The Design, Development and Use of Surveys in Healthcare Settings

Robert Lloyd, PhD
Katharine Luther, RN, MPM
Dave Williams, PhD

Faculty

Robert Lloyd, PhD,
Executive Director of Performance Improvement,
Institute for Healthcare Improvement

Katharine Luther, RN, MPM,
Vice President, north America,
Institute for Healthcare Improvement

David M. Williams, PhD,
IHI Improvement Advisor, Principal Truesimple Consulting
Discussion Topics

- How long will this survey madness go on?
- Why are you surveying?
- Key features of a survey
- Writing good questions
- Response formats and scales
- Data collection suggestions
- Analyzing and using survey results

How long will this survey madness go on?
“Swimming” in Surveys?

**Formal Surveys** (internal & external)

- Patients and families, safety climate/culture, physicians, staff
- Don’t know what to do with the results or what to do when a new problem appears
- And the answer is often: “Let’s do another survey!”

**Informal surveys frequently studied:**
Hand washing, handoff communication, medication availability, discharge process, clinic wait time, educational needs, feedback on programs, physician satisfaction with laboratory services,...and the list goes on!

And then...

“**Home-made**” surveys start to appear with:
- Poorly designed questions
- Confusing response scales
- Unclear instructions
- No follow-up

The “dots” remain disconnected!
Exercise #1
So, how many surveys does your organization try to manage?

- We would like you to initiate an inventory of surveys being conducted within your organization.
- Use the worksheet on the next page to start recording your findings. Please make additional copies of the worksheet if needed.

The four columns in the worksheet are as follows:
- Short title of the survey
- The target group for the survey (e.g., patients, family members, physicians, staff, managers)
- The number of subscales or domains in the survey (e.g., nursing care subscale, physician communication subscale or comfort and cleanliness subscale) and the total number of individual questions on the survey.
- Identify the frequency of each survey (e.g., is a one-time event or if it is repeated at specific periods of time).

---

Exercise #1
Survey Inventory

<table>
<thead>
<tr>
<th>Short Title of the Survey</th>
<th>Target Group for the survey</th>
<th>Number of subscales and questions on this survey</th>
<th>Frequency of Survey (one-time or repeat?)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Number of Subscales = Total questions on the survey =</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of Subscales = Total questions on the survey =</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of Subscales = Total questions on the survey =</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of Subscales = Total questions on the survey =</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of Subscales = Total questions on the survey =</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of Subscales = Total questions on the survey =</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of Subscales = Total questions on the survey =</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Number of Subscales = Total questions on the survey =</td>
<td></td>
</tr>
<tr>
<td>Short name of survey</td>
<td>Target group</td>
<td>Number of subscales and items on this survey</td>
<td>Frequency of survey (one-time or repeat?)</td>
</tr>
<tr>
<td>---------------------------</td>
<td>------------------</td>
<td>---------------------------------------------</td>
<td>------------------------------------------</td>
</tr>
<tr>
<td>Outpatient Services</td>
<td>Patients</td>
<td>Subscales = 19</td>
<td>78 one-time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Subscales = 8</td>
<td></td>
</tr>
<tr>
<td>Emergency Dept</td>
<td>Patients</td>
<td>Subscales = 45</td>
<td>45 one-time</td>
</tr>
<tr>
<td>Ambulatory Surgery</td>
<td>Patients</td>
<td>Subscales = 43</td>
<td>43 one-time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Subscales = 6</td>
<td></td>
</tr>
<tr>
<td>Inpatient Services (100)</td>
<td>Patients</td>
<td>Subscales = 10</td>
<td>10 one-time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Subscales = 8</td>
<td></td>
</tr>
<tr>
<td>Home grown</td>
<td>Subscales = 8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ED or small surg.</td>
<td>Patients</td>
<td>Subscales = 34</td>
<td>34 one-time</td>
</tr>
<tr>
<td>Home grown</td>
<td>Subscales = 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Support Services</td>
<td>Staff</td>
<td>Subscales = 5</td>
<td>5 to 7 twice per year</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Subscales = 3</td>
<td></td>
</tr>
<tr>
<td>Renal</td>
<td>Staff</td>
<td>Total items = 36</td>
<td>Once per year</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Subscales = 7</td>
<td></td>
</tr>
<tr>
<td>Culture of Safety</td>
<td>Staff</td>
<td>Total items = 42</td>
<td>Once per year</td>
</tr>
</tbody>
</table>

Source: Val Tobin - Service Excellence Manager, Providence Alaska Medical Center

---

<table>
<thead>
<tr>
<th>Type of Survey</th>
<th>Target Group</th>
<th>Number of subscales and items on this survey</th>
<th>Frequency of Survey (one time or repeat?)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AHRQ HSQPs Hospital Survey on Patient Safety</td>
<td>100% of direct care staff and physicians, sampling of others</td>
<td>Subscales = 12 composites Total items = 51</td>
<td>Every 1-2 years</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Note: 386 teaching hospitals participated in 2012 survey; 1128 hospitals overall</td>
</tr>
<tr>
<td>HCARES Hospital Consumer Assessment of Healthcare</td>
<td>Impatients post discharge</td>
<td>Subscales = 18 Total items = 27</td>
<td>Repeat</td>
</tr>
<tr>
<td>Providers and Systems</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Patient Experience Surveys (10)</td>
<td>Specific populations of patients post discharge or following treatment</td>
<td>Total items = 10-61</td>
<td>Variable</td>
</tr>
<tr>
<td>Staff Surveys (43)</td>
<td>Nurses and other clinicians</td>
<td>Total items = 6-40</td>
<td>Variable; most one time</td>
</tr>
<tr>
<td>Educational Evaluation surveys (&gt;100)</td>
<td>Course participants Grand Rounds</td>
<td>Total items =5-25</td>
<td>With each educational offering</td>
</tr>
</tbody>
</table>

Source: Carol Moore - Safety and Quality Education specialist, Christiana Care, Wilmington, DE
<table>
<thead>
<tr>
<th>Main Name of Survey</th>
<th>Target Group</th>
<th>Number of sections &amp; describes (questions) on survey</th>
<th>Frequency of survey</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radiology Providers</td>
<td>Providers w/ in Radiology</td>
<td>17 Items Questions (Radiology Domains)</td>
<td>Every 12 Months</td>
</tr>
<tr>
<td>Radiology Patient Experience</td>
<td>Radiology Pat Patients</td>
<td>18 Items Questions (Radiology Domains)</td>
<td>Quarterly (2nd quarter)</td>
</tr>
<tr>
<td>ICU Family Satisfaction</td>
<td>Families of ICU Patients</td>
<td>14 Items Questions</td>
<td>4 Subsections</td>
</tr>
<tr>
<td>Hospital Management (Patient Satisfaction)</td>
<td>Inpatient Patients</td>
<td>17 Items Questions, Subsections</td>
<td>Each class session</td>
</tr>
<tr>
<td>Gastroenterology</td>
<td>Gastroenterology Patients</td>
<td>17 Items Questions, Subsections</td>
<td>Each class session</td>
</tr>
<tr>
<td>Palliative Care Medicine</td>
<td>Patients, Staff</td>
<td>17 Items Questions, Subsections</td>
<td>Each class session</td>
</tr>
<tr>
<td>CNA Evaluation</td>
<td>CNA participants</td>
<td>17 Items Questions, Subsections</td>
<td>Each class session</td>
</tr>
<tr>
<td>Staff Nurses</td>
<td>Staff nurses</td>
<td>17 Items Questions, Subsections</td>
<td>Each class session</td>
</tr>
<tr>
<td>CNGS, TCO, QE, GI, Pediatric</td>
<td>Staff nurses</td>
<td>17 Items Questions, Subsections</td>
<td>Each class session</td>
</tr>
<tr>
<td>Green Hurdle</td>
<td>Staff of scoring &amp; recording nurse</td>
<td>17 Items Questions, Subsections</td>
<td>Yearly, may for breaches and another change on implementation</td>
</tr>
<tr>
<td>Infection Prevention</td>
<td>Nursing staff who are susceptible for new staff</td>
<td>17 Items Questions, Subsections</td>
<td>Yearly to current perceptions</td>
</tr>
<tr>
<td>Infection New Hire</td>
<td>New hire nursing staff</td>
<td>17 Items Questions, Subsections</td>
<td>Year each new hire orientation</td>
</tr>
</tbody>
</table>

Source: Renee Comeau – Patient Safety Analyst, North Shore Medical Center, Lynn, MA

---

**Driver Diagram of the Survey Process**

**Primary Drivers**
- Develop a clear purpose for the survey
- Write good questions & select appropriate response formats & scales
- Assess the validity & reliability of the survey items
- Data collection including stratification and sampling
- Analysis & interpretation of the data

**Secondary Drivers**
- Purpose and aim of survey is clearly defined
- Response format limits respondent burden
- Pilot testing is built into survey development process
- Developing specific, easy to understand questions (skills in writing questions)
- Testing the questions with the people who will actually take the survey
- Responses are consistent over time by the same people (test-retest)
- Different people can agree on the quality of the questions (inter-rater) At least 1-2 people agree the content in the survey is accurate (content)
- Different measures of the variables are consistent with each other (construct)
- Sampling method supports conclusions (probability vs. non-probability)
- Response rates, sampling error & bias are addressed in sampling scheme
- Stratification levels
- Plot the results on run & control charts
- Avoid the use of aggregated data & summary statistics
- Remember that processes change more quickly than outcomes

© 2013 Institute for Healthcare Improvement/R. Lloyd
Driver Diagram of the Survey Process

**Primary Drivers**

- Develop a clear purpose for the survey
- Write good questions & select appropriate response formats & scales
- Assess the validity & reliability of the survey items
- Data collection including stratification and sampling
- Analysis & interpretation of the data

**Secondary Drivers**

- Purpose and aim of survey is clearly defined
- Response format limits respondent burden
- Pilot testing is built into survey development process
- Developing specific, easy to understand questions (skills in writing questions)
- Testing the questions with the people who will actually take the survey
- Responses are consistent over time by the same people (test-retest)
- Different people can agree on the quality of the questions (inter-rater) At least 1-2 people agree the content in the survey is accurate (content)
- Different measures of the variables are consistent with each other (construct)
- Sampling method supports conclusions (probability vs. non-probability)
- Response rates, sampling error & bias are addressed in sampling scheme
- Stratification levels
- Plot the results on run & control charts
- Avoid the use of aggregated data & summary statistics
- Remember that processes change more quickly than outcomes

**Aim**

High quality survey data

A Survey Continuum: What defines your situation?

- Often grant funded
- Higher cost
- Involves experts
- Can take years to develop and test
- Larger samples
- Controlled Setting

- Limited resources
- Lower costs
- Involves non-experts
- Results needed quickly
- Smaller usually convenient samples
- Uncontrolled/chaotic setting

A tool for Academic Research and study

A tool for Improvement or Judgment
Surveying for Improvement
Linking the VOC and the VOP

Ideally you’d like the VOP to **EXCEED** the expectations of those you serve. At a minimum, however, you want the VOC and the VOP to be balanced.

Defining the Critical Points with the VOC

- **Pre-Service**
- **Point-of-Service**
- **Post-Service**

# Customer Feedback Tools & When to Use Them

<table>
<thead>
<tr>
<th>Tool/Approach</th>
<th>Pre</th>
<th>POS</th>
<th>Post</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surveys</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Focus Groups</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Observation</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Personal Interviews</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Unsolicited Feedback</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>High-Tech Tools</td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Mystery Shopper</td>
<td></td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>


# Exercise #2
Uses of Survey Data

<table>
<thead>
<tr>
<th>Possible Use</th>
<th>Frequency of Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>Making sure that the patient is central to the organization's purpose and strategic plan</td>
<td>Frequently</td>
</tr>
<tr>
<td>Assessing compliance with external requirements</td>
<td></td>
</tr>
<tr>
<td>Identifying customer requirements prior to service</td>
<td></td>
</tr>
<tr>
<td>Determining if our processes are efficient</td>
<td></td>
</tr>
<tr>
<td>Determining if we deliver service in a courteous manner</td>
<td></td>
</tr>
<tr>
<td>Developing pay for performance bonus systems, incentive payouts or other compensation strategies</td>
<td></td>
</tr>
<tr>
<td>Evaluating our current performance against comparative reference data or norms</td>
<td></td>
</tr>
<tr>
<td>Evaluating the effectiveness of improvement interventions</td>
<td></td>
</tr>
</tbody>
</table>

© 2013 Institute for Healthcare Improvement, R. Lloyd
What is a survey?

Key aspects of a survey

- It is a purposeful and systematic information collection system
- It focuses on a question or problem of interest which can be exploratory (what’s wrong?) or confirmatory (did we fix it?)
- It involves: question design, data collection, analysis, reporting
- It should have a degree of reliability and validity (reliable and valid in 1 context ≠ to all contexts)
- It creates interactions between the questioner and respondent and between the respondent and the instrument!
The three main components of a survey and related issues

**Questions (a.k.a. items)**
- What is the purpose of the questions (attitudes, opinions, behaviors, experience, events)?
- How many questions will be on the survey?

**Scales and Subscales** (Thurstone, Guttman, Likert, Wolins)
- Do the questions "hang together" to capture underlying concepts?
- Are you trying to make a single scale or does the survey have subscales?

**Response Format**
- Open-ended
- Closed-ended (a.k.a. forced-choice or checklist)
  - Nominal scale
  - Ordinal scale
  - Interval scale
  - Ratio scale

---

Typical sequence for developing and using a survey

**Step 1.** Develop the concept/plan (do I need subscales?)

**Step 2.** Create item pool (response variability is good)

**Step 3.** Pre-test and revise (length, complexity response format)

**Step 4.** Test the survey in a pilot setting

**Step 5.** Psychometric analysis (reliability-validity/subscales)

**Step 6.** “Final” instrument is ready for administration

**Step 7.** Measure at key points or continuously over time in conjunction with small tests of change!

**Step 8.** Analyze and interpret the data
But, be aware of Sources of Survey Error

Coverage error occurs when not all member of the population have a known, non-zero chance of being included in the sample and when those excluded are different from those who are included.

Sampling error occurs when the precision of the survey estimates is limited because not every person in the population is sampled. For example, a random sample of about 100 members of the US general public can produce margins of error of ±10% with a 95% confidence level (i.e., 95 out of 100 times the estimate will be within ±10%). But if you surveyed 2000 people the margin of error is reduced to about of ±2%.

Sources of Survey Error (continued)

Non-response error occurs when you do not get everyone in the sample to respond to the survey. This is a critical error when the non-respondents are different from those who do respond.

Measurement error occurs when a respondent’s answer to a survey item is inaccurate or imprecise. Measurement error is usually the result of poor question wording or construction. This is especially a problem with mailed or Internet surveys when there is no one from the surveying group around to explain the meaning of a question.
Example: Sources of Error

Survey Focus: Post Service

- Language and terminology are not clear
- Survey was not designed from the customer’s perspective
- Biased sampling plan
- Lack of skill in connecting the VOC with the VOP

Most surveys are designed to assess Post Service attitudes or experience

Did your results reflect respondent views of reality? Did you or anyone question the results?

Reactions to survey results:
- If results equal or better than benchmark = big celebration!
- If data matched expectations = no action & small celebration
- If undesirable results... blame the respondents or the caregiver.
And don’t forget the logistics!

<table>
<thead>
<tr>
<th>Who is going to...</th>
<th>Why it matters</th>
</tr>
</thead>
<tbody>
<tr>
<td>...identify actual target group(s)/service lines/physician groups etc.?</td>
<td>Need to keep groups straight for stratification during analysis phase; location codes and maps may need to be created and put on survey forms</td>
</tr>
<tr>
<td>...define the distribution process?</td>
<td>Do the potential respondents get the survey at beginning OR middle OR end of their experience?</td>
</tr>
<tr>
<td>...get the surveys back?</td>
<td>Without the data you are sunk!</td>
</tr>
<tr>
<td>...scan the results and/or enter them in a data repository?</td>
<td>Need central location for ongoing data retrieval and analysis</td>
</tr>
<tr>
<td>...do the data quality check?</td>
<td>Must ensure the responses can be processed (missing values, wrong units, weak tick marks, lost surveys)</td>
</tr>
<tr>
<td>...own the reporting process?</td>
<td>Need to close the loop and use results to drive improvement</td>
</tr>
</tbody>
</table>

Driver Diagram of the Survey Process

**Aim**

High quality survey data

**Primary Drivers**

- Develop a clear purpose for the survey
- Write good questions & select appropriate response formats & scales
- Assess the validity & reliability of the survey items
- Data collection including stratification and sampling
- Analysis & interpretation of the data

**Secondary Drivers**

- Purpose and aim of survey is clearly defined
- Response format limits respondent burden
- Pilot testing is built into survey development process
- Responses are consistent over time by the same people (test-retest)
- Different people can agree on the quality of the questions (inter-rater) At least 1-2 people agree the content in the survey is accurate (content)
- Different measures of the variables are consistent with each other (construct)
- Sampling method supports conclusions (probability vs. non-probability)
- Response rates, sampling error & bias are addressed in sampling scheme
- Stratification levels
- Plot the results on run & control charts
- Avoid the use of aggregated data & summary statistics
- Remember that processes change more quickly than outcomes

© 2013 Institute for Healthcare Improvement/R. Lloyd
**Writing Survey Questions**

“One must think about many things at once to write a good question and failure to do so can have significant effects on how the question performs.”

Dillman et al, 2009, page 67

**Factors involved in writing “good questions” include:**

- The type of question to write (open or closed-ended)
- How to word the question “stem” (main concept)
- What response options to offer and how to word them
- How to visually present the questionnaire
- What type of answer spaces will be provided (response format)
- Whether additional sources of information will be provided to the respondent

---

**Questions can focus on...**

- Attitudes
- Opinions
- Behaviors
- Experiences
- Events

Do you differentiate between these various concepts when you develop the questions for your surveys? Do you actually ask questions, make statements or list topics?
“Immaculate Perception” is the problem!

IF...

• Perception is framework or context dependent
• Information processing is not the same for all
• Humans are irrational

Then...

We can never assume a respondent will just “get” what we are asking!

Therefore...

• Be very careful with question design
• Pretest questions with a small sample of the target population

Guidelines for Choosing Words and Forming Questions
(Dillman, page 70-89)

1. Make sure the question applies to the respondent
2. Make sure the question is technically accurate
3. Ask one question at a time (no double-barreled questions)
4. Use simple and familiar words
5. Use specific and concrete words to specify the concepts clearly
6. Use as few words as possible to pose the question
7. Use complete sentences with simple sentence structures
8. Make sure “yes” means yes and “no” means no (avoid double negatives)
9. Be sure the question specifies the response task (don’t pull the old “bait and switch” trick)
**Guideline #3**

Ask one question at a time (no double-barreled questions!)

- Did we treat you with courtesy and respect?
- Did the nurses and the doctors respond to your needs and answer your questions?
- When you came to the clinic for your visit last week, did you or did you not think that the way you were treated demonstrated our corporate values of compassion and dignity?
- In the past 30 days, have you taken all the medicines we prescribed at time of discharge and had any unusual side effects?
- When you were ready to leave the clinic, did you receive full information about your co-pay and next appointment?
- As an employee of XYZ Medical Center, do you feel the benefits are reasonably priced and comprehensive?

---

**Guideline #4**

Use simple and familiar words

<table>
<thead>
<tr>
<th>Complex Word</th>
<th>Simple Word</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exhausted</td>
<td></td>
</tr>
<tr>
<td>Candid</td>
<td></td>
</tr>
<tr>
<td>Top priority</td>
<td></td>
</tr>
<tr>
<td>Leisure</td>
<td></td>
</tr>
<tr>
<td>Employment</td>
<td></td>
</tr>
<tr>
<td>Courageous</td>
<td></td>
</tr>
<tr>
<td>Rectify</td>
<td></td>
</tr>
</tbody>
</table>

Source: Dillman et al, page 83.
Guideline #6
Use as few words as possible to pose the question

<table>
<thead>
<tr>
<th>Many Words</th>
<th>Simple</th>
</tr>
</thead>
<tbody>
<tr>
<td>Due to the fact that</td>
<td></td>
</tr>
<tr>
<td>At this point in time</td>
<td></td>
</tr>
<tr>
<td>A small number of patients</td>
<td></td>
</tr>
<tr>
<td>A considerable number of patients</td>
<td></td>
</tr>
<tr>
<td>Have the ability to</td>
<td></td>
</tr>
<tr>
<td>Ascertain the location of</td>
<td></td>
</tr>
<tr>
<td>Concerning the matter of</td>
<td></td>
</tr>
<tr>
<td>If conditions are such that</td>
<td></td>
</tr>
<tr>
<td>In a majority of the diagnoses</td>
<td></td>
</tr>
<tr>
<td>Make a decision</td>
<td></td>
</tr>
<tr>
<td>Take into consideration</td>
<td></td>
</tr>
</tbody>
</table>

Adapted from Dillman et al, page 86.

Other examples of bad questions!

<table>
<thead>
<tr>
<th>Question</th>
<th>Guideline Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are you able to connect easily to the hospital’s Internet site?</td>
<td>Guideline #1: Make sure the question applies to the respondent (patient has no Internet)</td>
</tr>
<tr>
<td>Did you find the experience in our X-ray department acceptable?</td>
<td>Guideline #2: Make sure the question is technically accurate (the patient had a MRI scan)</td>
</tr>
<tr>
<td>How did things go?</td>
<td>Guideline #7: Use complete sentences with simple sentence structures</td>
</tr>
<tr>
<td>Do you favor or oppose not allowing the ICU to allow families in to see their loved one 24 hours a day?</td>
<td>Guideline #8: Make sure “yes” means yes and “no” means no (avoid double negatives)</td>
</tr>
<tr>
<td>How many days have you gone without smoking? (Response Options = Always or Most of the Time)</td>
<td>Guideline #9: Be sure the question specifies the response task (avoid the bait and switch problem)</td>
</tr>
</tbody>
</table>
Exercise #3
Writing Survey Questions

From the list below, pick 1-2 topics and write 4-6 questions aimed at assessing satisfaction with:

- Your company’s new health benefits plan
- The outpatient radiology testing process
- The registration process in the Emergency Department
- The cafeteria’s new “healthy” menu
- Your family’s recent vacation
- The used car you just bought
- Performance of local politicians

If none of these topics appeals to you feel free to select one of your own.

Now that you have developed appropriate questions for your survey, what response formats will you use?
Response Formats, Items and Scales
(see Appendix B for additional details)

• There are many different types of **response formats** (binary, visual analog, semantic differential, number range).

• There are many different types of **scales** (Thurstone, Guttman, Likert, Wolins) which are made up of **items** which “hang together.”

• Scales and response formats are not synonymous!

• The focus in healthcare is usually on response formats and items not on true scales!

---

Response Format Options

**Open-ended questions** – provide a blank space or a box where the respondents type or write in their responses using their own words.

**Closed-ended questions (forced-choice)**

- Nominal scalar format (list a bunch of response options that have no rank order or directionality)

- Ordinal scalar format (list response options that follow a progressive sequence or continuum)

- Partially closed question format “other” is added as an option (use an Ordinal Scale but add “other” at the end of the scale choices)

- Visual analog scalar (select a response on a progressive scale)
Most surveys in healthcare settings are usually here

- **Nominal** – attributes only (naming) – weakest approach
- **Ordinal** – attributes can be rank ordered (1<sup>st</sup>, 2<sup>nd</sup> or 3<sup>rd</sup>) but there are no equal appearing intervals
- **Interval** – equal appearing intervals between the numbers
- **Ratio** – absolute zero, highest level

It all starts with the Level of Measurement (scales) you select
(because, all measurement is not the same)
## Measurement Scales and Response Formats

<table>
<thead>
<tr>
<th>Scale</th>
<th>Response Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal (naming or categorical)</td>
<td>1 = male, 2 = female; A = under 60, B = 60 or older; * = no prior admission, ** = prior admission &lt;30 days</td>
</tr>
<tr>
<td>Ordinal (rank order)</td>
<td>1 = strongly agree, 2 = agree, 3 = undecided, 4 = disagree, 5 = strongly disagree</td>
</tr>
<tr>
<td>Interval (equal distance between markers)</td>
<td>True interval scales do not exist in survey research (an example however is temperature in Fahrenheit or Centigrade)</td>
</tr>
<tr>
<td></td>
<td><strong>NOTE:</strong> Ordinal measurement is frequently transformed into a “quasi-interval” scale which is a controversial transformation within the survey research literature.</td>
</tr>
<tr>
<td>Ratio (equal distance between markers plus a true zero point)</td>
<td>True ratio scales do not exist in survey research. They occur frequently in physics or when measuring weight, time, length, area, angles, etc. The cardinal numbering system is based on a ratio scale.</td>
</tr>
</tbody>
</table>

## Measurement Scales and Statistical Options

<table>
<thead>
<tr>
<th>Scale</th>
<th>Permissible Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nominal (naming or categorical)</td>
<td>number of cases, mode, percent contingency analysis (e.g., chi-square)</td>
</tr>
<tr>
<td>Ordinal (rank order but no equal appearing intervals between ranks)</td>
<td>median, mode, percentiles, quartiles and non-parametric statistical analysis</td>
</tr>
<tr>
<td>Interval (equal distance between markers on a scale)</td>
<td>mean, standard deviation, rank-order correlation, product-moment correlation, ANOVA, regression, factor, discriminant, etc.</td>
</tr>
<tr>
<td>Ratio (equal distance between markers plus a true zero point)</td>
<td>all statistics for interval plus coefficient of variation, logarithms, geometric and harmonic means</td>
</tr>
</tbody>
</table>

*While the outlawing the use of means with ordinal data would probably serve no good purpose, it is proper to point out that means and standard deviations computed on an ordinal scale are in error to the extent that the successive intervals on the scale are unequal in size.*

You Make the Call! (Good, Bad or Ugly?)

Instructions: Dear ICU nurse thank you for participating in the ICU MRSA reduction project two months ago. Please answer the questions below so that we can assess the quality and effectiveness of the project. Return the survey within 5 days. Thanks.

Your name:_____________________ Today’s date:_________ 

Q1. In general, I thought the MRSA project was?
   A.  the worst project I have ever been a part of.
   B.  somewhere between the worst and best project I have been a part of.
   C.  the best project I have ever been a part of.

Q2. In general, I felt the nurses and physicians consistently followed the MRSA reduction protocol:
   A. strongly disagree    B. disagree    C. neutral    D.  agree    E. strongly agree

Q3. During the project, I felt as if I had adequate access to the Infection Control team when questions arose about the protocol and their answers were useful.
   A. adequate access   B. undecided   C. not bad access   D. no access

Q4. During the project period, the microbiology laboratory took too long to get culture results back to the unit?
   A. time was ok  B. undecided   C. took too long   D. took way too long

Q5. The log book for recording which patients received MRSA screens on admission was effective and easy to locate?
   A. Agree  B. Disagree

Q6. All the patients found value in the screening process?
   A. strongly disagree  B. disagree   C. undecided   D. agree  E. strongly agree

Driver Diagram of the Survey Process

Primary Drivers
- Develop a clear purpose for the survey
- Write good questions & select appropriate response formats & scales
- Assess the validity & reliability of the survey items
- Data collection including stratification and sampling
- Analysis & interpretation of the data

Secondary Drivers
- Purpose and aim of survey is clearly defined
- Response format limits respondent burden
- Pilot testing is built into survey development process
- Responses are consistent over time by the same people (test-retest)
- Different people can agree on the quality of the questions (inter-rater) At least 1-2 people agree the content in the survey is accurate (content)
- Different measures of the variables are consistent with each other (construct)
- Sampling method supports conclusions (probability vs. non-probability)
- Response rates, sampling error & bias are addressed in sampling scheme
- Stratification levels
- Plot the results on run & control charts
- Avoid the use of aggregated data & summary statistics
- Remember that processes change more quickly than outcomes
Selecting an appropriate data collection method

<table>
<thead>
<tr>
<th>Benefits</th>
<th>Drawbacks</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mail</strong></td>
<td>• Inexpensive</td>
</tr>
<tr>
<td>• Inexpensive</td>
<td>• Slower turnaround time</td>
</tr>
<tr>
<td>• No interviewer bias</td>
<td>• Possible low response rates</td>
</tr>
<tr>
<td>• Can be anonymous</td>
<td>• May miss certain segments of the population</td>
</tr>
<tr>
<td><strong>Phone Surveys</strong></td>
<td>• Expensive</td>
</tr>
<tr>
<td>• Fast turnaround</td>
<td>• Interview skill and bias</td>
</tr>
<tr>
<td>• Personal touch</td>
<td>• Unlisted cell phone numbers</td>
</tr>
<tr>
<td><strong>Personally Distributed Questionnaires Onsite or in Office</strong></td>
<td>• Not anonymous</td>
</tr>
<tr>
<td>• Inexpensive</td>
<td>• Leads to positive response bias</td>
</tr>
<tr>
<td>• Easy to administer</td>
<td>• Can have a high sampling error</td>
</tr>
</tbody>
</table>


Key Terms and Definitions
(adapted from Dillman, pages 42-43)

• **Survey Population** – all units (patients family members, employees) to which your desire to generalize the results

• **Sample Frame** – the list from which a sample is to be drawn

• **Sample** – all units of the population that are drawn for inclusion in the survey

• **Completed Sample** – all the units that complete the survey

• **Coverage Error** – results from every unit in the survey population not having a known, nonzero chance of being included in the sample

• **Sampling Error** – the result of collecting data from only a subset, rather than all, of the members of the sampling frame.
How this sampling stuff works

Survey Population

Sample Frame

Sample

Completed Sample

Related to the Completed Samples are the two issues of:
- Coverage Error
- Sampling Error

All inpatients (n = 8,200)

All inpatients with verified mailing addresses

A random sample of inpatients with verified mailing addresses (n = 1,200)

All surveyed inpatients who returned a completed survey (n = 681 or 67%)
Sampling Methods


**Probability Sampling Methods**
- Simple random sampling
- Stratified random sampling
- Stratified proportional random sampling
- Systematic sampling
- Cluster sampling

**Non-probability Sampling Methods**
- Convenience sampling
- Quota sampling
- Judgment sampling

---

**Driver Diagram of the Survey Process**

- **Primary Drivers**
  - Develop a clear purpose for the survey
  - Write good questions & select appropriate response formats & scales
  - Assess the validity & reliability of the survey items
  - Data collection including stratification and sampling
  - Analysis & interpretation of the data

- **Secondary Drivers**
  - Purpose and aim of survey is clearly defined
  - Response format limits respondent burden
  - Pilot testing is built into survey development process
  - Developing specific, easy to understand questions (skills in writing questions)
  - Testing the questions with the people who will actually take the survey
  - Responses are consistent over time by the same people (test-retest)
  - Different people can agree on the quality of the questions (inter-rater)
  - At least 1-2 people agree the content in the survey is accurate (content)
  - Different measures of the variables are consistent with each other (construct)
  - Sampling method supports conclusions (probability vs. non-probability)
  - Response rates, sampling error & bias are addressed in sampling scheme
  - Stratification levels
  - Plot the results on run & control charts
  - Avoid the use of aggregated data & summary statistics
  - Remember that processes change more quickly than outcomes
So, you have collected VOC data. Now what do you do with it?

Analyzing and displaying survey data

Key Questions

• What type of data did you collect? Nominal, ordinal, interval or ratio? Do you have percentages (e.g., top box), a scores, percentiles or a composite index?

• Do you plan to analyze individual survey items (questions) or do you plan to create subscale scores?

• Will you present the data in tabular form (counts and amounts with summary statistics) or graphically?

• If you use graphics, will you compare two data points or look at a series of data points over time?
The Problem

Aggregated data presented in tabular formats or with summary statistics, will not help you measure the impact of process improvement efforts.

Aggregated data can only lead to judgment, not to improvement.

A Static Display of Survey Data


This year the average percent of patients marking “Excellent” on the survey (about 60%) was roughly the same as the average for last year. But we’ve worked so hard on improving customer service. How is this possible?
Many organizations choose to take a static approach to analyzing survey data. Typically bar graphs or pie charts are used to report the data on a quarterly basis.

Here we have 3 subscales which make up the Clinical Care dimension on a mailed survey (the bar graphs) and the overall quality of care question on a pie chart.

But, looking at data over time can give you a very different picture.

If we look at the data by month, however, we get a very different picture. Now what do you conclude?

Last year by month

This year by month
The average of a set of numbers can be created by many different distributions.

If you don’t understand the variation that lives in your data, you will be tempted to:

- Deny the data (It doesn’t fit my view of reality!)
- See trends where there are no trends
- Try to explain natural variation as special events
- Blame and give credit to people for things over which they have no control
- Distort the process that produced the data
- Hope the results improve (hope is NOT a plan!)
- Kill the messenger!
"A phenomenon will be said to be controlled when, through the use of past experience, we can predict, at least within limits, how the phenomenon may be expected to vary in the future" - Dr. Walter A Shewhart

"What is the variation in one system over time?" - Walter A. Shewhart - early 1920's, Bell Laboratories

Every process displays variation:
- **Common Cause variation**
  - Stable, consistent pattern of variation
  - Random variation in the process
- **Special Cause variation**
  - Unstable and not predictable variation
  - Requires investigation to find "assignable" causes
Control charts applied to survey data

The Key Questions

Is your VOC data:

1. Stable?
2. Predictable?
3. Capable?

A run or control chart will tell you if the process is stable and predictable.
You have to decide if the performance of the process is capable!
**Primary Drivers**
- Develop a clear purpose for the survey
- Write good questions & select appropriate response formats & scales
- Assess the validity & reliability of the survey items
- Data collection including stratification and sampling
- Analysis & interpretation of the data

**Secondary Drivers**
- Purpose and aim of survey is clearly defined
- Response format limits respondent burden
- Pilot testing is built into survey development process
- Developing specific, easy to understand questions (skills in writing questions)
- Testing the questions with the people who will actually take the survey
- Responses are consistent over time by the same people (test-retest)
- Different people can agree on the quality of the questions (inter-rater) At least 1-2 people agree the content in the survey is accurate (content)
- Different measures of the variables are consistent with each other (construct)
- Sampling method supports conclusions (probability vs. non-probability)
- Response rates, sampling error & bias are addressed in sampling scheme
- Stratification levels
- Plot the results on run & control charts
- Avoid the use of aggregated data & summary statistics
- Remember that processes change more quickly than outcomes

**Aim**
High quality survey data

---

**The Survey Checklist**

<table>
<thead>
<tr>
<th>Status</th>
<th>In Progress</th>
<th>Complete</th>
<th>N/A</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Getting Started</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Determine Purpose, Scope and Aim</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Determine Available Resources</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decide if Outside Vendor is Needed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decide on Paper, Electronic, Mail, Phone</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Determine Frequency of Survey (one time, ongoing)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Form a Team</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Create the Survey (if not using existing survey)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Determine # of Domains and Questions</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Determine Response Format</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assess for Reliability and Validity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do Small Pilot Test</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Selecting A Sample</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decide Who To Survey (Population)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Determine When to Survey (Pre-service, Point of Service)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decide How to Sample (Random, Purposive, Judgment)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Determine Sample Size</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Data Collection Method</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decide How Surveys Will Be Administered and Returned</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Establish Points of Contact (Clinics, Hospital, Community)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decide Where Data Will Be Stored</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Data Collection Procedure</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Decide Whether to Track Responses via Identifiers</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Send Pre-notification Letter (for large surveys)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assemble Survey Materials</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Data Analysis</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Create Graphic Display of Data</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Report Data To Appropriate Stakeholders</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

© 2013 Institute for Healthcare Improvement, R. Lloyd
Making Improvements

S + P + C* = O

Structure + Process + Culture = Outcomes

*R. Lloyd and R. Scoville added the “C” component to Donabedian’s original equation to highlight the important role that culture plays in driving change and improvement.


Relating the VOC to the VOP

A few final thoughts from the faculty

Don't forget to be clear about why you are surveying and also realize that there are many ways to listen to the VOC.

Be practical, be focused and make sure you are keeping the patient at the center of your work!

Make sure you are asking questions about reliability, validity, question construction and response formats.

Appendices

• Appendix A: Recommended Reading
• Appendix B: A note on Likert and his scales
• Appendix C: Faculty Bios
Appendix A: Recommended Reading


Appendix B

A note on Likert and his scales

Rensis Likert, (1903–1981) was an American educator and organizational psychologist best known for his research on management styles.

He is also famous for developing the Likert Scale, a psychometric scale commonly involved in research using questionnaires.

Likert was known for his support of interdisciplinary collaborations and emphasis on using social science research to effect positive change.
Characteristics of a Likert Scale

- Uni-dimensional, summative scale
- Scale contains several items (has to have >1 item)
- Response levels are:
  - arranged horizontally
  - anchored with consecutive integers
  - anchored with verbal labels that connote evenly-spaced gradients
- Verbal labels are:
  - bi-valent (two positive and two negative response options)
  - symmetrical around a neutral middle
- Scales always measures attitude in terms of level of agreement/disagreement of several target statements that can be aggregated to provide a respondent’s position on a concept or topic.

Be clear on the differences between a “Likert item” and a “Likert scale”

- An important distinction must be made between a Likert item and a Likert scale.
- A Likert scale is the sum of responses on several Likert items (a subscale or factor).
- Because Likert items are often accompanied by a visual analog scale, the items are sometimes called scales themselves, which is incorrect.
- This is the source of much confusion; it is better, therefore, to reserve the term “Likert scale” to apply to the summed scale, and “Likert item or Likert-type format” to refer to an individual question or item on a survey.
A “Likert-like” Response Format

If you use a 3, 5 or even 7 category response format, this does not necessarily mean that you have a Likert item or more importantly a Likert Scale.

What you have decided to do is to use a rank ordered response format; that is all!

Appendix C: Faculty Bio

Robert Lloyd, PhD is Executive Director of Performance Improvement for the Institute for Healthcare Improvement (IHI). Dr. Lloyd provides leadership in the areas of performance improvement strategies, statistical process control methods, development of strategic dashboards and building capacity and capability for quality improvement. He also serves as lead faculty for various IHI initiatives and demonstration projects in the US and abroad. Before joining the IHI, Dr. Lloyd served as the Corporate Director of Quality Resource Services for Advocate Health Care (Oak Brook, IL). He also served as Senior Director of Quality Measurement for Lutheran General Health System (Park Ridge, IL), directed the American Hospital Association’s Quality Measurement and Management Project (QMMP) and served in various leadership roles at the Hospital Association of Pennsylvania. The Pennsylvania State University awarded all three of Dr. Lloyd’s degrees. His doctorate is in agricultural economics and rural sociology. Dr. Lloyd has written many articles and chapters in books. He is also the co-author of the internationally acclaimed book, Measuring Quality Improvement in Healthcare: A Guide to Statistical Process Control Applications (American Society for Quality Press, 2001, 5th printing) and the author of Quality Health Care: A Guide to Developing and Using Indicators, 2004 by Jones and Bartlett (Sudbury, MA).
Appendix C: Faculty Bio

Katharine Luther, RN, MPM, Vice President, North America, Institute for Healthcare Improvement (IHI), is responsible for furthering IHI’s work to help hospital leaders and staff achieve bold aims. Key to this work is developing strategic partnerships that leverage innovation, pilot testing, implementation, and continuous learning across organizations, systems, professional societies, and entire countries. Previously, she served as Executive Director at IHI, designing new programs to impact cost and health care quality. Ms. Luther has over 25 years of experience in clinical and process improvement, focusing on large-scale change projects and program development, system improvement, rapid cycle change, developing and managing a portfolio of projects, and working with all levels of health care staff and leaders. Her clinical experience includes critical care, emergency room, trauma, and psychiatry. Prior to joining IHI, she held leadership positions at the University of Pittsburgh Medical Center, MD Anderson Cancer Center, and Memorial Hermann–Texas Medical Center. She has experience in Lean and is a Six Sigma Master Black Belt.

Appendix C: Faculty Bio

David M. Williams, PhD, is an Improvement Advisor. He founded and leads a consulting practice based in Austin, Texas, called TrueSimple Improvement. He advises education, emergency medical service (EMS), and healthcare organizations in the US and Europe on applying the science of improvement and integrating quality into business strategy. Dr. Williams started his career as a volunteer firefighter in high school, worked as a paramedic in several cities, and is a subject matter expert on emergency medical services systems. He earned a Bachelor’s degree in EMS Management, a Master’s degree in Emergency Health Services Management, and a PhD in Organizational Systems with a research focus on the obstacles to patient-centric EMS system design.