Quality Control Considerations in Naturalistic Driving Data Reduction Efforts

Julie McClafferty, Jon Hankey, Miguel Perez

Virginia Tech Transportation Institute

Data Reduction & Quality Control – What, Why, & How?

- Data Reduction:

  Translating video data into quantifiable data for statistical analysis.

- Quality Control:
  - What: Ensuring the accuracy and precision of that translation
  - Why: People make errors, and no two people are alike.
  - How:
    - Create new data: Know the risks, have a plan
    - Use existing data : Understand the risks and the steps that were taken

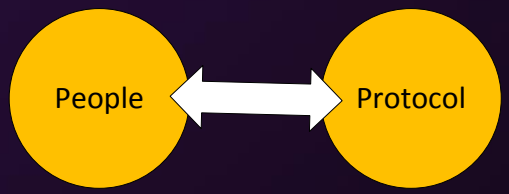
More "Why"s

- Results of data reduction
  - Event identification
  - Environmental conditions
  - Traffic conditions
  - Infrastructure information
  - Secondary tasks
  - Glance locations
  - Driver impairments
  - Vehicle and conflict parameters
  - Event sequence in context
- Results of quality control
  - Allow for multiple people help to code.
  - Confidence that coding is consistent
  - Confirmation that your definitions are clear and complete
  - No (bad) surprises in the end
  - Trust in your overall research results.

Data Reduction Components - People

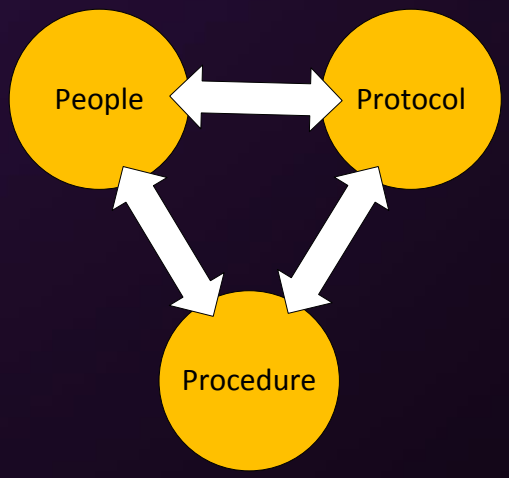
- Researcher(s)
- Project manager(s)
- Staff recruiter(s)
- Staff supervisor(s)
- Staff trainer(s)
- Communications liaison(s)
- Quality control technician(s)
- Data reductionist(s)
- Data analyst(s)

Data Reduction Components: Protocol



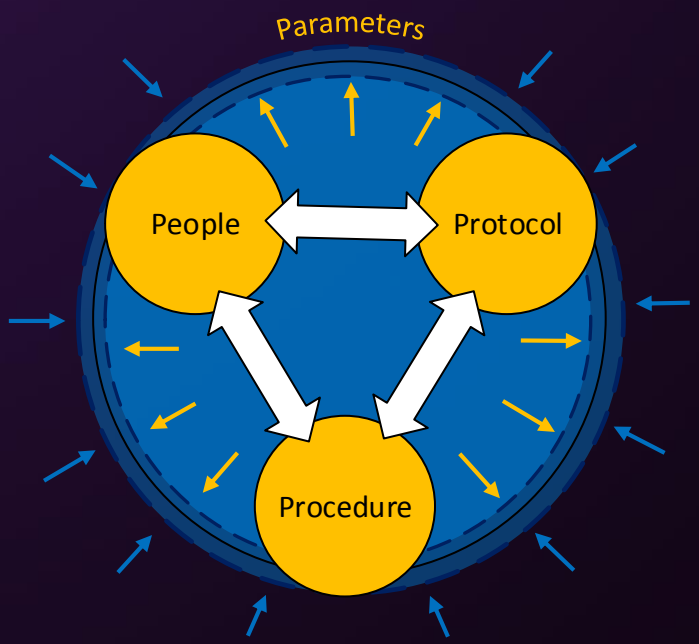
- Operationalizes the research question
- A framework for reduction task
- Relies on people for accurate and consistent interpretation

Data Reduction Components - Procedure



- Instructions for how each member of the team should operate
- Defines communication avenues
- Provides an infrastructure for continuous quality monitoring
- Sets limits to preserve quality, security, etc.

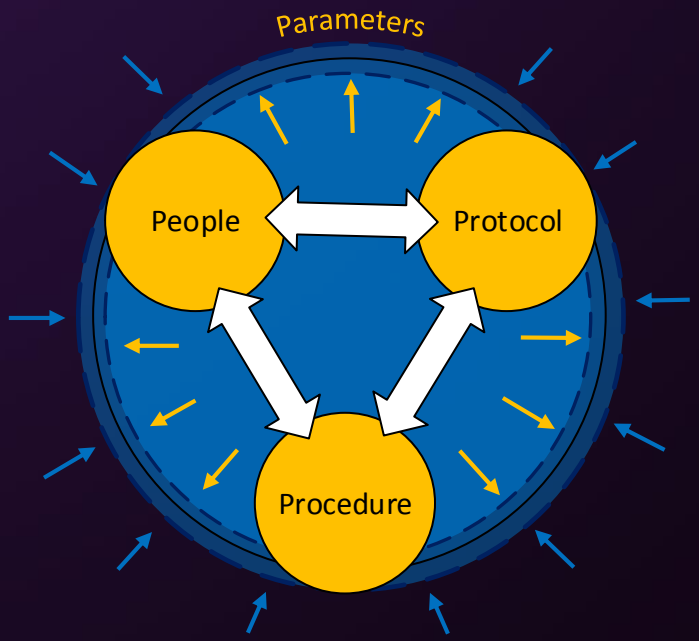
Data Reduction Components – Parameters



Time
Budget
Space
Video Specs
Equipment
Software
Network/Database
Training Tools

- Defines the workspace within which other components must operate (while maintaining quality)
- Often dynamic, either further constraining other components or allowing them to expand

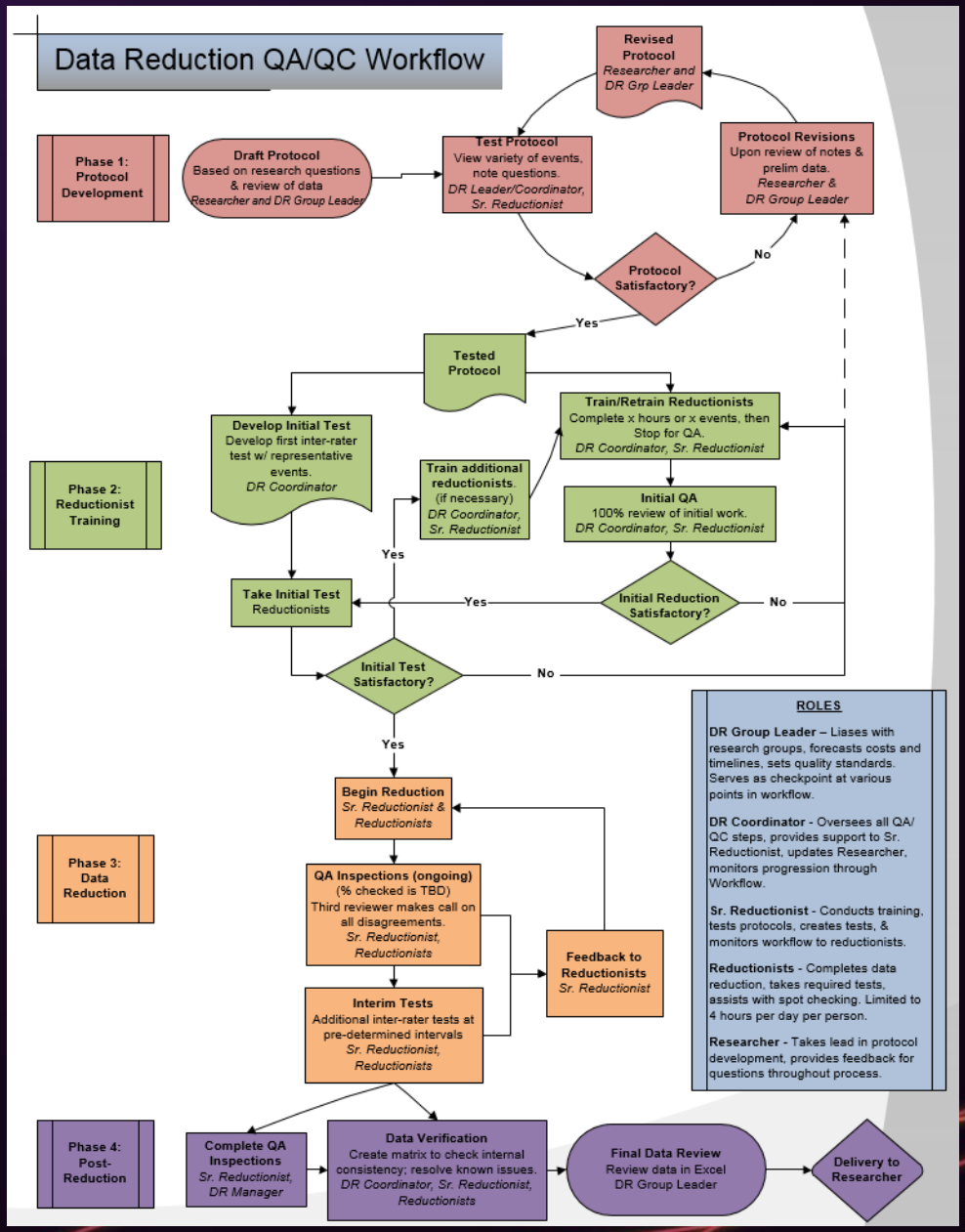
Data Reduction Components - Workspace



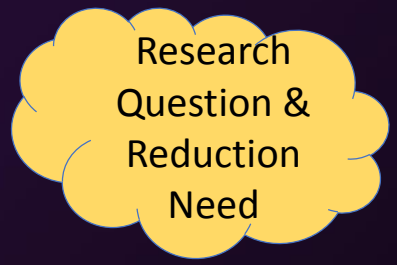
Components interact within a 3-dimensional, dynamic space.

The secret is making them work together!

Quality Control Workflow


Data Reduction QA/QC Workflow
Revised Protocol
Researcher and DR Grp Leader
Phase 1: Protocol Development
Draft Protocol
Based on research questions & review of data
Researcher and DR Group Leader
Test Protocol
View variety of events, note questions.
DR Leader/Coordinator, Sr. Reductionist
Protocol Revisions
Upon review of notes & prelim data.
Researcher & DR Group Leader
Protocol Satisfactory?
No
Yes
Tested Protocol
Train/Retrain Reductionists
Complete x hours or x events, then Stop for QA.
DR Coordinator, Sr. Reductionist
Develop Initial Test
Develop first inter-rater test w/ representative events.
DR Coordinator
Phase 2: Reductionist Training
Train additional reductionists (if necessary)
DR Coordinator, Sr. Reductionist
Initial QA
100% review of initial work.
DR Coordinator, Sr. Reductionist
Yes
Take Initial Test
Reductionists
Yes
Initial Reduction Satisfactory?
No
Initial Test Satisfactory?
No
Yes
ROLES
DR Group Leader – Liases with research groups, forecasts costs and timelines, sets quality standards. Serves as checkpoint at various points in workflow.
DR Coordinator - Oversees all QA/QC steps, provides support to Sr. Reductionist, updates Researcher, monitors progression through Workflow.
Sr. Reductionist - Conducts training, tests protocols, creates tests, & monitors workflow to reductionists.
Reductionists - Completes data reduction, takes required tests, assists with spot checking. Limited to 4 hours per day per person.
Researcher - Takes lead in protocol development, provides feedback for questions throughout process.
Begin Reduction
Sr. Reductionist & Reductionists
Phase 3: Data Reduction
QA Inspections (ongoing)
(% checked is TBD)
Third reviewer makes call on all disagreements.
Sr. Reductionist, Reductionists
Feedback to Reductionists
Sr. Reductionist
Interim Tests
Additional inter-rater tests at pre-determined intervals
Sr. Reductionist, Reductionists
Phase 4: Post-Reduction
Complete QA Inspections
Sr. Reductionist, DR Manager
Data Verification
Create matrix to check internal consistency; resolve known issues.
DR Coordinator, Sr. Reductionist, Reductionists
Final Data Review
Review data in Excel
DR Group Leader
Delivery to Researcher

Quality Control Workflow



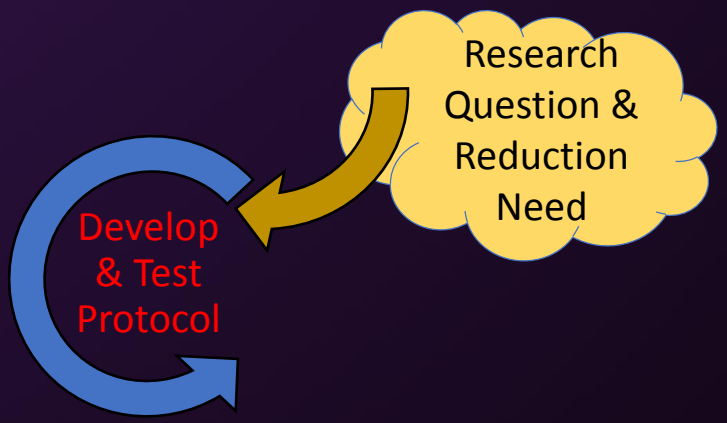
???

Reduced Data for Analysis

Two endpoints to define:

1. Research question & data needed
2. Data analysis plan

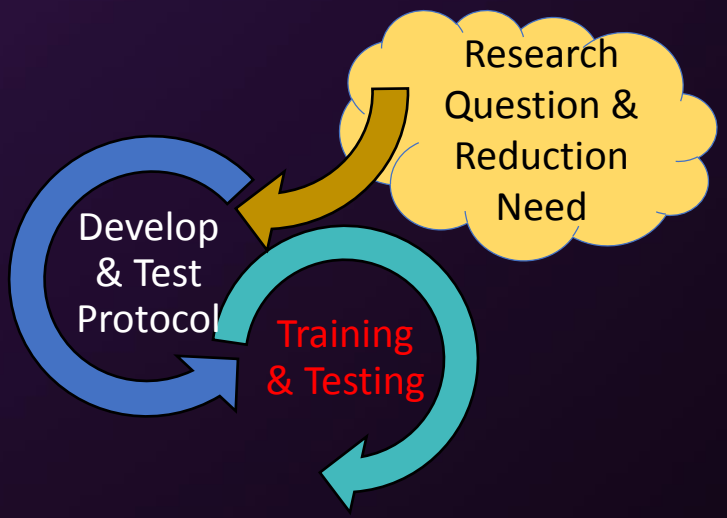
QC Workflow – Protocol



- Write, test, rewrite, retest
- Review ample video
- 3 Goals:
  1. Adequate information
  2. Accurate information
  3. Efficient retrieval
- Expect questions and revisions

Reduced Data for Analysis

QC Workflow – Training/Testing



- Develop a training dataset and a proficiency test
- Train small cohorts
- Train QC staff first
- Sample loop:
  - Review protocol on own
  - In-person training
  - Review training data
  - Complete proficiency test
  - Start over with next cohort

Reduced Data for Analysis

QC Workflow – Feedback

- Ongoing reduction phase
- Two key QC steps:
  1. Continuous checks and feedback
  2. Repeated testing, evaluation, & retraining
- Communication critical

QC Workflow - Verification

- Wrapping up all previous loops
- Check for "sanity" of reduced data
  1. Completeness
  2. Unexpected or unusual responses
  3. Internal consistency
- Prepare final dataset

QC Workflow – In Practice

←Research question + Video data in

←Four overlapping, bidirectional loops

←Reduced data out

Questions to Ask

- Research question + Existing/needed data?
- Analysis plan?
- Roles filled by which people?
- Resources and limitations?
- Protocol existing or needed?
- Key quality concerns?
- How will your team operate and communicate with each other and you?