

# **Video Consumer Mapping Study: Technical Appendix**

**Study Outline: Methodologies, Tools and Instruments**

**Methodological Constraints and Limitations**

**Observer Training Process**

**Observer Feedback Report**

**VCM Survey Instrument Sources and Acknowledgements**

---

# Table of Contents

---

<b>Study Outline: Methodologies, Tools and Instruments</b>	<b>3-5</b>
Core Sample	3
Acceleration Sample	4-5
<b>Methodological Constraints and Limitations</b>	<b>6-10</b>
General Constraints: Human Observation	6
General Constraints: Behavioral Focus	6
General Constraints: The Logging Tool	7
General Constraints: Observational Coding Typologies	7-8
Data Limitations/Imperfections: Category Systems	8-9
Data Limitations/Imperfections: Instruments	9-10
Data Limitations/Imperfections: Other Limitations	10
<b>VCM Observer Training Process</b>	<b>11-19</b>
General Statement of Position Function	11
Characteristic Duties and Responsibilities	11-13
Observer Training Overview	13
Summary of the Training Process	13-14
Examples of Training Support Materials	14-16
Evaluation of Trainee Performance	16-18
On-going Learning in the Field	18
Observer Feedback on Training	18-19
<b>Observer Feedback Report</b>	<b>20-25</b>
Background	20-21
Post-Session Questionnaire Results	21-25
<b>VCM Survey Instrument Sources and Acknowledgements</b>	<b>26-27</b>
Media Devices and Services Inventory	26
Big 5 Personality Survey	26
Innovativeness Profile	27
Yankelovich Media Usage and Gratifications Battery	27
Digital Transition Preparedness	27
Recall of Media Exposure	27
Observer Session Report	27

---

## Study Outline: Methodologies, Tools and Instruments

---

For the **Core Sample**, after recruitment the following represents the sequence of events:

1. Pre-Observation mailing by The Nielsen Company
  - a. Demographics questionnaire
  - b. Informed consent form
2. Pre-Observation Call, Wave 1 (Spring 2008):
  - a. Devices and Services Inventory
  - b. Scheduling the Observation Day
  - c. Reminder call 24 to 48 hours before scheduled observation.
3. Day of Observation (from 3/26/08 to 6/28/08, with two lagging cases in July):
  - a. Observer/Respondent reconstruction of early part of day before Observer arrival
  - b. Full day of Observation, split into two sessions between two Observers who each cover approximately 7 ½ hours of the day.
4. Wave 1 Follow-up Call (24-72 hours after Observation Day):
  - a. Reconstruction of remainder of Observation day after the Observer left.
5. Pre-Observation Call, Wave 2 (Fall 2008):
  - a. Scheduling the Observation Day
  - b. Reminder call 24 to 48 hours before scheduled observation.
6. Day of Observation (9/02/08 to 11/13/08):
  - a. Observer/Respondent reconstruction of early part of day before Observer arrival
  - b. Full day of Observation, typically split into two sessions between Observers who each cover approximately 7 ½ hours of the day.
  - c. Leave behind self-administered printed survey including:
    - i. Revised Devices and Services Inventory
    - ii. Big 5 Personality Survey
    - iii. Yankelovich Media Uses and Gratifications Battery
    - iv. Innovativeness Profile
7. Wave 2 Follow-up Call (24-72 hours after Observation Day)
  - a. Reconstruction of remainder of Observation day after the Observer left.
  - b. Recall of Media Exposure questionnaire
  - c. Digital Transition Preparedness survey
8. Observer "Session Reports" completed including any notes on unusual circumstances and comments on observer effect on respondent behavior.

For the **Acceleration Sample**, variations from core sample procedure are noted with an asterisk (\*). After recruitment, the sequence of events were as follows”

1. Pre-Observation Mailing by Ball State University, Wave 1 (Spring 2008):
  - a. Demographic form
  - b. Informed Consent form
  - c. Video Release form \*
  - d. Innovativeness Profile\*
2. Pre-Observation call, Wave 1:
  - a. Devices and Services Inventory
  - b. Scheduling the Observation Day
3. Day of Observation (5/10/08 to 6/27/08):
  - a. Observer/Respondent reconstruction of early part of day before the Observer arrival
  - b. Full day of Observation, split into two sessions between two Observers who each cover approximately 7 ½ hours of the day.
4. Wave 1 Follow-up Call (24-72 hours after Observation Day):
  - a. Reconstruction of remainder of Observation day after the Observer left.
5. One Month Window to Purchase New Media Device in Acceleration Panel Catalogue.\*
  - a. Participants typically had several telephone interactions with BSU CMD staff while finalizing payment and delivery details.
6. Pre-visit Call, Wave 2 (Fall 2008):
  - a. Scheduling the Observation Day
7. Day of Observation (9/16/09 to 11/10/09):
  - a. Observer/Respondent reconstruction of early part of day before the Observer arrival
  - b. Full day of Observation, split into two sessions between two Observers who each cover approximately 7 ½ hours of the day.
  - c. Leave behind self-administered printed survey including:
    - i. Revised Devices and Services Inventory
    - ii. Big 5 Personality Survey
    - iii. Yankelovich Media Uses and Gratifications Battery
8. Wave 2 Follow-up Call (24-72 hours after Observation Day)
  - a. Reconstruction of remainder of Observation day after the Observer left.
  - b. Recall of Media Exposure questionnaire
  - c. Digital Transition Preparedness survey

9. Observer “Session Reports” completed including any notes on unusual circumstances and comments on observer effect on respondent behavior.
10. After Follow-up Call, within four weeks after Wave 2 Observation, recording of video interview with a subset of Acceleration respondents.\*
  - a. Questions addressed such issues as major life changes between waves of observation, decision processes for choice of acceleration items, item delivery and setup outcomes, device or service likes and dislikes, patterns of use, and sources of content.

---

## Methodological Constraints and Limitations

---

Researchers learn firsthand the wisdom of the systems theory adage "If you want to learn anything you can't learn everything." Research design is a series of decisions about trade-offs between what is desirable and what is possible. This study is no different; the data are exceptionally comprehensive and granular but also necessarily constrained and imperfect. The general constraints stem from features of the research method including reliance on human observers; the behavioral focus of the method; technical restrictions of the data logging tool and restrictions of pre-defined typologies for location, activity and media exposure. The data limitations and imperfections stem from design decisions found, in retrospect, to have unanticipated negative consequences or to have overlooked more satisfying solutions.

### General Constraints

#### Human Observation

There are no electronic monitoring systems which can capture media exposure (much less participant activities) across all media platforms and all locations. We therefore rely on a "shadowing" method in which human observers follow participants throughout the day. Shadowing overcomes the memory limitations and social desirability bias of recall data and the compliance problems of participant diary data; however, the data are limited to what can be observed and logged by a trained and attentive observer. (Observer reliability and "observer effect" issues are addressed in the Observer Training Process section and the Observer Feedback Report.) The reliance on human observation limits sample size, the length and complexity of the observational category systems and the minimum time granularity of recorded data.

#### Behavioral Focus

We often apply the language of "media exposure" rather than "media use" in this report. To evaluate an event as "media use" is to make a complex interpretation of cues to a participant's cognitions and motivations. Media exposure, in contrast, is a concrete, behavioral observation. For example, if a participant is positioned to see and/or hear TV content a media exposure is taking place. We don't know if the participant is influenced by the content, "paying attention" to it or ignoring it; we do know the participant is exposed to the content. The most subjective observer judgment called for in this study is to judge, in cases of concurrent media exposure, which medium appears to have the participant's primary attention. The behavioral focus of the method means we know "what" but not, from observation alone, the "why" or "so what." These issues are addressed by supplemental self-report instruments and—in the acceleration module— qualitative interviews.

## The Logging Tool

The Alphasmart Dana™ smart keyboard is used for entry of observational data. Its advantages include light weight, long battery life and physical robustness (i.e., it can withstand the inquisitive attention of dogs and small children). The Media Collector™ software developed by BSU's Center for Media Design supports quick and easy touch-screen and keyboard data entry while the on-screen menus provide cognitive support to the observers by reminding them of available categories. The limitations imposed by the Dana result from relatively small screen size (constraining category menu length, as we seek to avoid excessive scrolling) and finite response time due to processor limitations.

Data are logged in near-real time; that is, observers make entries as required to indicate changes in location, activity and/or media exposure state. The current state is time-stamped and logged to a data file every 10 seconds. Each record contains the complete current state: location, activity, media and any observer comments. This allows a simple data file structure and avoids overloading the system during periods of rapid change. The tradeoffs are difficulty in recording very short events (e.g., phone calls or texting taking less than 10 seconds) and in aligning simultaneous changes in location, activity and media. Observers log simultaneous changes in location, activity and media in that order. However, a 10-second interval may elapse during the observer's data entry, triggering writing of a record so that the location change is recorded to one 10-second interval and the activity and/or media change is recorded in the next record. This peculiarity of the logging software can result in a small number of observations with incompatible values across the category domains.

The Media Collector™ software places a heavy load on the Dana™ CPU (especially its clock-related functions). One infrequent consequence is for the device to "freeze up" and require rebooting. These rare system crashes typically create a void of one to two minutes in the data. If the location, activity and media states are identical before and after the system failure and the "data gap" is short, we auto-fill the missing records with the values of the records before and after the gap.

## Observation Coding Typologies

The location, activity and media exposure typologies are the heart of the method. They define what can be logged during observations and consequently what constitutes available data for analysis. While it is a truism that "the map is not the territory" (i.e., our descriptions of the world should not be confused with the world itself), all we have for analysis is the map--in this case, is the parsing of the stream of lived behavior into discrete categories within coding typologies.

Design criteria for coding typologies include clarity, comprehensiveness and utility. Each category needs to be unambiguously defined and easily distinguished from other categories in a typology to allow for reliable application. Each typology should constitute a complete and valid

inventory of the target domain, designed so only a small portion of entries end up in a catch-all "Other" category (without expanding beyond the cognitive limits of observers). Each typology must be able to generate data useful for answering the research questions.

The coding typologies used in this study were based on systems used in prior field research by the Center for Media Design; these were updated and modified through extensive discussion and review by the Media Consumption and Engagement committee of the Council for Research Excellence. They represent a balance between providing broadly accessible and complete descriptions of locations, activities and media (to describe the interplay of the media ecosystem with daily routines) and capturing details of special interest for this research (to describe video's niche in the ecosystem). Although the typologies appear rich and detailed, we distill the "blooming, buzzing confusion" of the world down to seven location types, sixteen life activities (adapted from the U.S. Bureau of Labor Statistics American Time Use Studies activity categories) and eight major media categories (with over sixty minor categories). The frustratingly-finite typologies are easily second-guessed ("why differentiate X but not Y?") as they are not designed to meet every conceivable research need. They were the research sponsor's and providers' consensus inventory of key components of the location, activity and media domains prior to the execution of the research. These decisions were coded into the Media Collector™ interface and observer training materials and were not easily revised once implemented. Several weaknesses in the category systems were discovered during field research. These are discussed in the next section.

## **Data Limitations and Imperfections**

### **Category Systems**

**Location:** The seven top-level location types were unproblematic for observers. However, the five room-specific subcategories within "own home" and "other's home" posed problems in the field as observers frequently encountered home layouts in which it was difficult to classify particular rooms or spaces or to establish distinct boundaries between spaces. Our analysis has therefore concentrated on the top-level categories.

**Activities:** The activity coding system requires exclusive categorization; activity multitasking is not accommodated. In this respect the activity coding is not as complete or rich as the media coding (in which concurrent media are possible). Observers sometimes had to judge which, of multiple concurrent activities, should be coded. The prioritizing criteria were "What would the participants likely say if you asked them what they were doing?" and "Does one activity appear to supplement or support the other?" For example, snacking is a "personal needs" activity. However, snacking during child care would not be logged as it would be superseded by "care of another."

**Media:** Categorization of media exposure becomes more difficult as media converge and the content boundaries between TV, computer and smart phones blur. We define media categories by platform and by content type, in that order. TV content watched on a computer is coded as

computer video rather than TV. Multiple media can be logged as concurrent exposure. When this happens the observer is asked to judge which medium has primary attention. Overt behavioral cues (gaze direction for visual media, media integration into a work task, etc.) are the basis of these decisions. We cannot, of course, see inside the "black box" of actual attention or detect attention shifts occurring without behavioral cues.

Perhaps the most marked weakness in the media category system results from the fact that the media categories were developed before the video hierarchy (established later for analytical purposes). The initial classification of TiVO™ /DVR was as a video category rather than as a kind of TV exposure. The TV subcategories for content genres were not duplicated under TiVO™/DVR. Therefore we are not able to determine, for example, if altered speed playback was associated with particular TV genres (such as the ad/promotion subcategory). Given the importance of this issue to the future of measurement, this must be recognized as a major limitation despite the relatively small amount of total media spent in the TiVO™/DVR category.

Another video-related category weakness is the conflation of DVD and VCR video playback into a single coding category. This reflected the CMD's prior practice but overlooked the importance for this study of distinguishing VCR playback of commercially recorded tapes (e.g., movies) and playback of time-shifted TV through self-recording of content. Although it can be presumed that VCR-based time shifting has a limited future as the technology is supplanted by digital recording, we missed an opportunity to quantify its current reach and duration. In the second round of observation we asked observers to add a qualitative comment when self-recorded TV content was played back from a VCR; however, we had insufficient instances of such notes to be confident in their reliability and therefore did not incorporate them into the analysis.

The third video-related category limitation is our failure to distinguish VOD in the TV subcategories. VOD was instead coded as "TV – other," along with any other content genre not fitting the defined subcategories. A separate breakout of VOD reach and duration was therefore impossible. This oversight was noticed during preliminary analysis of round one results and observers were charged with noting VOD use via comments in the second round of observation.

## Instruments

The supplemental instruments are subject to all the familiar limitations of self-administered measures. Age was not reported by several participants and sex was miscoded for one participant; these errata were partly or wholly addressed after the analysis reported here. Future tabulations will report results from corrected data.

We noticed inconsistencies in "device and services inventory" responses from first round to second round (such as going from owner to non-owner for a technology or service). These inconsistencies likely result from participant carelessness or confusion over terminology (especially for TiVO™/DVR and access to cable or satellite HD channels).

The overall response rate for the supplemental instruments was good but lower for complex, multi-page instruments such as the Media Uses and Gratifications survey. Analyses based on instruments therefore vary in their sample size.

The Uses and Gratifications instrument was applied to a variety of video platforms; in retrospect this diminished the utility of that data. Content and platform similarities resulted in few noticeable differences in uses and gratifications among the target video media. Richer interpretation would have been possible had non-video media (such as recorded music) been included in the Uses and Gratifications surveys for purposes of comparison and benchmarking.

### **Other Limitations**

The timing of the two rounds of observation (spring and fall of 2008) is such that we cannot treat this as a true longitudinal study. Seasonal variations, possible lingering impacts of the Olympics on viewing patterns, the impact of the TV writer's strike on spring programming and possible heightened interest in news content in the late stages of the fall election are potential confounding factors in explaining over-time differences in media exposure. We have focused instead on analysis of the acceleration cohort and natural device adoption groups in the core cohort to inform speculation about future changes in the media ecosystem.

---

## Video Consumer Mapping Observer Training Process

---

The Video Consumer Mapping study is, at its core, an observational study; it is based primarily on data produced by human observers categorizing the observed locations, life activities and media exposures of participants throughout the course of the observed day. The study applies some traditional pencil-and-paper instruments (such as a personality profile and innovativeness scale) and telephone interview instruments (such as next-day recall of media exposure and a survey of readiness for the digital TV transition) but the most important data are generated by observers. The observer training process is designed to "calibrate" the observer as an effective research instrument. The optimal observer is adaptable, unobtrusive, able to politely and gently deflect interruptions, and capable of rapid, consistent and accurate logging of observations using the observational tool (the Alphasmart Dana™) and category systems detailed in the technical appendices to the study report.

The VCM study utilized full-time observers hired and trained by BSU specifically for the study. The official position posting outlined the duties, responsibilities and minimum and preferred qualifications for the position:

### **General Statement of Position Function**

Responsibilities include accurately conducting observational field research of human behavior with a hand-held computer device, performing follow-up telephone interviews with participants; attending a thorough orientation and training session; supporting team members; maintaining a flexible schedule and being able to adapt to new working environments daily; reviewing and cleaning data records using Excel and maintaining strong communication with team members and supervisors.

### **Characteristic Duties and Responsibilities**

1. Attend observational training sessions (40 hours +) and subsequent training events and meetings related to specific research projects, as needed.
2. Obtain necessary IRB certification in order to comply with standards governing human research protocols (completion of online NIH tutorial).
3. Maintain consistent and reliable communications with research staff during the data collection and analysis period of assigned projects. This includes having a reliable email account and being reachable by a designated phone number.
4. Be on time for scheduled observations and stay with participant until shift is over.
5. Have the ability to establish a rapport with participants while maintaining an objective, non-obtrusive research presence.
6. Schedule and adhere to research equipment pick-up schedule prior to assigned observations and return equipment after observations in a timely manner.

7. Review observational data files and enter corrections or additions as needed.
8. Perform follow-up telephone interviews with participants.
9. Be available for five eight-hour shifts in a given week.
10. Dress in “business casual” attire for all observations and represent Ball State University in a positive manner while on assignment.
11. Be familiar and interested in the latest consumer media devices and technology.
12. Perform other project-related duties as assigned.

Minimum Qualifications:

- Bachelor’s degree.
- Research experience and/or training in field research, observation or ethnographic research, focus group facilitation, face-to-face survey work and/or other research techniques.
- Must be willing to travel and be at field locations for extended periods of time (housing is provided while in non-local field locations).
- Must possess a current valid driver’s license.
- Must have excellent communication, organizational and project management skills.
- Must have a reliable email account and be reachable by a designated phone number.
- Must have strong interpersonal skills.
- Must be able to complete the NIH on-line tutorial on the rights of human subjects in research and provide certificate of completion from that tutorial.
- Must be willing to work weekends and/or evenings as needed.
- Cannot have extreme allergies to dogs or cats.

Preferred Qualifications:

- A Bachelor’s degree in a social science field.
- Bilingual (English/Spanish) applicants are encouraged to apply

Observer selection was based on review of qualifications and interviews of qualified candidates. In total, 49 observers were hired in spring of 2008; of these, 42 successfully completed training (86%; five dropped out and two failed performance tests). Twelve of the round I observers did not return for the second round of observations in fall of 2008 due to obtaining full-time ongoing employment elsewhere or other changes in circumstances (we regret to report the

death of one observer due to illness during the summer of 2008). All 14 new observers hired in fall of 2008 successfully completed training. Not all trained observers were extensively utilized; some served as reserve observers in the Indianapolis DMA to ease scheduling challenges.

### **Observer Training Overview**

Observer training prior to the first round of observation was completed at Ball State University; refresher training and training of replacement observers prior to Round II in Fall 2008 was completed in the DMAs selected for the study. All retained observers originally training for Round I in Spring 2008 completed refresher training.

The full training typically lasted four days and addressed goals in three domains:

1. Affective goals

- Observers will recognize importance of the study
- Observers will recognize importance of their performance
- Observers will feel motivated for top performance

2. Cognitive goals

- Observers will know relevant study policies and procedures
- Observers will master location, activity and media coding systems
- Observers will know how to write clarification and correction comments

3. Behavioral goals

- Observers will demonstrate basic operation of the Dana device
- Observers will demonstrate proper non-obtrusive behavior
- Observers will demonstrate acceptable inter-rater reliability

### **Summary of the Training Process**

The trainers were Bonnie Krupa, CMD Field Project Supervisor, and Wesley Withers, CMD Project Manager. Both had extensive prior experience in observational research project management. Both served as trainers in prior CMD observational research. CMD Insight & Research leadership (Mike Bloxham and Michael Holmes) participated early in the training to provide an overview of the history of CMD's observational research, discuss importance of the research to the media industry, and emphasize the crucial role of the observer in generating rigorous research.

Training materials and experiences incorporated into the four days of training included printed reference materials such as observer policies and procedures and category system definitions (see appendix), interactive materials including extensive training PowerPoint™ slides and training videos for in-session practice (also made available on a CDROM for later reference),

half-day practice observations with on-campus volunteer participants, paper-and-pencil tests of cognitive mastery of the material and a video-based real-time coding test of mastery of the observation device and category systems.

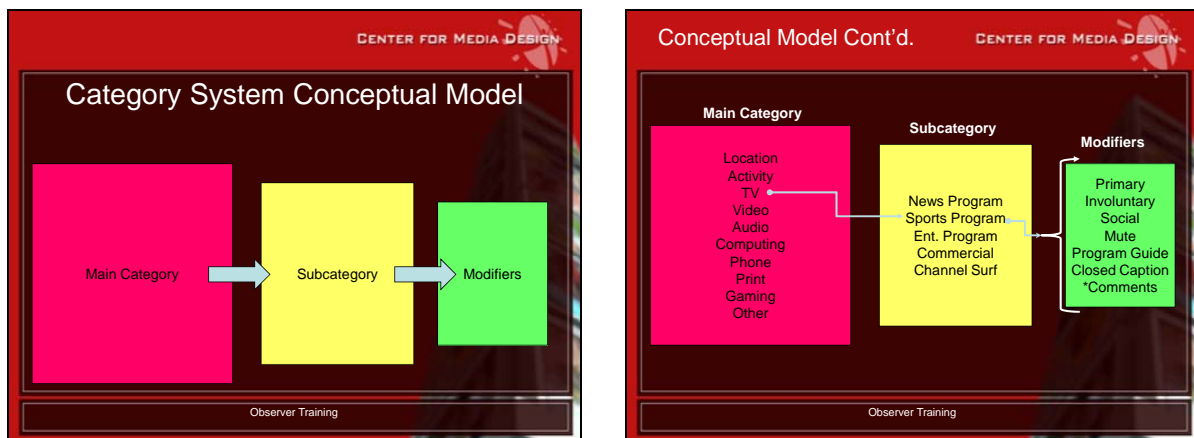
Training addressed a variety of topics beyond application of the coding systems. Observers were expected to be familiar with the observer responsibilities outlined above (including the rights of human research subjects), know and comply with guidelines for ensuring personal safety in the field, be aware of all grounds for termination of employment, be familiar with operation of the Dana™ device and Media Collector™ software, be able to operate the Observation Management System™ for scheduling of observations, and know how to reduce the obtrusiveness of their presence and minimize "observer effect."

Refresher training for the second round of observation was one day in length. PowerPoint™ materials for refresher training (and the decks for initial training of replacement observers) were modified to reflect learning from the experiences of round I – that is, category definitions further clarified with extra additional to challenging categories, procedural problems reviewed, additional examples added, minor changes to the software interface introduced and "rules of thumb" for dealing with unexpected situations encountered in the first round noted. Focus groups were held with observers after the first round of observation to identify areas for additional reinforcement or clarification in preparation for refresher training.

### Examples of Training Support Materials

The hundreds of slides used in the VCM training decks are too numerous to reproduce here. Selected slides are displayed here to indicate the range of cognitive and behavioral content addressed in training.

Example Slides: Category System Levels



Copyright (c) 2008 Ball State University

## Example Slides: Category Definition

CENTER FOR MEDIA DESIGN

### Locations (1)

1. Own home: Living/Family/TV room	9. Car
2. Own home: Kitchen	10. Public transportation
3. Own home: Bedroom	11. Work
4. Own home: Other room	12. School
5. Other's home: Living/family/TV room	13. Other*
6. Other's home: Kitchen	
7. Other's home: Bedroom	
8. Other's home: Other room	

Observer Training

CENTER FOR MEDIA DESIGN

### 1.1 - Location - Own home: Living/Family/TV room

- Definition
  - Living Room/Family Room – standard definition
  - TV Room – Primary shared television viewing room
  - If the living room serves multiple functions (i.e. also as a bedroom in a studio apartment), always default to Living Room for that space.
- Modifiers
  - Outside [OUT]

Observer Training

Copyright (c) 2008 Ball State University

## Example Slides: Introduction to the Dana™

CENTER FOR MEDIA DESIGN

Observer Training

CENTER FOR MEDIA DESIGN

### Home Screen

Observer Training

Copyright (c) 2008 Ball State University

## Example Slides: Introduction to the Media Collector™

CENTER FOR MEDIA DESIGN

Observer Training

The main Media Collector screen begins in an un-initialized state. To start the Media Collector recording records, you must indicate a baseline Location and Life Activity. Once this is done, the media category buttons will appear and records will be written every 10 seconds. Location and Activity are mutually exclusive categories and must always have a selection indicated.

CENTER FOR MEDIA DESIGN

### Indicating a Baseline Location

Step 1: Tap Category

Step 2: Select Subcategory

Step 3: Check 'On' to activate subcategory

Observer Training

Arrow indicates that the subcategory is 'ON'.

Subcategory-specific modifier codes appear once item is toggled on.

Copyright (c) 2008 Ball State University

## Example slides: Video Demonstration



Copyright (c) 2008 Ball State University

Video demonstration slides combined videos of media exposure examples with animated screen captures of the steps for recording the event in the Media Collector™ software.

### Evaluation of Trainee Performance

During the training process the trainers noted which trainees appeared to be struggling with the category system or the cognitive demands of real-time coding of training videos. Several of the trainees who withdrew during training were counseled to do so by the trainers.

All observers completed, at the close of training, a 59-item traditional paper-and-pencil exam with questions addressing study policies and procedures (25% of the exam) and application of the location, life activity and media coding systems (75% of the exam). Performance on the written exam was quite strong, with an average score of 96% (range 88% to 100%).

### Inter-Coder Reliability Testing

The last day of each training session included trainee coding of a test video of 60 minutes in length. This compiled video presented a wide variety of location, activity and media exposure changes, with changes presented somewhat more quickly than would typically be encountered in the field. The trainers generated a "canonical" coding of the video against which trainee performance was measured. Comparison was made on a record-by-record basis. Due to the

Video Consumer Mapping (VCM) Study from the Council for Research Excellence: © The Nielsen Company; conducted by Ball State University's Center for Media Design (CMD) and Sequent Partners.

impossibility of precise synchronization of Dana™ internal clocks some inter-coder disagreement is generated by differences between devices on precisely when a 10-second record is written relative to observer entries. This most impacts reliability calculations for categories with short episodes and frequent changes (e.g., life activities, genres within major media).

Simple agreement on location and activity categories were typically in the 85% to 90% range. Agreement varied more across media categories; some, such as TV, video, radio and software, tended to very high simple agreement (90% or better). In contrast, simple agreement for short-episode media such as phone use tended to lower agreement (65% to 75%). Differences in observer data entry speed and the difficulty in precise alignment of records noted above have more marked impact on these short-episode media. The number of observations is small in these instances and the measurement error introduced by Dana™ timer differences will have a bigger influence on results.

Simple agreement ratios are an unsatisfactory measure of reliability because they don't discount for coder agreement that could happen by chance. For example, if the test video and the observer both record considerable TV exposure, the observer could score relatively high on simple agreement even if his or her coding is random. We instead examined reliability using Cohen's  $\kappa$ . This a more rigorous measure as it adjusts for the agreement that would be expected by chance.

Evaluation of Cohen's traditionally applies the following reliability ranges:

.8 to 1	Very Good
.6 to .8	Good
.4 to .6	Moderate
.2 to .4	Fair
< .2	Poor

With few exceptions, trainees obtained values in the "Very Good" range for Location categories. For Activity categories they typically achieved reliability in the "Good" range, reflecting the impact of more rapid changes and a more difficult judgment task.

Reliability for major media varied as expected given the simple agreement levels already noted. Observers achieved values in the "Very Good" range for TV, video, radio and software categories. The "Good" range was typical for Web, email, landline phone and mobile phone. "Moderate" reliability was obtained for audio (excluding radio) and print. No media categories consistently generated "Fair" or "Poor" reliability.

Observers who generated lower reliability for one or two of the media were retained if the coding confusion could be easily corrected (that is, if the source of error was recognizable and easily addressed). The observers who failed and were not employed in the field were those who generated substantial errors across multiple media.

Re-testing during refresher training generated similar results for the returning observers. Reliability indexes tended to be the same or higher in re-testing; this was expected given their extensive field experience. Note, however, that the test video was the same as used in the first round of training. Months had elapsed between the two tests, as well hundreds of hours of observation in the field, but some observers may have retained enough memory of the original testing to influence their performance in the second test.

### **Ongoing Learning in the Field**

Formal training was limited to four days but informal training and reinforcement continued throughout fieldwork. As noted above, each observer also had access to the complete set of training materials for reference as needed. Designated team leaders in each DMA served as liaisons between the observers and project managers to pass along questions, clarifications and recommendations. Observers maintained shared knowledge by frequently sharing "tales from the field" with each other, discussing unusual situations or difficult coding decisions. Project managers alerted team leaders and observers to issues encountered during data cleaning (i.e., requests for confirmation or clarifications of a coding decision) to aid reliability. Observers received targeted reinforcement from project managers if data cleaning processes revealed they were struggling with particular coding judgments. This learning was also incorporated into the updated initial training and refresher training materials for the second round.

### **Observer Feedback on Training**

Post-session evaluation by trainees of training outcomes reflected a high level of confidence in their preparedness for the field. Readiness responses typically averaged 6 or higher on a scale from "1 = Not at All" to "7 = Very Much." The highest average score was for the item about the most straightforward and least complex content – familiarity with the purpose of the VCM study (6.8). The lowest average score was still high (5.9 on the 7-point scale); this was for confidence in use of the media codes. This is unsurprising given the scope and complexity of the media category systems.

Question	Mean (SD) (n=39)
I feel ready for my first observation	6.1 (0.79)
I am confident in my use of Location codes	6.4 (0.67)
I am confident in my use of Activity codes	6.2 (0.63)
I am confident in my use of Media codes	5.9 (0.77)
I am familiar with policies and procedures for being an observer	6.5 (0.72)
I feel ready to deal with unexpected events	6.1 (0.95)
I understand the purpose of the study	6.8 (0.43)

Open-ended comments expressed general satisfaction with the training, albeit noting the amount of information could be overwhelming and suggesting more time be allotted to practice observation sessions.

Videotaped qualitative interviews were performed after the first round of observation in order to provide the Media Consumption and Engagement Committee with perspectives on the observer experience. The questions addressed participant reactions to the observer, perceptions of media use patterns, difficult or unusual experiences, etc., rather than asking for reflections on training. However, no observers volunteered comments critical of their preparation for the field experience. One observer noted, however, that it would be valuable to have more opportunities in the on-campus practice observations to practice deflecting participant attempts at conversation.

---

## Observer Feedback Report

---

Research methods differ in what kinds of questions they are most suited to answer and in the trade-offs or compromises inherent in applying them. The "shadowing" approach to observational research is especially well-suited for providing a single-source, cross-platform, cross-location view of consumer behavior in the media ecosystem; it is also the best currently available method for studying concurrent media exposure across all possible platform combinations. These advantages come with a potential price—one which is quickly recognized and often raised by research audiences: "What about observer effect? Doesn't the presence of an observer change participants' media use?" The concern is that participant self-consciousness and self-monitoring will result in non-typical media experiences during the day of observation. We have several reasons for discounting significant observer-effect impacts on results. These include convergence with other measures, the often-overlooked social context of much media exposure, the cognitive load of mindfulness of being observed and observer session feedback questionnaires.

As noted in the report, there is remarkable convergence between this study's TV, computer video and mobile video exposure results and the Nielsen Three-Screen Study from the same time period, and close agreement between results for total ad/promotion exposure and Nielsen estimates. Convergence on these key measures suggests the observer's presence does not distort participants' video exposure in meaningful ways.

Another reason observer impact may not be discernable is that it is not uncommon for media exposure to occur with other people present; there is already a social impact on exposure and content choices so the presence of an observer does not have noteworthy additional impact. Additionally, the participant may not be the person in the setting making the media choice; someone else in the home, workplace or other setting may figuratively or literally "have the remote."

A subset of Acceleration cohort members were invited to participate in qualitative interviews. The interviews yielded anecdotal evidence about the difficulty of keeping an observer in mind, as participants often noted "I forget he was there" or "After a short while I just went about my day." The demands of home and work activities consumed the cognitive resources necessary to be constantly mindful of the observer.

Observer effect is impossible to measure directly or to quantify in a field study like this one; however, there are some behavioral cues to "mindfulness" towards the observer which can be appraised, such as direct comments by participants on changes in routine on the day of observation or repeated attempts to engage the observer in conversation or other participation in a social setting. At the Media Consumption and Engagement Committee's suggestion, we added a post-session observer feedback instrument in the second round of observation. Its

purpose was to gauge observers' perceptions of the impact their presence was having on participants. During "refresher" training for the second round observers were instructed to be sensitive to any cues indicating "observer effect" and to complete the instrument after each observation. The instrument included the following questions:

- Did anything surprising, funny or unusual occur during this observation? If so, please provide a description of the incident. (open-ended response)
- Your perception of participant's overall level of comfort with being observed (scale 1=Very comfortable, 5= Not at all comfortable)
- Frequency of participant's attempts to engage you in social conversation (scale 1=Not at all, 5=Frequently throughout the session)
- Your perception of the overall impact of "observer effect" on participant's behavior (scale 1=No impact, 5=Extreme impact)
- Your perception of the overall impact of "observer effect" on participant's media exposure during the session (scale 1=No impact, 5=Extreme impact)
- If the answer to the above question 4 is a "3" or above, please provide a brief description of the impact(s) (open-ended response)

For observation session with participants in the Acceleration cohort, additional questions addressed the participant's overall level of media exposure and suitability for the qualitative interview (not reported here).

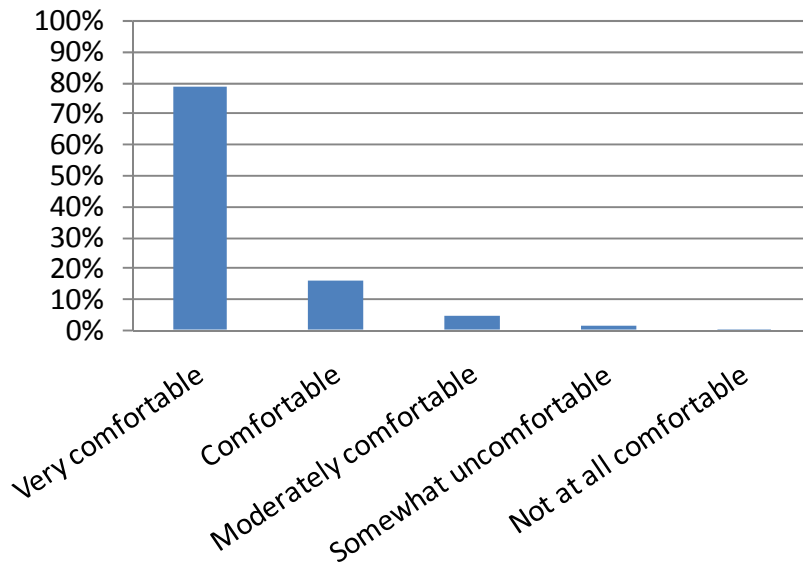
These are, of course, not objective measures and are themselves subject to the familiar cognitive and social desirability biases of self-report items; they do, however, provide additional insight and can inform appraisals of the impact of observer effect on the results of this study.

### **Post-Session Questionnaire Results**

The second round of observation (fall 2008) included 476 participants. The norm is two observers per day with a shift change in early- to mid-afternoon, or a total of 952 observation sessions. However, 86 observations did not include a shift change and were completed by a single observer due to challenges in scheduling for gender, ethnicity or language matches between participant and observer. The total session count is therefore 866. Feedback forms were completed for 802 sessions, a 93% compliance rate.

The first question in the instrument prompted observers to report surprising, funny or unusual events during the observation. These responses are informative about the lives of participants; they reveal, for example, instances of going fishing, spending days in meetings or long religious services, arguments with spouses, the presence of amusing or remarkable children or pets, etc. As they do not directly address observer effect these comments are not discussed here.

Figure 1. Perceived comfort with being observed

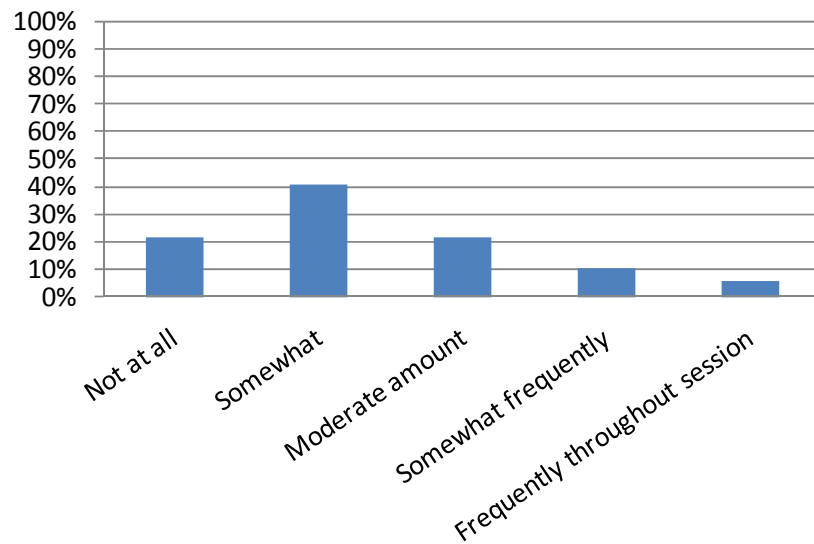


Observers described participants as comfortable with being observed (Figure 1), with a mean of 1.29 ( $SD=0.63$ ) on a five-point scale from 1=Very comfortable to 5=Very uncomfortable. Only 46 sessions (5.9%) were reported as having participants appearing less than comfortable; of these only 12 sessions involved participants rated less than moderately comfortable.

We performed a paired-sample t-test on this item for those participants for whom two session reports were available ( $n=323$  to test the common assumption that participant comfort may increase over the day. There was no significant difference in perception of participant comfort between first and second sessions ( $M1=1.27$ ,  $M2=1.32$ ;  $t=-1.34$ ,  $df=322$ ,  $p >.05$ ).

One overt form of observer effect is manifest as attempts by participants to engage observers in conversation despite prior instructions to ignore them. Observers are trained to politely but firmly decline such invitations to conversation. Given social politeness norms and the level of participant comfort with the observer's presence, it is unsurprising that conversation attempts are common and, in some cases, persistent (Figure 2). Observers informally reported that frequent conversation attempts tended to be associated with participants who were older, retired or unemployed and living alone. For these participants the presence of the observer offered an attractive opportunity for social interaction.

Figure 2. Perceived frequency of conversation attempts



The mean rating for perceived frequency of conversation attempts was 2.37 ( $SD=1.1$ ) on a five-point scale from 1=Not at all to 5=Frequently throughout the session. Almost 40% of sessions were reported as having at least a "moderate amount" of conversational attempts.

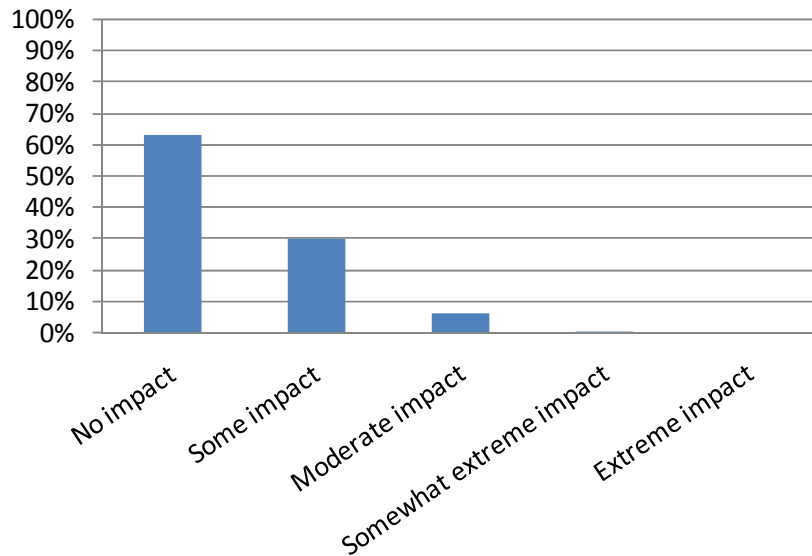
The results suggest it is difficult for some participants to ignore the opportunity to initiate conversations with observers. The "Somewhat frequently" and "Frequently" responses (15.9%) reflect the persistent participants who continued to pursue conversation with observers despite repeated deflections.

There was no difference between sessions for perceived frequency of conversation attempts ( $M1=2.46$ ,  $M2=2.34$ ,  $t=1.7$ ,  $df=320$ ,  $p > .05$ ).

Observers apparently did not see such attempts at conversation as having a major influence on participant behavior. The mean response to "Your perception of the overall impact of 'observer effect' on participant's behavior" was a low 1.45 ( $SD=0.65$ ) on a scale from 1="No impact" to 5="Extreme impact.": The contrast between this result and results for the prior question suggest simple social politeness acts and brief attempts at conversation may not require participants to alter their activities or media exposure. No observers noted "Extreme impact" and moderate impact or higher were reported for only 58 sessions (7.2%).

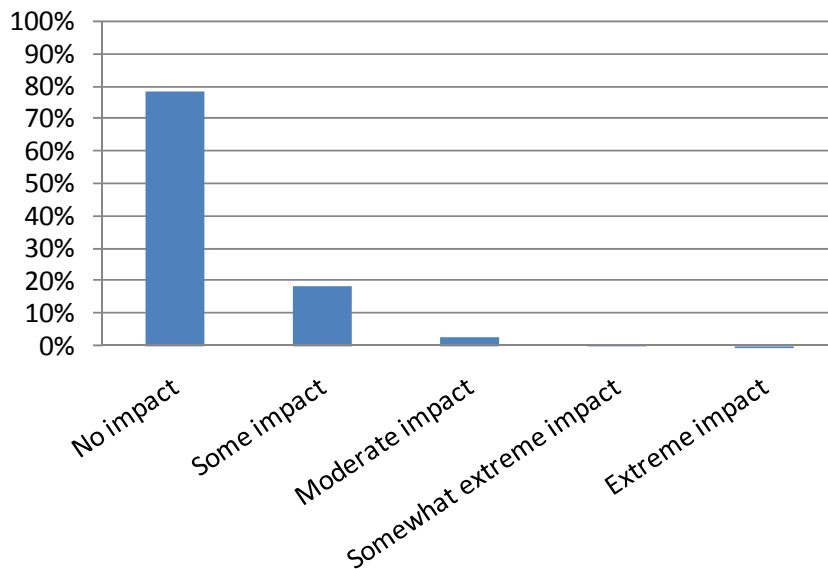
There was no difference between sessions for perceived impact of observer effect on participant general behavior ( $M1=1.46$ ,  $M2=1.45$ ,  $t=1.7$ ,  $df=320$ ,  $p > .05$ ).

Figure 3. Perception of observer effect on participant behavior



Results for the fourth question, "Your perception of the overall impact of 'observer effect' on the participant's media exposure during the session" also suggest any observer effects do not necessarily impact media exposure. For example, a deflected attempt at conversation while watching TV may not be associated with any discernable alteration in media platform or content choices. Observers consistently rated perceived impact on media exposure as low ( $M=1.26$ ,  $SD=.55$ ), reporting moderate or more severe impact for only 29 sessions (3.6%).

Figure 4. Perception of observer effect on media exposure



The final question prompted observers to describe, if the answer to the prior question was "3" (moderately comfortable) or greater, how observer effect was manifested during the session.

The small number of ratings in this range resulted in too few comments for in-depth analysis; however, the comments indicate several forms or consequences of observer effect. These include "social politeness" interactions (e.g., offering observers water, snacks or food; asking them for program preferences when the TV is turned on); persistent sociability (i.e., continued attempts to engage the observer in conversation despite reminders to "go about their day"); reported changes in routine (e.g., a participant remarking that he or she "would normally run errands this afternoon but is staying home instead"); and reported changes in media exposure (e.g., "I usually wouldn't read this much"). As noted, such manifestation reports are low in number and are associated with the small handful of observation sessions reported as having more than minimal observer effect. These are greatly outnumbered by observation sessions generally characterized as showing little or no observer effect.

---

## VCM Survey Instrument Sources and Acknowledgments

---

### Media Devices and Services Inventory

Developed by the Ball State University Center for Media Design for *Middletown Media Studies* (2004). Adapted for the CRE VCM study in collaboration with Sequent Partners and the CRE Media Consumption and Engagement committee.

### Big 5 Personality Survey

The “Big Five” personality model was independently developed in the 1970’s by Costa and McCrae (1994) and Goldberg (1994). According to this model, the five factors underlying personality are:

- Emotional Stability (how a person responds to stress);
- Extraversion (ability to tolerate stimulation from people and situations);
- Openness (to new experiences and new ways of doing things);
- Agreeableness (the degree to which a person defers to others);
- Conscientiousness (the degree to which a person pushes towards goals).

The cited researchers originally utilized the term “neuroticism” for the first factor above, but other scholars have substituted the positively-valenced “Emotional Stability” or “Need for Stability” rather than “neuroticism.”

Big Five instruments assign the respondent a value on a continuum for each factor, rather than assigning the respondent to a bimodal category. Unlike the well-known Myers-Briggs Type Indicator, for example, the emphasis is on an individual’s profile on these continuous scores rather than assigning the individual to a small set of pre-defined personality types.

The personality profile instrument used in the CRE VCM study is a short version developed and validated by Saucier (1994).

Costa, P. T., & McCrae, R. R. (1994). Set like plaster: Evidence for the stability of adult personality. In T. F. Heatherton & J. L. Weinberger (Eds.), *Can personality change?* (pp. 21-40). Washington, DC: American Psychological Association.

Goldberg, L.R. (1993). The structure of phenotypic personality traits. *American Psychologist*. January 1993, 48(1), 26-34.

Saucier, G. (1994). Mini-markers: A brief version of Goldberg’s unipolar big-five markers. *Journal of Personality Assessment*, 63(3), 506-516.

## **Innovativeness Profile**

Adapted for the CRE VCM study from a version provided by Bruce Goerlich of ZenithOptimedia (<http://www.zenithoptimedia.com>). Originally developed by Goldsmith and Hofacker:

Goldsmith, R.E., & Hofacker, C.F. (1991). Measuring consumer innovativeness. *Journal of the Academy of Marketing Science*, 19(3), 209-221.

## **Yankelovich Media Uses and Gratifications Battery**

Developed by Yankelovich, Inc. (<http://www.yankelovich.com>) for the "When Advertising Works" study. Used with permission.

## **Digital Transition Preparedness**

Developed by Ball State University Center for Media Design in collaboration with Sequent Partners and the Media Consumption and Engagement Committee. Upgrade path questions were adapted from the National Association of Broadcasters' recommended measures for the FCC's Wilmington, NC DTV experiment. (<http://www.dtvanswers.com/presskit/080606.html>)

## **Recall of Media Exposure**

Developed by the Ball State University Center for Media Design for *Middletown Media Studies* (2004). Question format adapted from media use recall questions applied in the biennial media consumption survey of the Pew Research Center for People and the Press.

## **Observer Session Report**

Developed by Ball State University Center for Media Design and revised for the CRE VCM study in collaboration with Sequent Partners and the Media Consumption and Engagement Committee.