrange the timing of a new data batch to correspond with the Students arrival in Keypunch.

7.3.11 The Project Data Coordinator sends the Student to Keypunch for 1 to 2 weeks. The student will learn how data entry codes all field and diary forms.

7.3.12 The Student will learn how data is entered and re-entered (verified) by Keypunch. When the Student has completed a complete batch of all packets, training in Data Entry will end and the student will report to the Project Data Coordinator.

Note: By this stage the Student has completed about 4 weeks (80 hours) of general training.

7.3.13 The Project Data Coordinator assigns the Student a dummy "batch" of data for entry, verification and validation. The data contains typical problems experienced and is designed to promote typical mistakes by the student.

7.3.14 The Student begins cleaning actual project data batches (Using SOP # UA-D-16.0) under the direct supervision of the Project Data Coordinator.

7.3.15 After completing one batch of data the Student will be considered trained for routine tasks.

7.3.16 Specialized tasks will be assigned to the trained Student as they arise and with specific one-on-one training by the appropriate personnel.

7.4 Calculations

None

7.5 Special QA Checks

7.5.1 Tolerance Limits

The Project Data Coordinator will review all data cleaning steps completed by the new Student Data Assistant. Any errors will be brought to the student's attention and she or he will be asked to correct them.

7.5.2 Detection Limits

If the new student followed data cleaning and correction procedures appropriately, then all errors are detectable. If, however, he or she changed electronic data without documenting them, then detection is more difficult. In this case, errors could be detected in an electronic data QA check (see UA-D-26.0)

7.5.3 Corrective Actions

Once a new student has completed the training sequence presented in this SOP, then data he or she handles and cleans should pass electronic data QA checks. If they fail QA checks, then the problem area(s) will be identified and the student will be coached on the technique(s).

8.0 Records

8.1 Figures

Figure 1. Student Data Assistant Training Record

8.2 Training records are filed in the Project Data Coordinator's office, NHEXAS Arizona, Univ. Arizona, Tucson, Az.

Figure 1. Student Data Assistant Training Record

STUDENT DATA ASSISTANT TRAINING RECORD

Form ID: UA-T-5.0-1.0

Student's Name: _____

INSTRUCTION TYPE	STUDENT COMPLETION Initials:	STUDENT COMPLETION Date:	TRAINING PERSONNEL Initials:	TRAINING PERSONNEL Date:
Office Procedure and Orientation				
Overview Presentation				
Field Questionnaires Completed				
Field Visit(s): _____				
Lab Observatns: _____				
Coding/Keypunch: _____				
LAN Overview SOP#_____				
Subject Confid SOP#_____				
Flow & Cust(L&F) SOP#_____				
Batching (L&F) SOP#_____				
Global Coding SOP#_____				
Data Cleaning SOP#_____				
Data Correction SOP#_____				
Data QA Check SOP#_____				
R:BASE or _____ Tutorial				
Data Entry: B#___ of_____				
Data Verif: B#___ of_____				
Trained on: Duplicate Check				
Range Check				
Logic Check				
Performed: Duplicate Check				
B#___ Range Check				
of_____ Logic Check				
Other: 1)_____				
2)_____				
3)_____				
4)_____				

Comments: _____

_____ **has completed orientation and preliminary, basic training for data handling, cleaning, and validation techniques used by NHEXAS Arizona.**

_____
Project Data Coordinator **Date**