Topics in Measurement for the Social and Health Sciences

Gary King
Institute for Quantitative Social Science
Harvard University

(talk in "Foundations in Global Health" class, HSPH, 9/16/11)

1. Survey Research

1. Survey Research

2. Aggregate Government (& other) Statistics

The Evidence Base of Social Science: The Last 50 Years

1. Survey Research

2. Aggregate Government (& other) Statistics

3. One-off studies of individual places, people, or events

The Evidence Base of Social Science: The Last 50 Years

1. **Survey Research**
   - Occasional snapshots, of random selections, of isolated individuals, from unknown geographic locations

2. **Aggregate Government (& other) Statistics**

3. **One-off studies of individual places, people, or events**

The Evidence Base of Social Science: The Last 50 Years

1. Survey Research
   - Occasional snapshots, of random selections, of isolated individuals, from unknown geographic locations
   - Interpersonal incomparability, "non-opinions," Hawthorne effects, no direct connection to behavior

2. Aggregate Government (& other) Statistics

3. One-off studies of individual places, people, or events

The Evidence Base of Social Science: The Last 50 Years

1. **Survey Research**
   - Occasional snapshots, of random selections, of isolated individuals, from unknown geographic locations
   - Interpersonal incomparability, "non-opinions," Hawthorne effects, no direct connection to behavior
   - Crumbling foundation: overuse & cell phones $\leadsto$ 90+% nonresponse

2. **Aggregate Government (& other) Statistics**

3. **One-off studies of individual places, people, or events**

The Evidence Base of Social Science: The Last 50 Years

1. **Survey Research**
   - Occasional snapshots, of random selections, of isolated individuals, from unknown geographic locations
   - Interpersonal incomparability, "non-opinions," Hawthorne effects, no direct connection to behavior
   - Crumbling foundation: overuse & cell phones $\rightsquigarrow$ 90+% nonresponse
   - Web survey opportunity: marginal cost ≈ 0, but what about selection?
2. **Aggregate Government (& other) Statistics**

3. **One-off studies of individual places, people, or events**

The Evidence Base of Social Science: The Last 50 Years

1. **Survey Research**
   - Occasional snapshots, of random selections, of isolated individuals, from unknown geographic locations
   - Interpersonal incomparability, "non-opinions," Hawthorne effects, no direct connection to behavior
   - Crumbling foundation: overuse & cell phones $\leadsto$ 90+% nonresponse
   - Web survey opportunity: marginal cost ≈ 0, but what about selection?
2. **Aggregate Government (& other) Statistics**
   - Highly aggregated over time and space

3. **One-off studies of individual places, people, or events**

The Evidence Base of Social Science: The Last 50 Years

1. **Survey Research**
   - Occasional snapshots, of random selections, of isolated individuals, from unknown geographic locations
   - Interpersonal incomparability, "non-opinions," Hawthorne effects, no direct connection to behavior
   - Crumbling foundation: overuse & cell phones $\rightsquigarrow$ 90+% nonresponse
   - Web survey opportunity: marginal cost $\approx$ 0, but what about selection?

2. **Aggregate Government (& other) Statistics**
   - Highly aggregated over time and space
   - Individuals not identified

3. **One-off studies of individual places, people, or events**

The Evidence Base of Social Science: The Last 50 Years

1. **Survey Research**
   - Occasional snapshots, of random selections, of isolated individuals, from unknown geographic locations
   - Interpersonal incomparability, "non-opinions," Hawthorne effects, no direct connection to behavior
   - Crumbling foundation: overuse & cell phones $\rightsquigarrow$ 90+% nonresponse
   - Web survey opportunity: marginal cost ≈ 0, but what about selection?

2. **Aggregate Government (& other) Statistics**
   - Highly aggregated over time and space
   - Individuals not identified
   - No investigator control

3. **One-off studies of individual places, people, or events**

The Evidence Base of Social Science: The Last 50 Years

1. **Survey Research**
   - Occasional snapshots, of random selections, of isolated individuals, from unknown geographic locations
   - Interpersonal incomparability, "non-opinions," Hawthorne effects, no direct connection to behavior
   - Crumbling foundation: overuse & cell phones $\rightsquigarrow$ 90+% nonresponse
   - Web survey opportunity: marginal cost ≈ 0, but what about selection?
2. **Aggregate Government (& other) Statistics**
   - Highly aggregated over time and space
   - Individuals not identified
   - No investigator control
   - Dubious impartiality: Governments, newspapers, NGOs, etc.
3. **One-off studies of individual places, people, or events**

The Evidence Base of Social Science: The Last 50 Years

1. **Survey Research**
   - Occasional snapshots, of random selections, of isolated individuals, from unknown geographic locations
   - Interpersonal incomparability, "non-opinions," Hawthorne effects, no direct connection to behavior
   - Crumbling foundation: overuse & cell phones $\rightsquigarrow$ 90+% nonresponse
   - Web survey opportunity: marginal cost ≈ 0, but what about selection?

2. **Aggregate Government (& other) Statistics**
   - Highly aggregated over time and space
   - Individuals not identified
   - No investigator control
   - Dubious impartiality: Governments, newspapers, NGOs, etc.

3. **One-off studies of individual places, people, or events**
   - do not scale

The Evidence Base of Social Science: The Last 50 Years

1. **Survey Research**
   - Occasional snapshots, of random selections, of isolated individuals, from unknown geographic locations
   - Interpersonal incomparability, "non-opinions," Hawthorne effects, no direct connection to behavior
   - Crumbling foundation: overuse & cell phones ⤳ 90+% nonresponse
   - Web survey opportunity: marginal cost ≈ 0, but what about selection?

2. **Aggregate Government (& other) Statistics**
   - Highly aggregated over time and space
   - Individuals not identified
   - No investigator control
   - Dubious impartiality: Governments, newspapers, NGOs, etc.

3. **One-off studies of individual places, people, or events**
   - do not scale
   - are not representative

The Evidence Base of Social Science: The Last 50 Years

1. **Survey Research**
   - Occasional snapshots, of random selections, of isolated individuals, from unknown geographic locations
   - Interpersonal incomparability, "non-opinions," Hawthorne effects, no direct connection to behavior
   - Crumbling foundation: overuse & cell phones $\rightsquigarrow$ 90+% nonresponse
   - Web survey opportunity: marginal cost ≈ 0, but what about selection?

2. **Aggregate Government (& other) Statistics**
   - Highly aggregated over time and space
   - Individuals not identified
   - No investigator control
   - Dubious impartiality: Governments, newspapers, NGOs, etc.

3. **One-off studies of individual places, people, or events**
   - do not scale
   - are not representative
   - do not measure long-term change.

The Last 50 Years:

The Changing Evidence Base of Social Science Research

The Last 50 Years:

- Survey research

The Changing Evidence Base of Social Science Research

The Last 50 Years:

- Survey research
- Aggregate government statistics

The Changing Evidence Base of Social Science Research

The Last 50 Years:

- Survey research
- Aggregate government statistics
- In depth studies of individual places, people, or events

The Changing Evidence Base of Social Science Research

The Last 50 Years:

- Survey research
- Aggregate government statistics
- In depth studies of individual places, people, or events

The Next 50 Years: Spectacular increases in new data sources, due to...

The Changing Evidence Base of Social Science Research

The Last 50 Years:

- Survey research
- Aggregate government statistics
- In depth studies of individual places, people, or events

The Next 50 Years: Spectacular increases in new data sources, due to. . .

- Much more of the above — improved, expanded, and applied

The Changing Evidence Base of Social Science Research

The Last 50 Years:

- Survey research
- Aggregate government statistics
- In depth studies of individual places, people, or events

The Next 50 Years: Spectacular increases in new data sources, due to...

- Much more of the above — improved, expanded, and applied
- Shrinking computers & the growing Internet: data everywhere

The Changing Evidence Base of Social Science Research

The Last 50 Years:

- Survey research
- Aggregate government statistics
- In depth studies of individual places, people, or events

The Next 50 Years: Spectacular increases in new data sources, due to...

- Much more of the above — improved, expanded, and applied
- Shrinking computers & the growing Internet: data everywhere
- The replication movement: academic data sharing (e.g., Dataverse)

The Changing Evidence Base of Social Science Research

The Last 50 Years:

- Survey research
- Aggregate government statistics
- In depth studies of individual places, people, or events

The Next 50 Years: Spectacular increases in new data sources, due to...

- Much more of the above — improved, expanded, and applied
- Shrinking computers & the growing Internet: data everywhere
- The replication movement: academic data sharing (e.g., Dataverse)
- Governments encouraging data collection, distribution, experimentation (e.g., GovData)

The Changing Evidence Base of Social Science Research

The Last 50 Years:

- Survey research
- Aggregate government statistics
- In depth studies of individual places, people, or events

The Next 50 Years: Spectacular increases in new data sources, due to...

- Much more of the above — improved, expanded, and applied
- Shrinking computers & the growing Internet: data everywhere
- The replication movement: academic data sharing (e.g., Dataverse)
- Governments encouraging data collection, distribution, experimentation (e.g., GovData)
- Advances in statistical methods, informatics, & software

The Changing Evidence Base of Social Science Research

The Last 50 Years:

- Survey research
- Aggregate government statistics
- In depth studies of individual places, people, or events

The Next 50 Years: Spectacular increases in new data sources, due to. . .

- Much more of the above — improved, expanded, and applied
- Shrinking computers & the growing Internet: data everywhere
- The replication movement: academic data sharing (e.g., Dataverse)
- Governments encouraging data collection, distribution, experimentation (e.g., GovData)
- Advances in statistical methods, informatics, & software
- *The march of quantification*: through academia, professions, government, & commerce (*SuperCrunchers*, *The Numerati*, *MoneyBall*)

Examples of what's now possible

- Opinions of activists:

Examples of what's now possible

- Opinions of activists: A few thousand interviews

Examples of what's now possible

- Opinions of activists: A few thousand interviews ⤳ millions of political opinions in social media posts (1B tweets/5 days)

Examples of what's now possible

- Opinions of activists: A few thousand interviews ⤳ millions of political opinions in social media posts (1B tweets/5 days)
- Exercise:

Examples of what's now possible

- Opinions of activists: A few thousand interviews ⤳ millions of political opinions in social media posts (1B tweets/5 days)
- Exercise: A survey: "How many times did you exercise last week?

Examples of what's now possible

- Opinions of activists: A few thousand interviews ⤳ millions of political opinions in social media posts (1B tweets/5 days)
- Exercise: A survey: "How many times did you exercise last week? ⤳ 500K people carrying cell phones with accelerometers

Examples of what's now possible

- Opinions of activists: A few thousand interviews ⇝ millions of political opinions in social media posts (1B tweets/5 days)
- Exercise: A survey: "How many times did you exercise last week? ⇝ 500K people carrying cell phones with accelerometers
- Social contacts:

Examples of what's now possible

- Opinions of activists: A few thousand interviews $\rightsquigarrow$ millions of political opinions in social media posts (1B tweets/5 days)
- Exercise: A survey: "How many times did you exercise last week? $\rightsquigarrow$ 500K people carrying cell phones with accelerometers
- Social contacts: A survey: "Please tell me your 5 best friends"

Examples of what's now possible

- Opinions of activists: A few thousand interviews ⤳ millions of political opinions in social media posts (1B tweets/5 days)
- Exercise: A survey: "How many times did you exercise last week? ⤳ 500K people carrying cell phones with accelerometers
- Social contacts: A survey: "Please tell me your 5 best friends" ⤳ continuous record of phone calls, emails, text messages, bluetooth, social media connections, electronic address books

Examples of what's now possible

- Opinions of activists: A few thousand interviews ⤳ millions of political opinions in social media posts (1B tweets/5 days)
- Exercise: A survey: "How many times did you exercise last week? ⤳ 500K people carrying cell phones with accelerometers
- Social contacts: A survey: "Please tell me your 5 best friends" ⤳ continuous record of phone calls, emails, text messages, bluetooth, social media connections, electronic address books
- Economic development in developing countries:

Examples of what's now possible

- **Opinions of activists:** A few thousand interviews ⇝ millions of political opinions in social media posts (1B tweets/5 days)

- **Exercise:** A survey: "How many times did you exercise last week? ⇝ 500K people carrying cell phones with accelerometers

- **Social contacts:** A survey: "Please tell me your 5 best friends" ⇝ continuous record of phone calls, emails, text messages, bluetooth, social media connections, electronic address books

- **Economic development in developing countries:** Dubious or nonexistent governmental statistics

Examples of what's now possible

- Opinions of activists: A few thousand interviews ⤳ millions of political opinions in social media posts (1B tweets/5 days)
- Exercise: A survey: "How many times did you exercise last week? ⤳ 500K people carrying cell phones with accelerometers
- Social contacts: A survey: "Please tell me your 5 best friends" ⤳ continuous record of phone calls, emails, text messages, bluetooth, social media connections, electronic address books
- Economic development in developing countries: Dubious or nonexistent governmental statistics ⤳ satellite images of human-generated light at night, or networks of roads and other infrastructure

Examples of what's now possible

- **Opinions of activists:** A few thousand interviews ⇝ millions of political opinions in social media posts (1B tweets/5 days)

- **Exercise:** A survey: "How many times did you exercise last week? ⇝ 500K people carrying cell phones with accelerometers

- **Social contacts:** A survey: "Please tell me your 5 best friends" ⇝ continuous record of phone calls, emails, text messages, bluetooth, social media connections, electronic address books

- **Economic development in developing countries:** Dubious or nonexistent governmental statistics ⇝ satellite images of human-generated light at night, or networks of roads and other infrastructure

- **Expert-vs-Statistician contests:** Whenever enough information is quantified (& a right answer exists), stats wins every time

Examples of what's now possible

- **Opinions of activists:** A few thousand interviews ⤳ millions of political opinions in social media posts (1B tweets/5 days)

- **Exercise:** A survey: "How many times did you exercise last week? ⤳ 500K people carrying cell phones with accelerometers

- **Social contacts:** A survey: "Please tell me your 5 best friends" ⤳ continuous record of phone calls, emails, text messages, bluetooth, social media connections, electronic address books

- **Economic development in developing countries:** Dubious or nonexistent governmental statistics ⤳ satellite images of human-generated light at night, or networks of roads and other infrastructure

- **Expert-vs-Statistician contests:** Whenever enough information is quantified (& a right answer exists), stats wins every time

- Many, many more...

Some New Data Types

Some New Data Types

1. **Unstructured text:** emails (1 LOC every 10 minutes), speeches, government reports, blogs, social media updates, web pages, newspapers, scholarly literature

Some New Data Types

1. **Unstructured text:** emails (1 LOC every 10 minutes), speeches, government reports, blogs, social media updates, web pages, newspapers, scholarly literature
2. **Commercial activity:** credit cards, sales data, and real estate transactions, product RFIDs

Some New Data Types

1. **Unstructured text:** emails (1 LOC every 10 minutes), speeches, government reports, blogs, social media updates, web pages, newspapers, scholarly literature
2. **Commercial activity:** credit cards, sales data, and real estate transactions, product RFIDs
3. **Geographic location:** cell phones, Fastlane or EZPass transponders, garage cameras

Some New Data Types

1. **Unstructured text:** emails (1 LOC every 10 minutes), speeches, government reports, blogs, social media updates, web pages, newspapers, scholarly literature
2. **Commercial activity:** credit cards, sales data, and real estate transactions, product RFIDs
3. **Geographic location:** cell phones, Fastlane or EZPass transponders, garage cameras
4. **Health information:** digital medical records, hospital admittances, google/MS health, and accelerometers and other devices being included in cell phones

Some New Data Types

1. **Unstructured text:** emails (1 LOC every 10 minutes), speeches, government reports, blogs, social media updates, web pages, newspapers, scholarly literature
2. **Commercial activity:** credit cards, sales data, and real estate transactions, product RFIDs
3. **Geographic location:** cell phones, Fastlane or EZPass transponders, garage cameras
4. **Health information:** digital medical records, hospital admittances, google/MS health, and accelerometers and other devices being included in cell phones
5. **Biological sciences:** effectively becoming social sciences as genomics, proteomics, metabolomics, and brain imaging produce huge numbers of *person-level variables*.

Some New Data Types

1. **Unstructured text:** emails (1 LOC every 10 minutes), speeches, government reports, blogs, social media updates, web pages, newspapers, scholarly literature
2. **Commercial activity:** credit cards, sales data, and real estate transactions, product RFIDs
3. **Geographic location:** cell phones, Fastlane or EZPass transponders, garage cameras
4. **Health information:** digital medical records, hospital admittances, google/MS health, and accelerometers and other devices being included in cell phones
5. **Biological sciences:** effectively becoming social sciences as genomics, proteomics, metabolomics, and brain imaging produce huge numbers of *person-level variables*.
6. **Satellite imagery:** increasing in scope, resolution, and availability.

Some New Data Types

1. **Unstructured text:** emails (1 LOC every 10 minutes), speeches, government reports, blogs, social media updates, web pages, newspapers, scholarly literature
2. **Commercial activity:** credit cards, sales data, and real estate transactions, product RFIDs
3. **Geographic location:** cell phones, Fastlane or EZPass transponders, garage cameras
4. **Health information:** digital medical records, hospital admittances, google/MS health, and accelerometers and other devices being included in cell phones
5. **Biological sciences:** effectively becoming social sciences as genomics, proteomics, metabolomics, and brain imaging produce huge numbers of *person-level variables*.
6. **Satellite imagery:** increasing in scope, resolution, and availability.
7. **Electoral activity:** ballot images, precinct-level results, individual-level registration, primary participation, and campaign contributions

Some More New Data Examples

Some More New Data Examples

⑧ Social media: facebook, twitter, social bookmarking, blog comments, product reviews, virtual worlds, game behavior, crowd sourcing

Some More New Data Examples

8. **Social media:** facebook, twitter, social bookmarking, blog comments, product reviews, virtual worlds, game behavior, crowd sourcing

9. **Web surfing artifacts:** clicks, searches, and advertising clickthroughs. (Google collects 1 petabyte/72 minutes on human behavior!)

Some More New Data Examples

8. **Social media:** facebook, twitter, social bookmarking, blog comments, product reviews, virtual worlds, game behavior, crowd sourcing

9. **Web surfing artifacts:** clicks, searches, and advertising clickthroughs. (Google collects 1 petabyte/72 minutes on human behavior!)

10. **Multiplayer web games and virtual worlds:** Billions of highly controlled experiments on human behavior

Some More New Data Examples

8. **Social media:** facebook, twitter, social bookmarking, blog comments, product reviews, virtual worlds, game behavior, crowd sourcing

9. **Web surfing artifacts:** clicks, searches, and advertising clickthroughs. (Google collects 1 petabyte/72 minutes on human behavior!)

10. **Multiplayer web games and virtual worlds:** Billions of highly controlled experiments on human behavior

11. **Government bureaucracies:** moving from paper to electronic data bases, increasing availability

Some More New Data Examples

8. **Social media:** facebook, twitter, social bookmarking, blog comments, product reviews, virtual worlds, game behavior, crowd sourcing

9. **Web surfing artifacts:** clicks, searches, and advertising clickthroughs. (Google collects 1 petabyte/72 minutes on human behavior!)

10. **Multiplayer web games and virtual worlds:** Billions of highly controlled experiments on human behavior

11. **Government bureaucracies:** moving from paper to electronic data bases, increasing availability

12. **Governmental policies:** requiring more data collection, such e.g., "No Child Left Behind Act"; allowing randomized policy experiments; Obama pushing data distribution

Some More New Data Examples

8. **Social media:** facebook, twitter, social bookmarking, blog comments, product reviews, virtual worlds, game behavior, crowd sourcing

9. **Web surfing artifacts:** clicks, searches, and advertising clickthroughs. (Google collects 1 petabyte/72 minutes on human behavior!)

10. **Multiplayer web games and virtual worlds:** Billions of highly controlled experiments on human behavior

11. **Government bureaucracies:** moving from paper to electronic data bases, increasing availability

12. **Governmental policies:** requiring more data collection, such e.g., "No Child Left Behind Act"; allowing randomized policy experiments; Obama pushing data distribution

13. **Scholarly data:** the replication movement in academia is massively increasing data sharing

Enormous Emerging Opportunities for Social Scientists

- For the first time: technologies, policies, data, and methods are making it feasible to attack some of the most vexing problems that afflict human society

Enormous Emerging Opportunities for Social Scientists

- For the first time: technologies, policies, data, and methods are making it feasible to attack some of the most vexing problems that afflict human society

- A massive change from studying problems to understanding and solving problems

Enormous Emerging Opportunities for Social Scientists

- For the first time: technologies, policies, data, and methods are making it feasible to attack some of the most vexing problems that afflict human society
- A massive change from studying problems to understanding and solving problems
- Opportunities require a change in our job descriptions, with new:

Enormous Emerging Opportunities for Social Scientists

- For the first time: technologies, policies, data, and methods are making it feasible to attack some of the most vexing problems that afflict human society
- A massive change from studying problems to understanding and solving problems
- Opportunities require a change in our job descriptions, with new:
  1. Computer-assisted methods:

Enormous Emerging Opportunities for Social Scientists

- For the first time: technologies, policies, data, and methods are making it feasible to attack some of the most vexing problems that afflict human society
- A massive change from studying problems to understanding and solving problems
- Opportunities require a change in our job descriptions, with new:
  1. Computer-assisted methods: Traditional quantitative-only or qualitative-only approaches are infeasible

Enormous Emerging Opportunities for Social Scientists

- For the first time: technologies, policies, data, and methods are making it feasible to attack some of the most vexing problems that afflict human society
- A massive change from studying problems to understanding and solving problems
- Opportunities require a change in our job descriptions, with new:
  1. Computer-assisted methods: Traditional quantitative-only or qualitative-only approaches are infeasible
  2. Large-scale, interdisciplinary, collaborative research

Enormous Emerging Opportunities for Social Scientists

- For the first time: technologies, policies, data, and methods are making it feasible to attack some of the most vexing problems that afflict human society
- A massive change from studying problems to understanding and solving problems
- Opportunities require a change in our job descriptions, with new:
  1. Computer-assisted methods: Traditional quantitative-only or qualitative-only approaches are infeasible
  2. Large-scale, interdisciplinary, collaborative research
  3. New statistical methods & engineering required

Enormous Emerging Opportunities for Social Scientists

- For the first time: technologies, policies, data, and methods are making it feasible to attack some of the most vexing problems that afflict human society
- A massive change from studying problems to understanding and solving problems
- Opportunities require a change in our job descriptions, with new:
  1. Computer-assisted methods: Traditional quantitative-only or qualitative-only approaches are infeasible
  2. Large-scale, interdisciplinary, collaborative research
  3. New statistical methods & engineering required
  4. Better theory: to respond to massive new evidence, privacy challenges, data-driven science

Enormous Emerging Opportunities for Social Scientists

- For the first time: technologies, policies, data, and methods are making it feasible to attack some of the most vexing problems that afflict human society
- A massive change from studying problems to understanding and solving problems
- Opportunities require a change in our job descriptions, with new:
  1. Computer-assisted methods: Traditional quantitative-only or qualitative-only approaches are infeasible
  2. Large-scale, interdisciplinary, collaborative research
  3. New statistical methods & engineering required
  4. Better theory: to respond to massive new evidence, privacy challenges, data-driven science
- And then there's you & me:

Enormous Emerging Opportunities for Social Scientists

- For the first time: technologies, policies, data, and methods are making it feasible to attack some of the most vexing problems that afflict human society
- A massive change from studying problems to understanding and solving problems
- Opportunities require a change in our job descriptions, with new:
  1. Computer-assisted methods: Traditional quantitative-only or qualitative-only approaches are infeasible
  2. Large-scale, interdisciplinary, collaborative research
  3. New statistical methods & engineering required
  4. Better theory: to respond to massive new evidence, privacy challenges, data-driven science
- And then there's you & me:
  - In most legislatures, courts, academic departments, ..., change comes from replacement not conversion

Enormous Emerging Opportunities for Social Scientists

- For the first time: technologies, policies, data, and methods are making it feasible to attack some of the most vexing problems that afflict human society
- A massive change from studying problems to understanding and solving problems
- Opportunities require a change in our job descriptions, with new:
  1. Computer-assisted methods: Traditional quantitative-only or qualitative-only approaches are infeasible
  2. Large-scale, interdisciplinary, collaborative research
  3. New statistical methods & engineering required
  4. Better theory: to respond to massive new evidence, privacy challenges, data-driven science
- And then there's you & me:
  - In most legislatures, courts, academic departments, ..., change comes from replacement not conversion
  - Will we wait to be replaced? or put in the effort to convert and learn how to use the new information?

- Interpersonal comparability

- Interpersonal comparability
- Cross-cultural comparability

- Interpersonal comparability
- Cross-cultural comparability
- Can we be understood by a survey respondent if we can't be understood by our spouse, siblings, and parents?

Readings on Anchoring Vignettes

- Daniel Hopkins and Gary King "Improving Anchoring Vignettes: Designing Surveys to Correct Interpersonal Incomparability," *Public Opinion Quarterly*, 2010, 1–22.

- Gary King and Jonathan Wand. "Comparing Incomparable Survey Responses: Evaluating and Selecting Anchoring Vignettes," *Political Analysis*, 15, 1 (Winter, 2007): 46–66.

- Gary King; Christopher J.L. Murray; Joshua A. Salomon; and Ajay Tandon. "Enhancing the Validity and Cross-cultural Comparability of Measurement in Survey Research," *American Political Science Review*, Vol. 98, No. 1 (February, 2004): 191–207.

- More information: http://GKing.Harvard.edu/vign

The Importance of Survey Research

The Importance of Survey Research

- In political science: 1/2 of all quantitative articles

The Importance of Survey Research

- In political science: $1/2$ of all quantitative articles
- Other social sciences and related professional areas: Very widely used

The Importance of Survey Research

- In political science: $1/2$ of all quantitative articles
- Other social sciences and related professional areas: Very widely used
- A large fraction of our information base over the last half century

The Importance of Survey Research

- In political science: 1/2 of all quantitative articles
- Other social sciences and related professional areas: Very widely used
- A large fraction of our information base over the last half century
- A multi-billion dollar industry

The Importance of Survey Research

- In political science: $1/2$ of all quantitative articles
- Other social sciences and related professional areas: Very widely used
- A large fraction of our information base over the last half century
- A multi-billion dollar industry
- Of widespread public interest

Examples of Interpersonal Incomparability

Examples of Interpersonal Incomparability

"How healthy are you? Excellent, Good, Fair, or Poor"

Examples of Interpersonal Incomparability

"How healthy are you? Excellent, Good, Fair, or Poor"

- Suppose an otherwise healthy 25-year-old woman with a cold and a backache answers "fair" and a 90-year-old man just able to get out of bed says "excellent"

Examples of Interpersonal Incomparability

"How healthy are you? Excellent, Good, Fair, or Poor"

- Suppose an otherwise healthy 25-year-old woman with a cold and a backache answers "fair" and a 90-year-old man just able to get out of bed says "excellent"
- Is the young woman less healthy or are the two interpreting the same question differently?

Examples of Interpersonal Incomparability

"How healthy are you? Excellent, Good, Fair, or Poor"

- Suppose an otherwise healthy 25-year-old woman with a cold and a backache answers "fair" and a 90-year-old man just able to get out of bed says "excellent"
- Is the young woman less healthy or are the two interpreting the same question differently?
- In some countries, responses to this survey question correlate negatively with objective measures of health status (Sen, 2002).

Examples of Interpersonal Incomparability

"How healthy are you? Excellent, Good, Fair, or Poor"

- Suppose an otherwise healthy 25-year-old woman with a cold and a backache answers "fair" and a 90-year-old man just able to get out of bed says "excellent"
- Is the young woman less healthy or are the two interpreting the same question differently?
- In some countries, responses to this survey question correlate negatively with objective measures of health status (Sen, 2002).

"Do you approve of how George W. Bush is handling his job?"

Examples of Interpersonal Incomparability

"How healthy are you? Excellent, Good, Fair, or Poor"

- Suppose an otherwise healthy 25-year-old woman with a cold and a backache answers "fair" and a 90-year-old man just able to get out of bed says "excellent"
- Is the young woman less healthy or are the two interpreting the same question differently?
- In some countries, responses to this survey question correlate negatively with objective measures of health status (Sen, 2002).

"Do you approve of how George W. Bush is handling his job?"

- On 9/10/2001, 55% of Americans approved of the way George W. Bush was "handling his job as president".

Examples of Interpersonal Incomparability

"How healthy are you? Excellent, Good, Fair, or Poor"

- Suppose an otherwise healthy 25-year-old woman with a cold and a backache answers "fair" and a 90-year-old man just able to get out of bed says "excellent"
- Is the young woman less healthy or are the two interpreting the same question differently?
- In some countries, responses to this survey question correlate negatively with objective measures of health status (Sen, 2002).

"Do you approve of how George W. Bush is handling his job?"

- On 9/10/2001, 55% of Americans approved of the way George W. Bush was "handling his job as president".
- The next day — which the president spent in hiding — 90% approved.

Examples of Interpersonal Incomparability

"How healthy are you? Excellent, Good, Fair, or Poor"

- Suppose an otherwise healthy 25-year-old woman with a cold and a backache answers "fair" and a 90-year-old man just able to get out of bed says "excellent"
- Is the young woman less healthy or are the two interpreting the same question differently?
- In some countries, responses to this survey question correlate negatively with objective measures of health status (Sen, 2002).

"Do you approve of how George W. Bush is handling his job?"

- On 9/10/2001, 55% of Americans approved of the way George W. Bush was "handling his job as president".
- The next day — which the president spent in hiding — 90% approved.
- Was this massive opinion change, or was the same question interpreted differently?

How much say [does 'name' / do you] have in getting the government to address issues that interest [him / her / you]?

(a) Unlimited say, (b) A lot of say, (c) Some say, (d) Little say, (e) No say at all

Anchoring Vignettes & Self-Assessments: Political Efficacy (about voting)

- "[Alison] lacks clean drinking water. She and her neighbors are supporting an opposition candidate in the forthcoming elections that has promised to address the issue. It appears that so many people in her area feel the same way that the opposition candidate will defeat the incumbent representative."

How much say [does 'name' / do you] have in getting the government to address issues that interest [him / her / you]?

(a) Unlimited say, (b) A lot of say, (c) Some say, (d) Little say, (e) No say at all

Anchoring Vignettes & Self-Assessments: Political Efficacy (about voting)

- "[Alison] lacks clean drinking water. She and her neighbors are supporting an opposition candidate in the forthcoming elections that has promised to address the issue. It appears that so many people in her area feel the same way that the opposition candidate will defeat the incumbent representative."

- "[Jane] lacks clean drinking water because the government is pursuing an industrial development plan. In the campaign for an upcoming election, an opposition party has promised to address the issue, but she feels it would be futile to vote for the opposition since the government is certain to win."

How much say [does 'name' / do you] have in getting the government to address issues that interest [him / her / you]?
(a) Unlimited say, (b) A lot of say, (c) Some say, (d) Little say, (e) No say at all

Anchoring Vignettes & Self-Assessments: Political Efficacy (about voting)

- "[Alison] lacks clean drinking water. She and her neighbors are supporting an opposition candidate in the forthcoming elections that has promised to address the issue. It appears that so many people in her area feel the same way that the opposition candidate will defeat the incumbent representative."

- "[Jane] lacks clean drinking water because the government is pursuing an industrial development plan. In the campaign for an upcoming election, an opposition party has promised to address the issue, but she feels it would be futile to vote for the opposition since the government is certain to win."

- "[Moses] lacks clean drinking water. He would like to change this, but he can't vote, and feels that no one in the government cares about this issue. So he suffers in silence, hoping something will be done in the future."

How much say [does 'name' / do you] have in getting the government to address issues that interest [him / her / you]?
(a) Unlimited say, (b) A lot of say, (c) Some say, (d) Little say, (e) No say at all
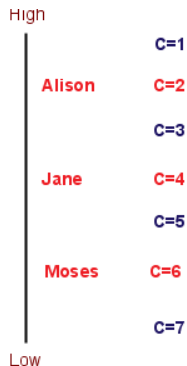
- The only reason different respondents do not agree on vignette positions: Different standards
- Why assumptions hold: investigator creates Alison, Jane & Moses

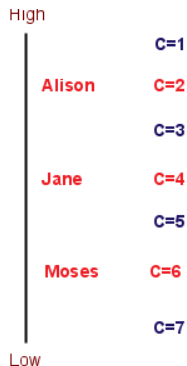# A Simple, Nonparametric Method

A Simple, Nonparametric Method

Define self-assessments *relative* to vignettes:

A Simple, Nonparametric Method

Define self-assessments *relative* to vignettes:

- (Special procedures for vignette ties and inconsistencies)

Comparing China and Mexico
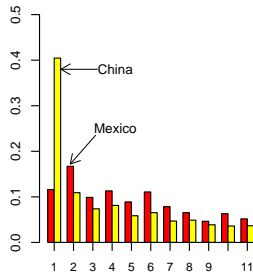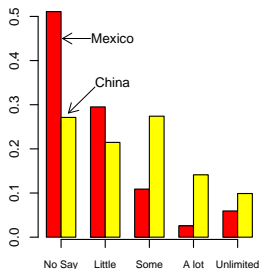
Comparing China and Mexico

Opposition leader Vicente Fox elected President.
71-year rule of PRI party ends.
Peaceful transition of power begins.

Plenty of political efficacy

China: How much say do you have in getting the government to address issues that interest you?

Nonparametric Estimates of Political Efficacy

- The left graph is a histogram of the observed categorical self-assessments.
- The right graph is a histogram of C, our nonparametric DIF-corrected estimate of the same distribution.

A Methodological Problem

A Methodological Problem

- For every question on your survey now: add 3-12 vignettes

A Methodological Problem

- For every question on your survey now: add 3-12 vignettes
- Too expensive, especially for public health surveys in many countries.

A Methodological Problem

- For every question on your survey now: add 3-12 vignettes
- Too expensive, especially for public health surveys in many countries.
- A second method allows: vignettes asked of a small subset of respondents, or in a different survey

Categorizing Years of Age
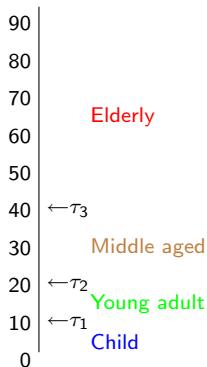
Respondent 1 / Respondent 2
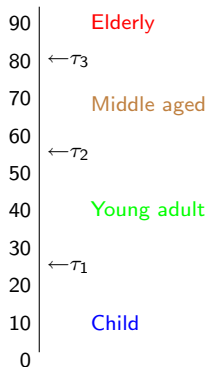
- If thresholds vary, categorical answers are meaningless.
- Our parametric model works by estimating the thresholds.
- Vignettes provide identifying information for the τ's.

Key Measurement Assumptions

Key Measurement Assumptions

1. *Response Consistency*: Each respondent uses the self-assessment and vignette categories in approximately the same way across questions.

Key Measurement Assumptions

1. *Response Consistency*: Each respondent uses the self-assessment and vignette categories in approximately the same way across questions. (DIF occurs across respondents, not across questions for any one respondent.)

Key Measurement Assumptions

1. *Response Consistency*: Each respondent uses the self-assessment and vignette categories in approximately the same way across questions. (DIF occurs across respondents, not across questions for any one respondent.)

2. *Vignette Equivalence*:

Key Measurement Assumptions

1. *Response Consistency*: Each respondent uses the self-assessment and vignette categories in approximately the same way across questions. (DIF occurs across respondents, not across questions for any one respondent.)

2. *Vignette Equivalence*:
   (a) The actual level for any vignette is the same for all respondents.

Key Measurement Assumptions

1. *Response Consistency*: Each respondent uses the self-assessment and vignette categories in approximately the same way across questions. (DIF occurs across respondents, not across questions for any one respondent.)

2. *Vignette Equivalence*:
   (a) The actual level for any vignette is the same for all respondents.
   (b) The quantity being estimated exists.

Key Measurement Assumptions

1. *Response Consistency*: Each respondent uses the self-assessment and vignette categories in approximately the same way across questions. (DIF occurs across respondents, not across questions for any one respondent.)
2. *Vignette Equivalence*:
   (a) The actual level for any vignette is the same for all respondents.
   (b) The quantity being estimated exists.
   (c) The scale being tapped is perceived as unidimensional.

Key Measurement Assumptions

1. *Response Consistency*: Each respondent uses the self-assessment and vignette categories in approximately the same way across questions. (DIF occurs across respondents, not across questions for any one respondent.)

2. *Vignette Equivalence*:
   (a) The actual level for any vignette is the same for all respondents.
   (b) The quantity being estimated exists.
   (c) The scale being tapped is perceived as unidimensional.

3. In other words: we allow response-category DIF but assume stem question equivalence.

Self-Assessments v. Medical Tests

Self-Assessment:
In the last 30 days, how much difficulty did [you/name] have in seeing and recognizing a person you know across the road (i.e. from a distance of about 20 meters)?

Self-Assessments v. Medical Tests

Self-Assessment:
In the last 30 days, how much difficulty did [you/name] have in seeing and recognizing a person you know across the road (i.e. from a distance of about 20 meters)? (A) none, (B) mild, (C) moderate, (D) severe, (E) extreme/cannot do

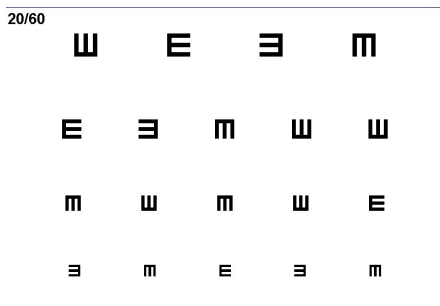## Self-Assessments v. Medical Tests

Self-Assessment:
In the last 30 days, how much difficulty did [you/name] have in seeing and recognizing a person you know across the road (i.e. from a distance of about 20 meters)? (A) none, (B) mild, (C) moderate, (D) severe, (E) extreme/cannot do

The Snellen Eye Chart Test:
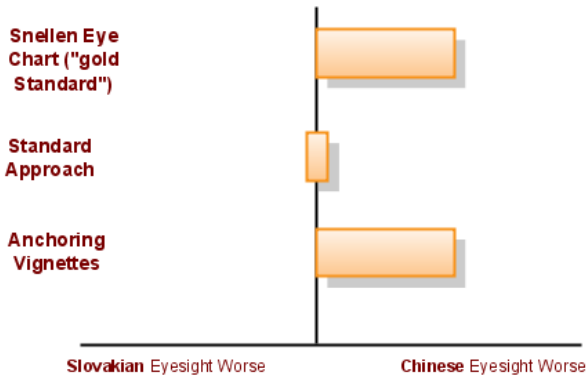
Vision

Fixing DIF in Self-Assessments of Visual (Non)acuity

	Snellen Eye Chart		Ordinal Probit		Chopit	
	Mean	(s.e.)	μ	(s.e.)	μ	(s.e.)
Slovakia	8.006	(.272)	.660	(.127)	.286	(.129)
China	10.780	(.148)	.673	(.073)	.749	(.081)
Difference	−2.774	(.452)	−.013	(.053)	−.463	(.053)

Fixing DIF in Self-Assessments of Visual (Non)acuity

	Snellen Eye Chart		Ordinal Probit		Chopit	
	Mean	(s.e.)	μ	(s.e.)	μ	(s.e.)
Slovakia	8.006	(.272)	.660	(.127)	.286	(.129)
China	10.780	(.148)	.673	(.073)	.749	(.081)
Difference	−2.774	(.452)	−.013	(.053)	−.463	(.053)

- The medical test shows Slovakians see much better than the Chinese

Fixing DIF in Self-Assessments of Visual (Non)acuity

	Snellen Eye Chart		Ordinal Probit		Chopit	
	Mean	(s.e.)	μ	(s.e.)	μ	(s.e.)
Slovakia	8.006	(.272)	.660	(.127)	.286	(.129)
China	10.780	(.148)	.673	(.073)	.749	(.081)
Difference	−2.774	(.452)	−.013	(.053)	−.463	(.053)

- The medical test shows Slovakians see much better than the Chinese
- Ordinal probit finds no difference

Fixing DIF in Self-Assessments of Visual (Non)acuity

	Snellen Eye Chart		Ordinal Probit		Chopit	
	Mean	(s.e.)	μ	(s.e.)	μ	(s.e.)
Slovakia	8.006	(.272)	.660	(.127)	.286	(.129)
China	10.780	(.148)	.673	(.073)	.749	(.081)
Difference	−2.774	(.452)	−.013	(.053)	−.463	(.053)

- The medical test shows Slovakians see much better than the Chinese
- Ordinal probit finds no difference
- Chopit reproduces the same result as the medical test (though on different scale)

Applications: Intended and Unintended

- Direct applications

- Direct applications
  - By Academic researchers:

Applications: Intended and Unintended

- Direct applications
  - By Academic researchers:
    - public health, medicine, political science, psychology, education, sociology, law, marketing research, economics.

Applications: Intended and Unintended

- Direct applications
  - By Academic researchers:
    - public health, medicine, political science, psychology, education, sociology, law, marketing research, economics.
  - Survey Organizations:

Applications: Intended and Unintended

- Direct applications
    - By Academic researchers:
        - public health, medicine, political science, psychology, education, sociology, law, marketing research, economics.
    - Survey Organizations:
        - World Health Organization in several waves in about 80 countries

Applications: Intended and Unintended

- Direct applications
  - By Academic researchers:
    - public health, medicine, political science, psychology, education, sociology, law, marketing research, economics.
  - Survey Organizations:
    - World Health Organization in several waves in about 80 countries
    - U.S. Government doing 120 surveys in 70 countries/year

Applications: Intended and Unintended

- Direct applications
    - By Academic researchers:
        - public health, medicine, political science, psychology, education, sociology, law, marketing research, economics.
    - Survey Organizations:
        - World Health Organization in several waves in about 80 countries
        - U.S. Government doing 120 surveys in 70 countries/year
        - major marketing research firms

Applications: Intended and Unintended

- Direct applications
  - By Academic researchers:
    - public health, medicine, political science, psychology, education, sociology, law, marketing research, economics.
  - Survey Organizations:
    - World Health Organization in several waves in about 80 countries
    - U.S. Government doing 120 surveys in 70 countries/year
    - major marketing research firms
- Methodological Developments

Applications: Intended and Unintended

- Direct applications
  - By Academic researchers:
    - public health, medicine, political science, psychology, education, sociology, law, marketing research, economics.
  - Survey Organizations:
    - World Health Organization in several waves in about 80 countries
    - U.S. Government doing 120 surveys in 70 countries/year
    - major marketing research firms
- Methodological Developments
  - Statistics, political methodology, public health, econometrics

Applications: Intended and Unintended

- Direct applications
  - By Academic researchers:
    - public health, medicine, political science, psychology, education, sociology, law, marketing research, economics.
  - Survey Organizations:
    - World Health Organization in several waves in about 80 countries
    - U.S. Government doing 120 surveys in 70 countries/year
    - major marketing research firms
- Methodological Developments
  - Statistics, political methodology, public health, econometrics
- Forging agreement when none seemed possible

Applications: Intended and Unintended

- Direct applications
  - By Academic researchers:
    - public health, medicine, political science, psychology, education, sociology, law, marketing research, economics.
  - Survey Organizations:
    - World Health Organization in several waves in about 80 countries
    - U.S. Government doing 120 surveys in 70 countries/year
    - major marketing research firms
- Methodological Developments
  - Statistics, political methodology, public health, econometrics
- Forging agreement when none seemed possible
  - Privacy in the Information Age (National Research Council)

Applications: Intended and Unintended

- Direct applications
  - By Academic researchers:
    - public health, medicine, political science, psychology, education, sociology, law, marketing research, economics.
  - Survey Organizations:
    - World Health Organization in several waves in about 80 countries
    - U.S. Government doing 120 surveys in 70 countries/year
    - major marketing research firms
- Methodological Developments
  - Statistics, political methodology, public health, econometrics
- Forging agreement when none seemed possible
  - Privacy in the Information Age (National Research Council)
  - Legal scholarship

Applications: Intended and Unintended

- Direct applications
  - By Academic researchers:
    - public health, medicine, political science, psychology, education, sociology, law, marketing research, economics.
  - Survey Organizations:
    - World Health Organization in several waves in about 80 countries
    - U.S. Government doing 120 surveys in 70 countries/year
    - major marketing research firms
- Methodological Developments
  - Statistics, political methodology, public health, econometrics
- Forging agreement when none seemed possible
  - Privacy in the Information Age (National Research Council)
  - Legal scholarship
  - Formalizing Philosophy

Conclusions

Conclusions

- Our approach can fix DIF, if response consistency and vignette equivalence hold — and the survey questions are good

Conclusions

- Our approach can fix DIF, if response consistency and vignette equivalence hold — and the survey questions are good
- Anchoring vignettes will not eliminate all DIF, but problems would have to occur at unrealistically extreme levels to make the unadjusted measures better than the adjusted ones.

Conclusions

- Our approach can fix DIF, if response consistency and vignette equivalence hold — and the survey questions are good

- Anchoring vignettes will not eliminate all DIF, but problems would have to occur at unrealistically extreme levels to make the unadjusted measures better than the adjusted ones.

- Expense can be held down to a minimum by assigning each vignette to a smaller subsample. E.g., 4 vignettes asked for 1/4 of the sample each adds only one question/respondent.

Conclusions

- Our approach can fix DIF, if response consistency and vignette equivalence hold — and the survey questions are good

- Anchoring vignettes will not eliminate all DIF, but problems would have to occur at unrealistically extreme levels to make the unadjusted measures better than the adjusted ones.

- Expense can be held down to a minimum by assigning each vignette to a smaller subsample. E.g., 4 vignettes asked for 1/4 of the sample each adds only one question/respondent.

- Writing vignettes aids in the clarification and discovery of additional domains of the concept of interest — even if you do not do a survey.

Conclusions

- Our approach can fix DIF, if response consistency and vignette equivalence hold — and the survey questions are good

- Anchoring vignettes will not eliminate all DIF, but problems would have to occur at unrealistically extreme levels to make the unadjusted measures better than the adjusted ones.

- Expense can be held down to a minimum by assigning each vignette to a smaller subsample. E.g., 4 vignettes asked for 1/4 of the sample each adds only one question/respondent.

- Writing vignettes aids in the clarification and discovery of additional domains of the concept of interest — even if you do not do a survey.

- We do not provide a solution for other common survey problems: Question wording,

Conclusions

- Our approach can fix DIF, if response consistency and vignette equivalence hold — and the survey questions are good

- Anchoring vignettes will not eliminate all DIF, but problems would have to occur at unrealistically extreme levels to make the unadjusted measures better than the adjusted ones.

- Expense can be held down to a minimum by assigning each vignette to a smaller subsample. E.g., 4 vignettes asked for 1/4 of the sample each adds only one question/respondent.

- Writing vignettes aids in the clarification and discovery of additional domains of the concept of interest — even if you do not do a survey.

- We do not provide a solution for other common survey problems: Question wording, Accurate translation,

Conclusions

- Our approach can fix DIF, if response consistency and vignette equivalence hold — and the survey questions are good

- Anchoring vignettes will not eliminate all DIF, but problems would have to occur at unrealistically extreme levels to make the unadjusted measures better than the adjusted ones.

- Expense can be held down to a minimum by assigning each vignette to a smaller subsample. E.g., 4 vignettes asked for 1/4 of the sample each adds only one question/respondent.

- Writing vignettes aids in the clarification and discovery of additional domains of the concept of interest — even if you do not do a survey.

- We do not provide a solution for other common survey problems: Question wording, Accurate translation, Question order,

Conclusions

- Our approach can fix DIF, if response consistency and vignette equivalence hold — and the survey questions are good

- Anchoring vignettes will not eliminate all DIF, but problems would have to occur at unrealistically extreme levels to make the unadjusted measures better than the adjusted ones.

- Expense can be held down to a minimum by assigning each vignette to a smaller subsample. E.g., 4 vignettes asked for 1/4 of the sample each adds only one question/respondent.

- Writing vignettes aids in the clarification and discovery of additional domains of the concept of interest — even if you do not do a survey.

- We do not provide a solution for other common survey problems: Question wording, Accurate translation, Question order, Sampling design,

Conclusions

- Our approach can fix DIF, if response consistency and vignette equivalence hold — and the survey questions are good

- Anchoring vignettes will not eliminate all DIF, but problems would have to occur at unrealistically extreme levels to make the unadjusted measures better than the adjusted ones.

- Expense can be held down to a minimum by assigning each vignette to a smaller subsample. E.g., 4 vignettes asked for 1/4 of the sample each adds only one question/respondent.

- Writing vignettes aids in the clarification and discovery of additional domains of the concept of interest — even if you do not do a survey.

- We do not provide a solution for other common survey problems: Question wording, Accurate translation, Question order, Sampling design, Interview length,

Conclusions

- Our approach can fix DIF, if response consistency and vignette equivalence hold — and the survey questions are good

- Anchoring vignettes will not eliminate all DIF, but problems would have to occur at unrealistically extreme levels to make the unadjusted measures better than the adjusted ones.

- Expense can be held down to a minimum by assigning each vignette to a smaller subsample. E.g., 4 vignettes asked for 1/4 of the sample each adds only one question/respondent.

- Writing vignettes aids in the clarification and discovery of additional domains of the concept of interest — even if you do not do a survey.

- We do not provide a solution for other common survey problems: Question wording, Accurate translation, Question order, Sampling design, Interview length, Social backgrounds of interviewer and respondent, etc.

http://GKing.Harvard.edu/vign

Includes:

- Academic papers
- Anchoring vignette examples by researchers in many fields,
- Frequently asked questions,
- Videos
- Conferences
- Statistical software