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

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

Parameters:
- Time
- Budget
- Space
- Video Specs

Components:
- Equipment
- Software
- Network/Database
- Training Tools

People

Procedure

Protocol

Virginia Tech Transportation Institute
Data Reduction Components - Workspace

Components interact within a 3-dimensional, dynamic space.

The secret is making them work together!
Quality Control Workflow
Advancing Transportation Through Innovation
Quality Control Workflow

Two endpoints to define:
1. Research question & data needed
2. Data analysis plan

Reduced Data for Analysis
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
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
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

Develop & Test Protocol

Training & Testing

Continuous Feedback

Data Verification

Research Question & Reduction Need

Reduced Data for Analysis
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?