A PRACTICAL GUIDE FOR SURVEYS BASED ON NONPROBABILITY SAMPLES
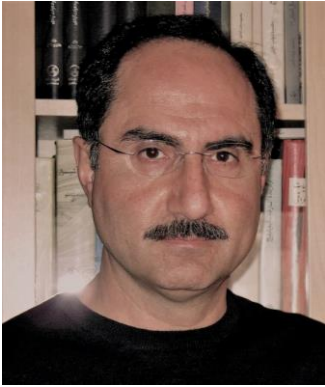
Mansour Fahimi

Frances M. Barlas

Randall K. Thomas

13 February, 2018

Presenters

Mansour Fahimi, Ph.D.
Chief Statistician & Senior Vice President

Frances M. Barlas, Ph.D.
Senior Research Scientist & Vice President

Randall K. Thomas
Chief Survey Methodologist & Senior Vice President
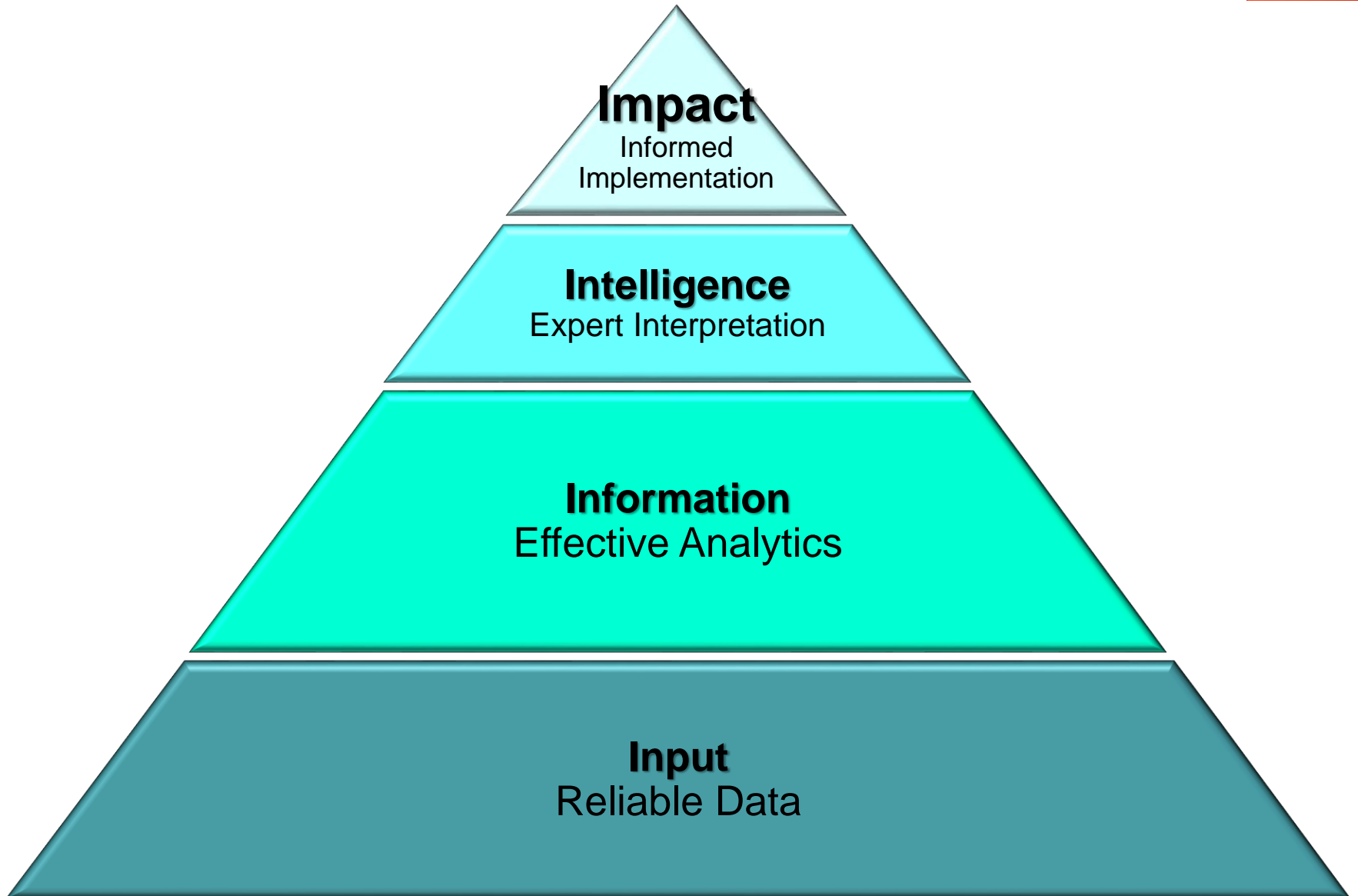
Webinar Objectives

1. **Demystifying nonprobability samples and their pragmatic utilities.**

2. **Reviewing best practices for nonprobability sample sourcing and selection methodologies.**

3. **Identifying and dealing with the hidden hazards of nonprobability samples.**

Webinar Agenda

1. **Definition and types of nonprobability samples**

2. **Guidelines for sample sourcing and selection**

3. **Fitness for use – pros and cons**

4. **Data quality and factors affecting it**

5. **Peeking outside of the box**

From Input to Impact

Impact
Informed
Implementation

Intelligence
Expert Interpretation

Information
Effective Analytics

Input
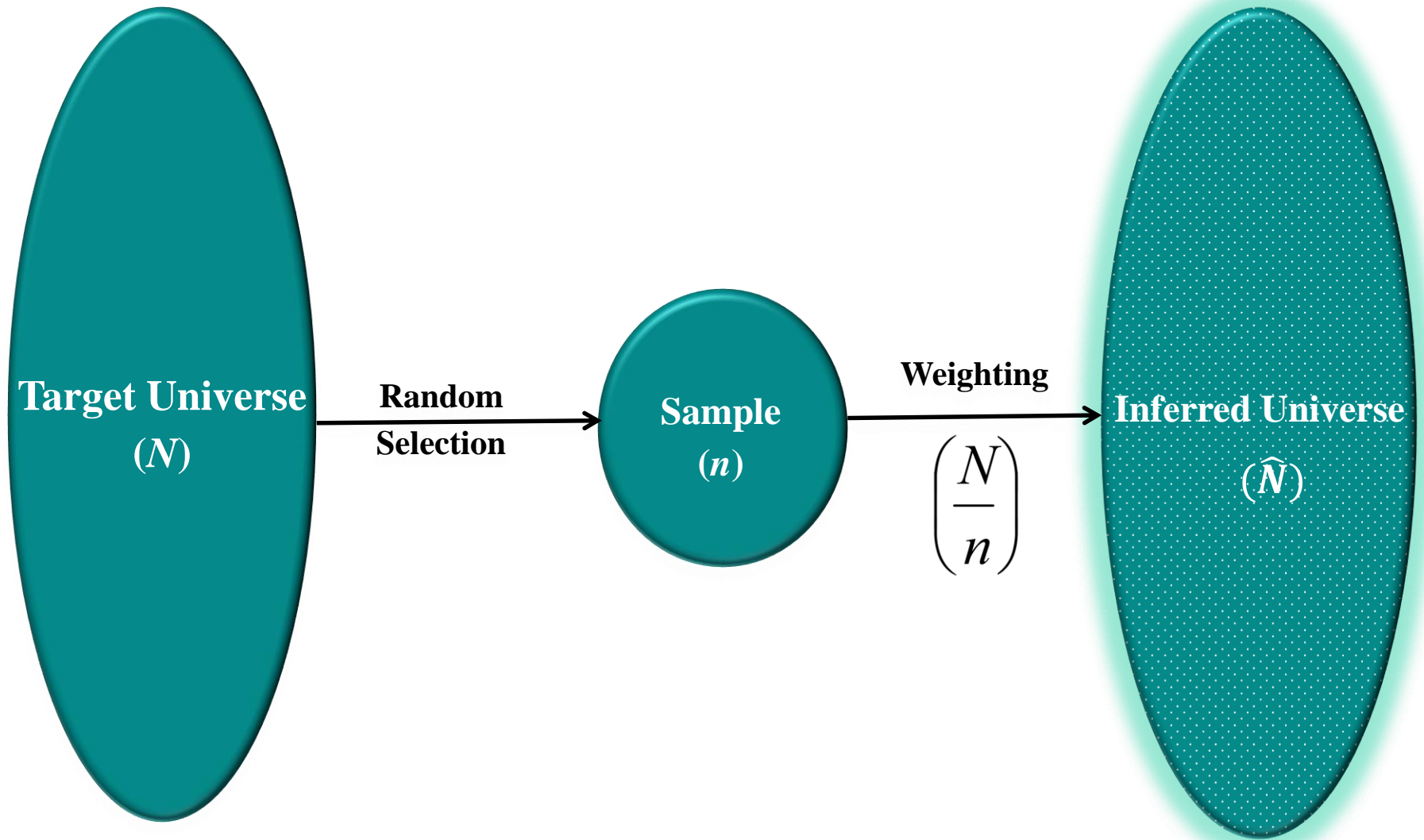Reliable Data

Sample Types & Characteristics

1. **Probability Samples:**

   a. Every unit has a known and nonzero chance of selection

   b. Availability of a sampling frame representative of the population

   c. Random nonresponse uncorrelated with outcome measures

   d. Calculable response rates

   e. Estimable error margins
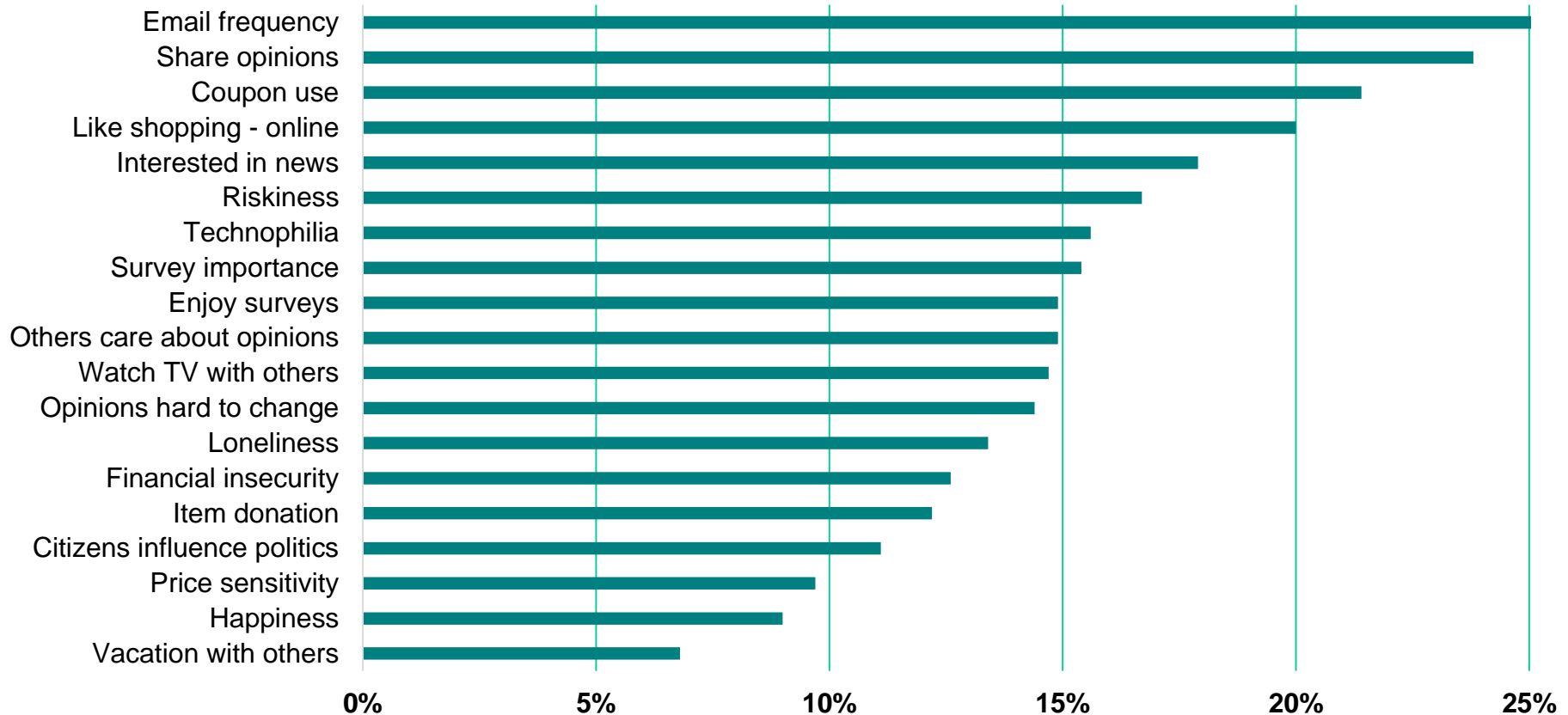
2. **Nonprobability Samples (none of the above):**

   a. Sampling units do not have known selection probabilities

   b. Available sampling frames have unknown coverage properties

   c. Response rates are inestimable and uninformative

   d. Error margins are inestimable

Reasonable Substitute for the Real Deal?

Target Universe (N) → **Random Selection** → Sample (n) → **Weighting** $\left(\dfrac{N}{n}\right)$ → Inferred Universe ($\hat{N}$)
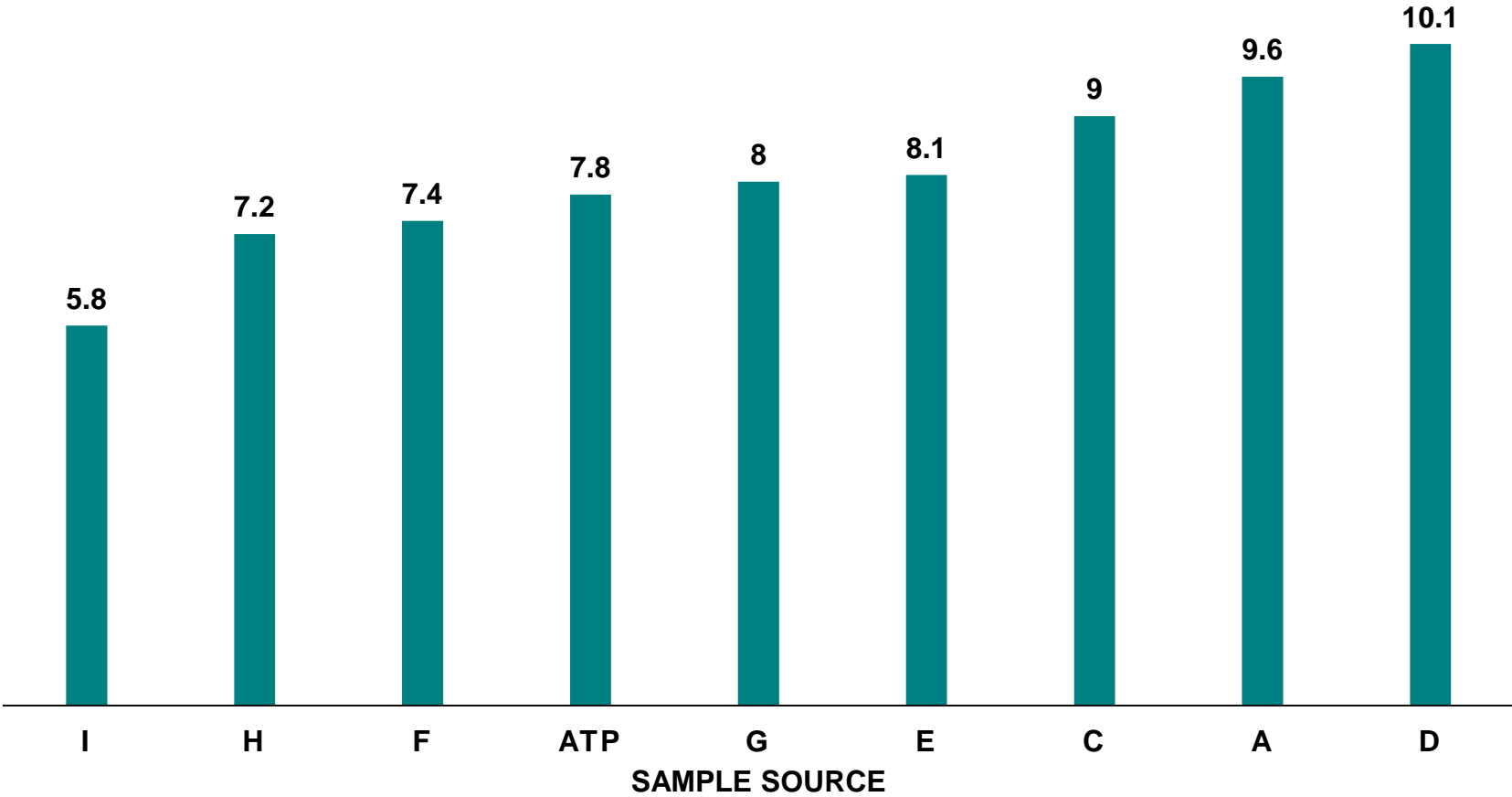
Probability vs. Nonprobability Samples

Significant Differences Between Weighted Estimates from Probability and Nonprobability Samples

Fahimi, M., F. Barlas, R. Thomas, and N. Buttermore (2015). Scientific Surveys Based on Incomplete Sampling Frames and High Rates of Nonresponse. *Survey Practice*, Vol. 8, no 5, 2015, December Issue.
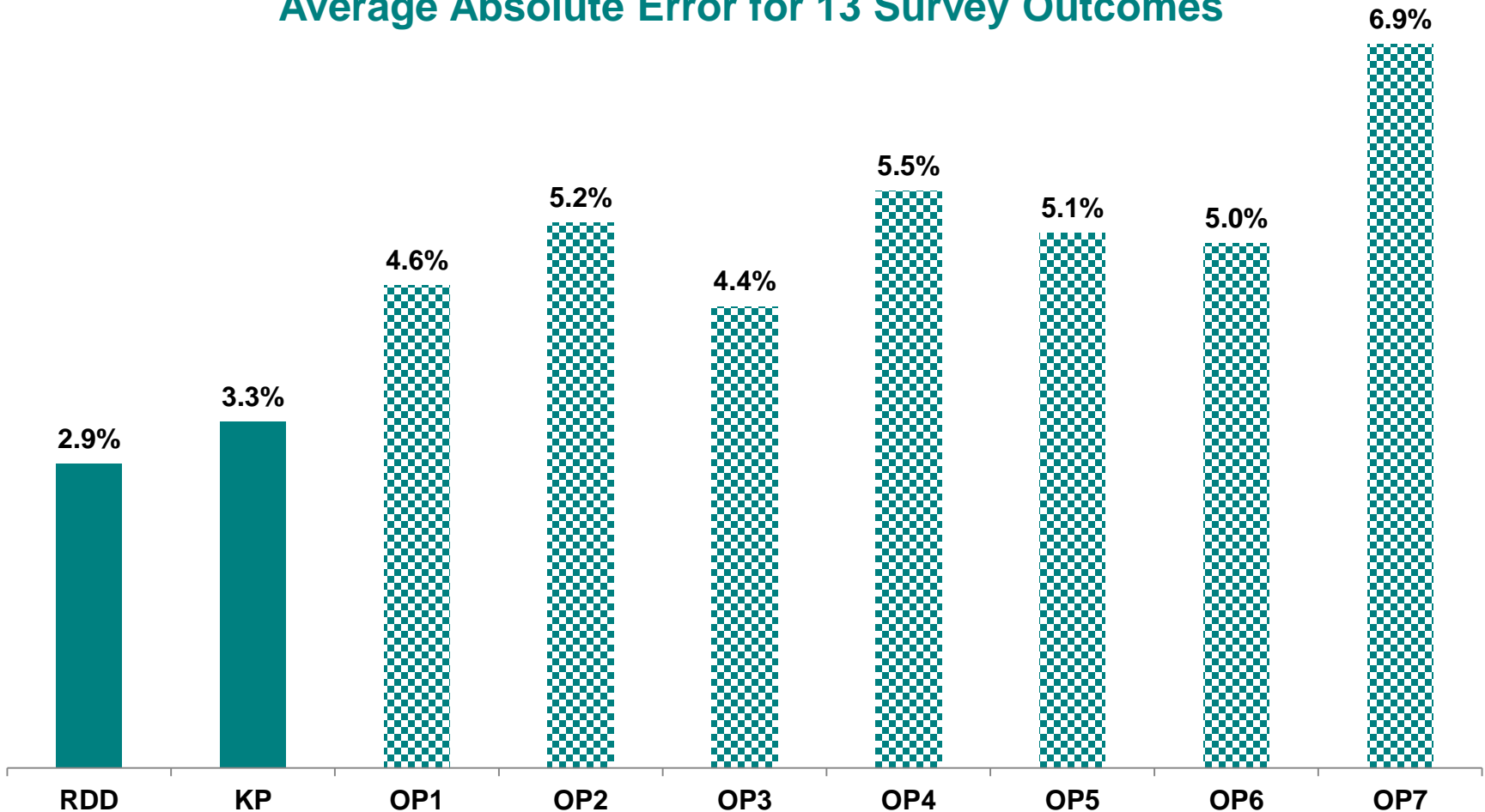
Probability vs. Nonprobability Samples

Average Bias of 20 Weighted Estimates by Sample (Pew)

Source: Pew Research Center, May 2016, "Evaluating Online Nonprobability Surveys."

Probability vs. Nonprobability Samples

Average Absolute Error for 13 Survey Outcomes

Source: Yeager, Krosnick, et. al., 2011. Comparing the Accuracy of RDD Telephone Surveys and Internet Surveys Conducted with Probability and Non-Probability Samples. Public Opinion Quarterly. 75:709-747.

Probability vs. Nonprobability Samples

Comparing KnowledgePanel to Nonprobability

1. **A general population study found approximately 1.5 × as much bias across 30 estimates from nonprobability sample compared to KnowledgePanel**

   Barlas, F. M., Thomas, R. K., & Fahimi, M. (2017). Effects of data cleaning on bias reduction. Paper presented at the 7th Conference of the European Survey Research Association, Lisbon, Portugal.

2. **In predicting election outcomes in 19 races across 8 states, we found the error was nearly 2 × as large using nonprobability sample compared to KnowledgePanel**

   Thomas, R. K., Barlas, F. M., & Weber, A., Pace, D., & Swanson, E. (2016). Experiments in election polling 2014-2016.

Why Nonprobability Samples?

1. **Challenges with probability samples for rare subgroups**

2. **Increasing cost & time constraints of probability samples**

3. **Proliferation of inexpensive data from "digital" sources**

4. **Interest in data unattainable from structured surveys**

5. **Challenging requirements for probability samples:**

    a. Full coverage of the target population

    b. Growing rates of nonresponse

    c. Increasing demands for doing more with less

Types of Nonprobability Samples

1. **Convenience Samples:**

   a. Random samples from incomplete frames

   b. Nonrandom samples from complete/incomplete frames

   c. Intercepts

2. **Synthesized Representation (Incomplete Frames):**

   a. Granular stratification from list frames

   b. Geodemographic balancing of respondents via quotas

3. **Snowballing:**

   a. Start with a probability or nonprobability seed sample

   b. Identify other eligible respondents from the seed sample

Online Samples

1. **Growing Majority of Nonprobability Samples:**
   a. Panel-based
   b. On the fly or river sampling
   c. Mixed sources

2. **Online Panels with:**
   a. Varying extent of profile data for sample balancing
   b. Varying quality and recency of profile data

3. **River Samples:**
   a. Potential respondents roped into surveys while surfing the web
   b. Oftentimes with no profile data

4. **Mixed Source Samples:**
   a. Panel and river samples mixed manually
   b. Panel and river samples mixed via routers

What is a Router?

1. **Technology (software) for constructing and managing online samples by:**

    a. Facilitating survey set-up & quota control
    b. Managing supplier allocations
    c. Implementing certain pre-screeners
    d. Administering certain QC and lock out protocols
    e. Opening projects up to a larger list of sample sources
    f. Tapping into a more diverse pool of potential respondents

2. **Potential concerns:**

    a. A router is only as good as how it is used to manage samples
    b. Certain sample sources are not available via routers
    c. Low transparency about available sample sources

3. **Examples of router providers:**

    a. Fulcrum: https://luc.id/fulcrum
    b. Cint: https://www.cint.com

Guidelines for Sample Sourcing & Selection

1. **A single source may not address all project needs:**

   a. Coverage issues due to varying recruitment methods

   b. Varying profile data available for sample balancing/stratification

   c. Large or tracker surveys can require large samples

2. **Geodemographically unbalanced samples are risky:**

   a. Depending on recruitment method samples can be very skewed

   b. Low cooperation rates can exasperate the above

   c. It is advisable to use a diverse set of sources

3. **Sample allocation across sources and waves:**

   a. Within source sample balancing

   b. Pragmatic sample allocation across sources

   c. Consistent sampling specifications wave-to-wave

AAPOR
AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

Pragmatic Sample Balancing

1. **Highly-nested sample balancing can be:**

   a. Expensive

   b. Impractical

   c. Difficult to implement towards the end

2. **Pragmatic sample balancing:**

   a. Gender nested with age

   b. Race-ethnicity

   c. Geography

   d. Education or Income

3. **Leave finer balancing for weighting:**

   a. Multivariate alignments can only be achieved via weighting

   b. Available profile data for sample balancing can be sketchy

Characteristics of Nonprobability Samples
Fitness for Use

1. **We estimate that there are between 6M to 12M individuals in the U.S. who are active, non-duplicated, online, non-probability survey takers.**

   a. Based on efforts to screen for various low incidence groups and typical average cooperation rates.

   b. This excludes user surveys like Amazon reviews and list surveys (for surveys like employee surveys).

2. **How are these active NPS survey takers different from the 330M individuals who live in the U.S.?**

Characteristics of Nonprobability Samples
Fitness for Use

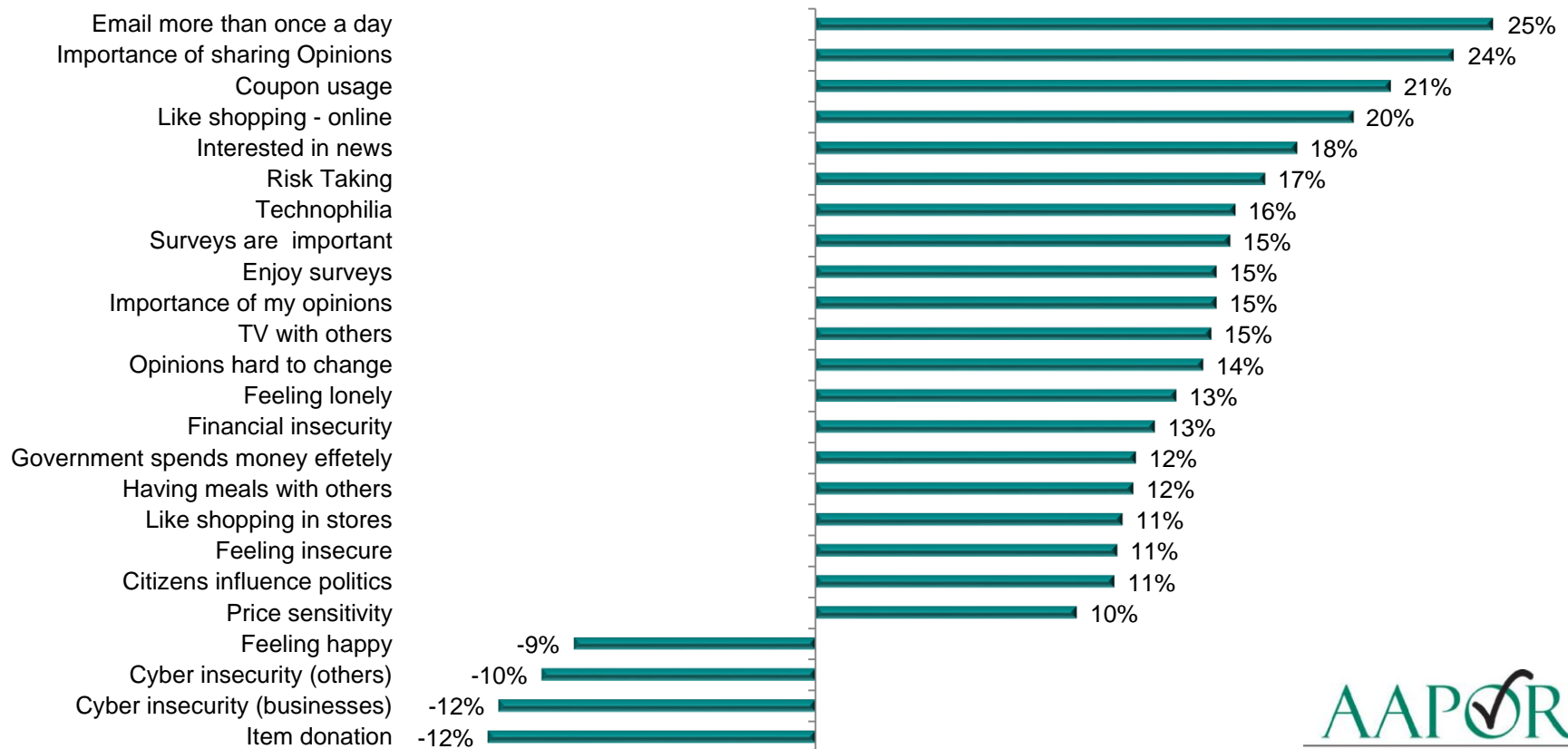**How are these NPS people different than the larger population?**

1. Lower in income

2. Higher in education

3. Fewer minorities – specifically blacks and Hispanics

4. Higher levels of technology-orientation (knowledge, purchase, etc.)

5. Higher levels of consumerism (shop more and more online)

6. Higher levels of media exposure (online more, TV more)

7. Higher in expressiveness (like to express opinions on surveys)

8. Lower levels of health (more disabled, more smokers, less activity)

9. More liberal politically

Characteristics of Nonprobability Samples
Fitness for Use

These differences from the general population lead to NPS estimates often different from population benchmarks, <u>even with demographic weighting</u>.

Significant Differences Between Nonprobability and Benchmarks

Characteristic	Value
Email more than once a day	25%
Importance of sharing Opinions	24%
Coupon usage	21%
Like shopping - online	20%
Interested in news	18%
Risk Taking	17%
Technophilia	16%
Surveys are important	15%
Enjoy surveys	15%
Importance of my opinions	15%
TV with others	15%
Opinions hard to change	14%
Feeling lonely	13%
Financial insecurity	13%
Government spends money effetely	12%
Having meals with others	12%
Like shopping in stores	11%
Feeling insecure	11%
Citizens influence politics	11%
Price sensitivity	10%
Feeling happy	-9%
Cyber insecurity (others)	-10%
Cyber insecurity (businesses)	-12%
Item donation	-12%

Pros & Cons of Nonprobability Samples
Fitness for Use

1. **Pros of Online Nonprobability Samples:**

   a. Less costly per complete

   b. Can obtain higher numbers of completes for rare subgroups

   c. Online studies can be more quickly fielded

2. **Cons of Online Nonprobability Samples:**

   a. Demographic weighting may not reduce the bias in NPS

   b. Sampling error in estimates cannot be adequately computed

3. **Fitness for Use:**

   a. When probability sample and mode (e.g., mail) takes too long, expensive, or infeasible

   b. When other considerations like cost are more important than having the lowest bias possible

   c. When interested in experimental effects

Pros & Cons of Nonprobability Samples
Fitness for Use

Type of Survey		Recontactability	Validity	Bias
Ad Hoc	Single, one-time survey	Not Applicable	Predictors can be just as high as PS, though estimates can be different	Often higher bias than PS. Can be some significant variability due to sample sources.
Tracker	Cross-sectional - same survey, different samples	Not Applicable	Can pick up changes equally well as PS over time, but estimates can be systematically different though	Often higher bias than PS, results can vary over time due to changes in sample composition, not real changes
Longitudinal	Same sample, repeated surveys	Low to Moderate	Predictors can be just as high as PS	May increase over time more than PS due to higher differential loss of respondents

AAPOR
AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

Pros & Cons of Nonprobability Samples
Fitness for Use

Topic of Survey	Comparisons with Probability Samples
Political polling	Can provide accurate results if the attitudinal biases of the sample are corrected, not just with demographic weighting
Political attitudes	Can trend similarly across time as PS, if biases are consistent over time
Advertising	NPS is more media oriented than PS, will be more likely exposed to advertising, so recognition is higher
New product concepts	NPS is more consumer oriented. Tends to like new products more than PS, but is consistent across products
Financial	Can vary depending on topics. Harder to obtain high net worth individuals. NPS higher in part-time workers.
Health	NPS can have more negative health behaviors (smoking, sitting, drug use, etc.)
Technology	NPS higher than PS in technology orientation, purchase and knowledge of technology overall
Brand positioning	Brands are most often rated in similar order for both NPS and PS, though there can be absolute differences

AAPOR
AMERICAN ASSOCIATION FOR PUBLIC OPINION RESEARCH

Data Quality

1. **Many researchers review data to identify and perhaps remove poor quality respondents.**

2. **Examples of suboptimal response include:**
   a. Extent of non-differentiation or straight-lining in grids
   b. Speeding
   c. Item nonresponse
   d. Trap failure
   e. Consistency checks

3. **The ARF FOQ2 study with 17 different sample providers affords an opportunity to compare suboptimal response among providers.**

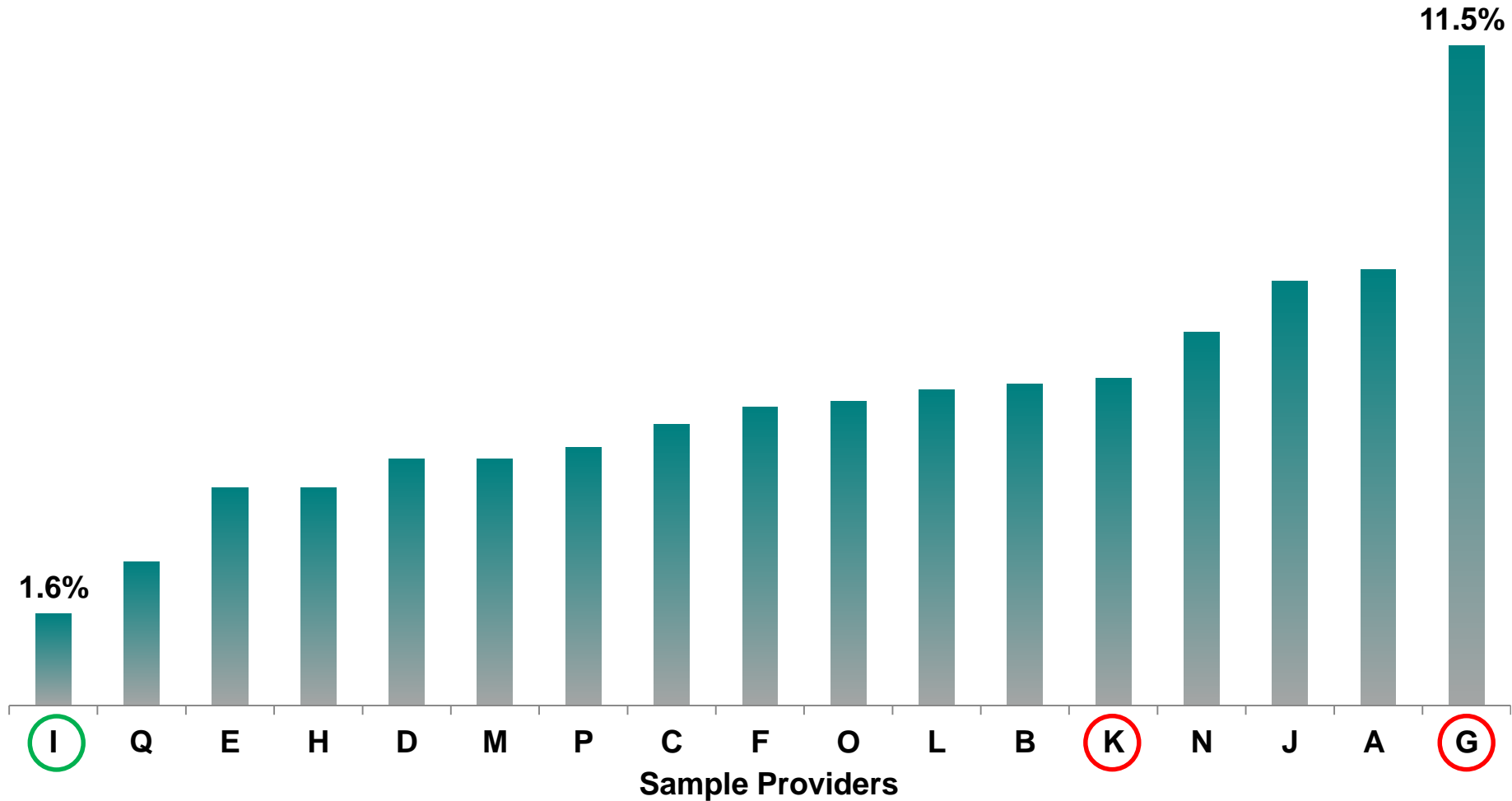# ARF FOQ2
Percent Speeders
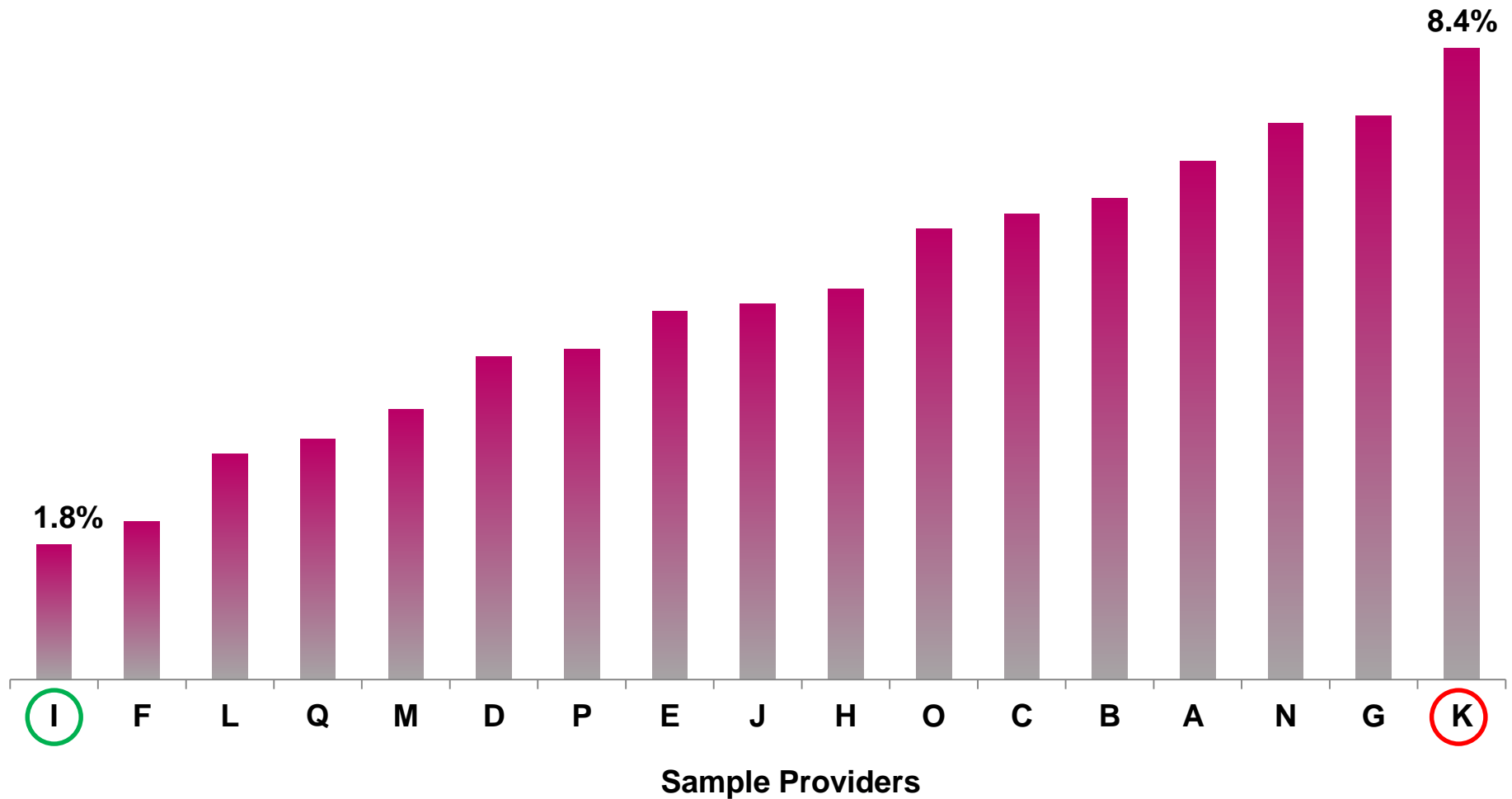
Sample Providers

Source: Thomas, R.K. (2014). Fast and furious or much ado about nothing? Sub-optimal respondent behavior and data quality. *Journal of Advertising Research Foundation*, 54(1), 17-31.
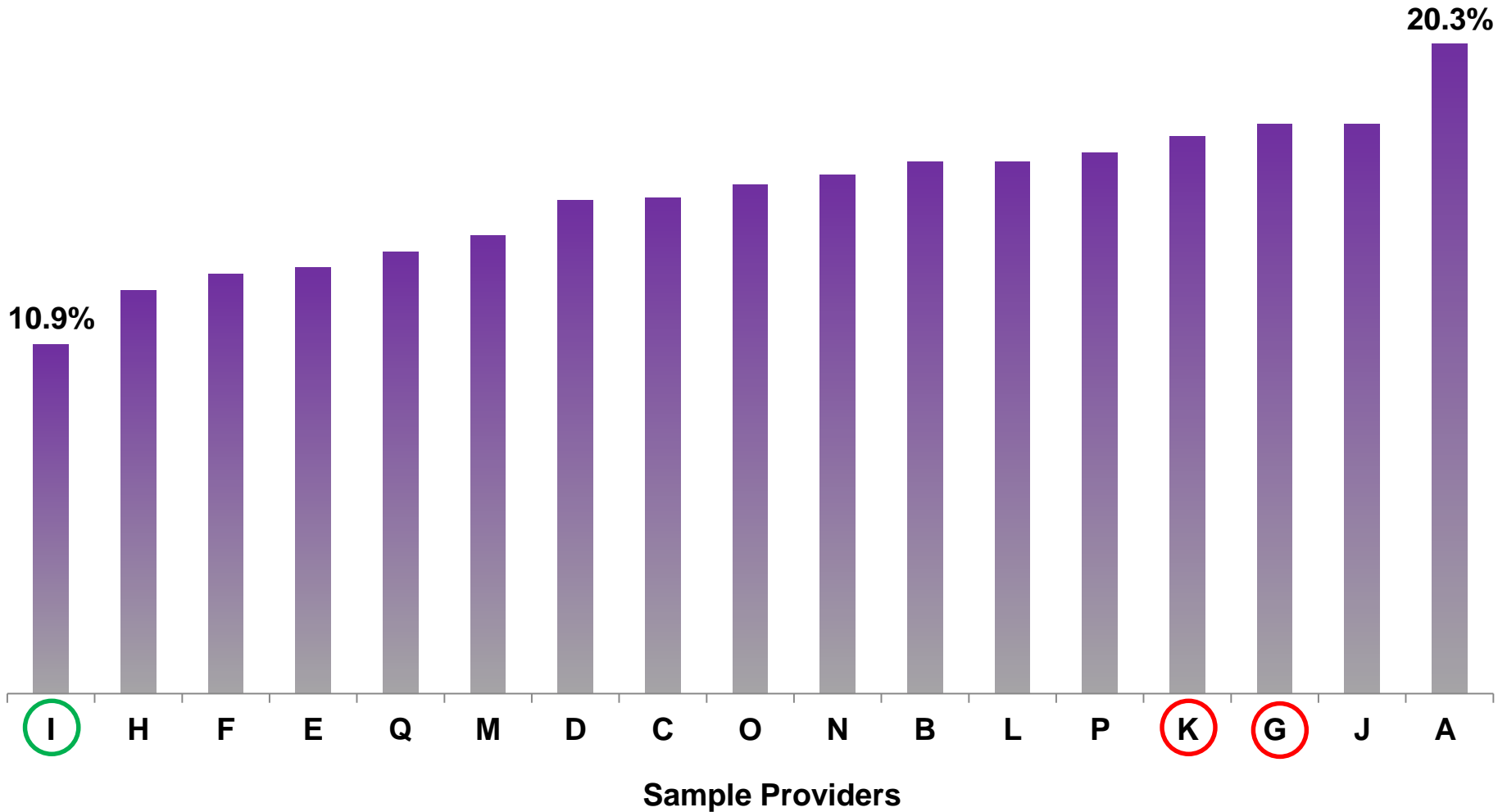
ARF FOQ2
Non-differentiation in 5 or More Grids

Source: Thomas, R.K. (2014). Fast and furious or much ado about nothing? Sub-optimal respondent behavior and data quality. *Journal of Advertising Research Foundation*, 54(1), 17-31.

ARF FOQ2
At Least 1 Trap Question Failure

Source: Thomas, R.K. (2014). Fast and furious or much ado about nothing? Sub-optimal respondent behavior and data quality. *Journal of Advertising Research Foundation*, 54(1), 17-31.

Data Quality

1. **We often do not use provider variable in analyses, but given variation in quality, it is important to pay careful attention to quality among the mix of providers you are using.**

2. **This can have cost and timing implications as replacement sample will be required if cleaning to remove respondents is being done.**

3. **We have found that quality is not always correlated with provider costs.**

Assessing Quality

1. **Extent of bias not necessarily related to suboptimal response.**

2. **Recommend selecting a small but diverse set of items with external benchmarks from high quality federally-funded, large-scale surveys, if possible for target population.**

3. **Examples include:**

    a. Secondary demographics – marital status, household size

    b. ACS – e.g., number of bedrooms or vehicles

    c. Health studies – e.g., NHIS, NSDUH, BRFSS

    d. CPS Supplements – e.g., Civic Engagement Supplement

4. **Calculate average absolute deviation from benchmarks to assess extent of bias and impact of weighting.**

Bots
Automated Respondent

Bots are an issue with nonprobability sample – generally do not affect probability sample where there is more control over panel recruitment and survey invitations.

THE ULTIMATE SURVEY BOT V2.0!

MAKING $$$ HAS NEVER BEEN THIS EASY!

DOWNLOAD NOW

SURVEY

ULTIMATE SURVEY BOT V2.0

Brought to life by **Bots4U™** a world class development team based out of the US comes one of their greatest brain storms to date.

With improved visuals, new undetectable mode, redesigned algorithms and now supporting surveys from platforms like: **Paymentwall, RadiumOme, SuperRewards, SuperSonic, Trialpay and Persona.ly.**

The Ultimate Survey Bot V2.0 is unmistakably setting itself up to be the single greatest **automated cash machine** of the decade!

The Ultimate Survey Bot V2.0 is unmistakably setting itself up to be the single greatest **automated cash machine** of the decade!

Bots
Prevention Efforts

Sample providers are working to prevent bots with:

1. **Identity verification**

   a. Validate name and address

   b. Mail incentives to physical mailing address

2. **Digital fingerprinting**

   a. IP address; cookies; system and browser information

   b. Remove duplicates

3. **reCAPTCHA**

   a. Can be used to verify that respondent is not a bot

   b. Need to be cautious that we don't fatigue respondents by asking them to prove that they are human too many times – coordinate with providers
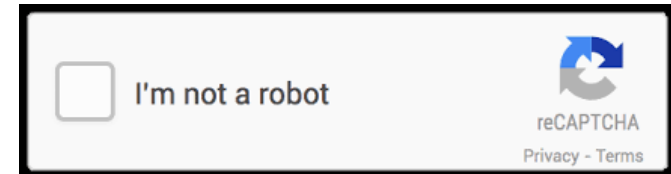
Bots
Detection Efforts

There are a couple of ways that bots can be detected:

1. Duplication

a.   The least sophisticated bots provide a repeating response pattern either the entire way through or in beginning of survey to get through screening

b.   Check for duplicates

2. Illogical response patterns

a.   A large number of 18 year olds with long history of smoking

b.   A large number of shoppers at a store that does not exist in their area

3. Repeated open ends

a.   Some bots alter responses to closed-ended questions but repeat open-ended responses

b.   Sort open-ends alphabetically and look for duplicates

Examples of Bots

In a few sentences, please describe what was most important to you when making a final decision on the smartphone you recently purchased.

1. **39 instances of "Is a excellent company"**

   a. Varying IP address, though some repetition

   b. All had same user agent string
      Mozilla/5.0 (Windows NT 10.0; Win64; x64; rv:52.0) Gecko/20100101 Firefox/52.0

   c. Same age and gender

   d. Varied between 2 states and 2 education levels

2. **Caution with more common types of responses, such as "I like it", "Don't know", "N/A"**

   a. Review demographics and paradata for patterns

   b. With such responses, we would expect them to be fairly evenly distributed across providers if multiple sample providers, red flag when they all come from same provider

Peeking Outside of the Box
The Divide

1. **The times, they are a-changing:**

   a. Traditional methods have issues – coverage, nonresponse, and cost

   b. Alternative methods are emerging and becoming more effective

   c. Informational needs are evolving

   d. Digital/Big data are becoming ubiquitous

2. **Ultra right (market research):**

   a. Traditional survey research is dead

   b. Data should be acquired digitally or via fusion

   c. Structured survey data miss key nuances such as POS information

3. **Ultra left (survey research):**

   a. Nonprobability samples are void of inferential integrity

   b. Market research methods are beneath survey research

   c. Weighting can make up for the inadequacies of sample surveys

Peeking Outside of the Box
Closing the Divide

1. ## The times, they are a-changing:

   a. Business as usual would not work

   b. Perfect coverage and high response rates are gone

   c. The days of lengthy surveys are over

   d. Some data are not securable via structured surveys

   e. Alternative methods and data sources are here to grow

   f. But don't throw out the baby with the bathwater

2. ## Centrism:

   a. Fundamental strengths of traditional methods need to be respected

   b. Potential gains of alternative methods and data sources need to be recognized

   c. Shortfalls of alternative methods have to be dealt with transparently

   d. Pragmatic amalgamation of data to support evolving informational needs:

   ➢ Shorter surveys based on solid principals

   ➢ Supplementary data carefully incorporated to complement the above

Thank you!

Mansour Fahimi
mansour.fahimi@gfk.com

Frances M. Barlas
frances.barlas@gfk.com

Randall K. Thomas
randall.thomas@gfk.com