

How accurate are surveys of objective phenomena?

LinChiat Chang

Jon A. Krosnick

Methods to Assess Accuracy

- Match each respondent's self-report with objective individual records of the same phenomena
- Match one-time aggregate survey percentages and means with available benchmarks from non-survey data
- Correlate individuals' self-reports on surveys with secondary objective data
- Correlate trends over time in self-reports with trends in objective benchmarks

Methods used to Assess Accuracy

- Match each respondent's self-report with objective individual records of the same phenomena – **555 comparisons**
- Match one-time aggregate survey percentages and means with available benchmarks from non-survey data – **399 comparisons**
- Correlate self-reports on surveys and secondary objective data – **168 comparisons**
- Correlate trends over time in self-reports and objective benchmarks – **6 instances**
- **Total of over 1,000 instances of validation against objective measures**

Review Across Diverse Domains

- Alcohol Use
- Crime & Deviance
- Demographics
- Dental Care
- Economic Indicators
- Education
- Estimated Frequency of Events
- Home Value Appreciation
- Healthcare Utilization
- Height & Weight
- Illegal Substance Use
- Work absenteeism
- Media Use
- Medical Costs/Payments
- Recent Medical Diagnosis
- Medical Screening Tests
- Chronic Ailments
- Consumer Purchases
- Philanthropy
- Prescription Drug Use
- Smoking
- Social Services Utilization
- Tax Evasion
- Voter Turnout

Method 1

- Match each respondent's self-report with objective individual records of the same phenomena
- Total number of estimates reviewed = 555
- Total number of individual matches = 520,131
- What percentage of individual self-reports yielded perfect agreement with objective records/measures?
 - Average Percent of Perfect Match = 85%
 - Median Percent of Perfect Match = 88%

Matching Self-reports vs. Objective Records: Summary Statistics

| <i>Topic</i> | <i>Mean Percent of Perfect Match between Self-reports & Objective Records</i> | <i>Median Percent of Perfect Match between Self-reports & Objective Records</i> | <i>Total Number of Individual Comparisons</i> |
|-----------------------------|---|---|---|
| Smoking | 93% | 95% | 9,755 |
| Crime & Deviance | 87% | 93% | 1,242 |
| Demographics | 87% | 92% | 7,148 |
| Recent Medical Diagnosis | 86% | 86% | 34,268 |
| Alcohol Use | 86% | 86% | 561 |
| Chronic Ailments | 86% | 90% | 282,716 |
| Healthcare Utilization | 85% | 89% | 63,531 |
| Dental Care | 84% | 88% | 29,138 |
| Labor Force Statistics | 84% | 88% | 5,051 |
| Consumer Purchases | 83% | 83% | 2,448 |
| Illegal Substance Use | 81% | 85% | 31,785 |
| Medical Screening Tests | 78% | 77% | 6,241 |
| Prescription Medication Use | 76% | 78% | 5,057 |
| Medical Costs/Payments | 75% | 75% | 534 |
| Height & Weight | 74% | 75% | 3,290 |
| Voter registration/turnout | 73% | 72% | 5,520 |
| Philanthropy | 66% | 66% | 920 |
| School grades & test scores | 57% | 54% | 30,926 |
| Overall Average | 85% | 88% | 520,131 |

Matching Self-reports vs. Objective Records: Specific Examples

| <i>Source</i> | <i>Survey Measure</i> | <i>Validation Method</i> | <i>Proportion of Exact Matches</i> | <i>Sample Size</i> |
|---------------------------------|--------------------------------------|--|------------------------------------|--------------------|
| <u>Petrou</u> et al. (2002) | Whether was hospitalized | Medical records | 100% | 82 |
| King et al. (1990) | Whether had mammogram in past year | Radiology records | 100% | 99 |
| <u>Guerriere</u> et al. (2006) | Whether had Physiotherapy | Administrative records | 100% | 110 |
| <u>Carsjo</u> et al. (1994) | Whether was hospitalized | Healthcare utilization records | 98% | 157 |
| McKenna et al. (1992) | Whether had pap smear | Medical records | 97% | 105 |
| <u>Vernacchio</u> et al. (2007) | Whether diagnosed with acute otitis | Medical records | 97% | 272 |
| Wolford et al. (2008) | Whether has PTSD | Clinician-Administered PTSD Scale (CAPS) | 91% | 58 |
| Roberts et al. (1996) | How many ambulatory physician visits | Medical records | 91% | 495 |
| Killeen et al. (2004) | Whether open social service case | Social service delivery records | 90% | 139 |
| <u>Weissman</u> et al. (1996) | Hours of homecare | Medical and financial records | 89% | 106 |

Method 2

- Match one-time aggregate survey percentage or mean estimates with available benchmarks from non-survey data
- Total number of estimates reviewed = 399
- Many different units of measurement
 - Percentages
 - Means in cm, kg, days, hours, times, teeth, drinks etc.
- Difference between survey means vs. objective means?
 - 8% perfect match
 - 38% almost perfect match (less than 1 unit difference)
 - 73% very close (less than 5 units difference)

Matching Aggregate Percentages

| <i>Source</i> | <i>Survey Measure</i> | <i>Survey Mean (X)</i> | <i>Objective Mean (Y)</i> | <i>Absolute Difference (X - Y)</i> |
|----------------------------------|--|------------------------|---------------------------|------------------------------------|
| Ford et al. (1997) | Whether smoked before pregnancy (%) | 31.1 | 31.3 | 0.2 |
| Bauman & Ennett (1994) | Whether recently smoked cigarettes (%) | 4.6 | 4.3 | 0.3 |
| Kreuter et al. (2010) | Whether is currently employed (%) | 27 | 26 | 1 |
| Means et al. (1994) | Whether smoked 5 days prior (%) | - | - | 1.93 |
| Murphy et al. (2000) | Whether smoked marijuana today or yesterday (%) | 17 | 15 | 2 |
| Gilbert et al. (2002) | Whether lost teeth or had teeth removed in six months before interview (%) | 34 | 36 | 2 |
| Pyy-Martikainen & Rendtel (2009) | Unemployment spell ended in person exiting labor force (%) | 22 | 25 | 3 |
| Goetzmann & Peles (1997) | Annual return on mutual fund (%) | - | - | 3.4 |

Matching Aggregate Means

| <i>Source</i> | <i>Survey Measure</i> | <i>Survey Mean (X)</i> | <i>Objective Mean (Y)</i> | <i>Absolute Difference (X - Y)</i> |
|------------------------------|---|------------------------|---------------------------|------------------------------------|
| Dauphinot et al. (2009) | Height (cm) | 163 | 163 | 0 |
| Morrissey et al. (2006) | Height (cm) | 166 | 166 | 0 |
| Gilbert et al. (1997) | Number of teeth | 21.1 | 21.2 | 0.1 |
| Embree & Whitehead (1991) | Number of alcohol drinks consumed | 4.13 | 3.96 | 0.17 |
| Chatwin et al. (1968) | Number of times visited the dentist in past year | 1.69 | 1.52 | 0.17 |
| Preston-Martin et al. (1985) | Number of full mouth X-rays at dentists first visited within 20 years of reference year | - | - | 0.4 |
| Villanueva (2001) | Weight (kg) | 82.1 | 81.6 | 0.5 |
| Cleary & Jette (1984) | Number of times received outpatient medical care in past year | - | - | 0.5 |
| Revicki et al. (1994) | Number of leave days took from work in the past 3 months | 6.11 | 5.59 | 0.52 |
| Douglass et al. (1991) | Number of teeth in maxilla | - | - | 0.61 |
| Yaffe et al. (1978) | How much charged for dentist during 7 month study period (\$) | 27.9 | 27.25 | 0.65 |

Method 3

- Correlate self-reports on surveys and secondary objective data
- Total number of estimates reviewed = 168

| <i>Statistic Reported</i> | <i>Mean Association</i> | <i>Median Association</i> | <i>Total Number of Survey Items Validated using this method</i> |
|---|-------------------------|---------------------------|---|
| Yule's coefficient of association | .77 | .76 | 10 |
| <u>Intraclass Correlation (ICC)</u> | .70 | .69 | 21 |
| Pearson's product-moment coefficient | .69 | .73 | 47 |
| Spearman's rank correlation coefficient | .60 | .67 | 9 |
| Cohen's kappa coefficient | .57 | .68 | 65 |
| Spearman-Brown split-half reliability coefficient | .32 | .22 | 16 |

Correlating Self-reports vs. Objective Records: Specific Examples

| <i>Source</i> | <i>Survey Measure</i> | <i>Validation Method</i> | <i>Correlation Type</i> | <i>Correlation Value</i> | <i>Sample Size</i> |
|--------------------------|---|---|-------------------------|--------------------------|--------------------|
| Gilbert et al. (2003) | Whether had dental cap or implant | Dental charts | Yule's q | .77 | 618 |
| Gilbert et al. (2003) | Whether had amalgam restoration | Dental charts | Yule's q | .75 | 618 |
| Fonseca et al. (2004) | Height | Actual measurement | ICC | .98 | 3713 |
| van Poppel et al. (2002) | How many days of sick leave taken | Individuals' company sick leave records | ICC | .87 | 31 |
| Revicki et al. (1994) | How many days of sick leave taken | Individuals' time card records | ICC | .86 | 36 |
| Killeen et al. (2004) | Whether had been arrested in the past 12 months | Individuals' records of service delivery from the criminal justice system | ICC | .69 | 139 |
| Shields et al. (2008) | Weight (females only) | Actual measurement | Pearson | .97 | 79 |
| Elgar et al. (2005) | Weight (girls only) | Actual measurement | Pearson | .95 | 211 |
| Shields et al. (2008) | Height (females only) | Actual measurement | Pearson | .88 | 79 |
| Spector & Bedell (1982) | Number of days of hospital stay in past 18 months | Records at one specific mental hospital in FL | Pearson | .73 | 201 |

Method 4

- Correlate trends over time in self-reports and objective benchmarks
- Data drawn from 6 high quality representative surveys conducted over decades

Method 4

| <i>Survey</i> | <i>Survey Measure</i> | <i>Number of Years of Available Data</i> | <i>Objective Benchmark</i> | <i>Pearson's r</i> |
|--|---|--|--|--------------------|
| Monitoring the Future (MTF) Youth Survey | Teen Drinking & Driving | 17 | NHTSA counts of teenage drivers in alcohol-related fatal crashes | .96 |
| American National Elections Study (ANES) | Voter Turnout | 40 | FEC records | .94 |
| National Crime Victimization Survey (NCVS) | Crime Victimization | 27 | FBI crime records | .91 |
| Survey of Consumer Attitudes (SCA) | Consumer Assessment of National Economy | 35 | Real Gross Domestic Product (GDP) | .90 |
| Survey of Consumer Attitudes (SCA) | Home Purchase Intent | 25 | Actual Home Sales from JEC report to US Congress | .77 |
| Survey of Consumer Attitudes (SCA) | Automobile Purchase Intent | 25 | Actual Vehicle Sales from JEC report to US Congress | .73 |

Conclusion:

Vast majority
of survey measures
of objective phenomena are
highly accurate