

Fall 2013

GRIT Consumer Participation in Research Report

A Study of Survey-Takers in 200+ Countries and Regions
around the World

GREENBOOK.

www.GreenBook.org/GRITCPR



QUALTRICS INSIGHT SUMMIT

February 18-20, 2014

Salt Lake City, UT



**Customer, Market and Employee Insights.
Make data-driven decisions.**

Register NOW!
www.qualtrics.com/summit
beright@qualtrics.com
801-669-5380



Table of Contents

Contacts

Leonard F. Murphy

Chief Editor & Principal Consultant
(770) 985-4904
lmurphy@greenbook.org

Lukas Pospichal

Managing Director
(212) 849-2753
lpospichal@greenbook.org

Diane Liebenson

Publisher, GreenBook
(310) 545-0918
dliebenson@greenbook.org

GreenBook

New York AMA Communication Services Inc.
116 East 27th Street, Floor 6
New York, NY 10016

Introduction	3
Executive Summary (Infographic)	5
About the Report	5
Detailed Findings	6
Respondent Breakdown	6
Respondent Freshness	7
Respondent Compensation	9
Operating Systems, Desktop and Mobile Status	11
Major Limitations	13
Conclusion	14
Toward the Future	15
Meta-Data Inferential Analytics	15
Improving Geo-Location	17
Appendix	18
About GreenBook	20

Complete GRIT CPR data sets for 200+ countries and regions are available at www.GreenBook.org/GRITCPR

Explore, add comments, analyze or benchmark your own data sets. Enjoy!

Introduction

Respondents are the lifeblood of market research. Whether it's qual or quant, surveys or communities, neuromarketing or 'Big Data' and everything in between, knowing how to reach, engage, and understand people is the very bedrock of insights.

In our interconnected world, achieving that goal is in some ways easier – but in many more ways it is harder – and little data has existed to help researchers understand this basic question: *How do we get consumers to engage with us, and what do those consumers look like?*

Until now.

The GRIT Consumer Participation in Research (CPR) Report is our effort to answer the who, what, when, where, and why of global consumer participation. Based on millions of data points across over 100,000 interviews, it uses a revolutionary random sampling micro-survey model. It was developed by The RIWI Corporation – winner of the Insight Innovation Competition 2013 in North America.

GRIT CPR is the tool that researchers can use to benchmark their efforts in over 200 countries and start crafting new strategies to grow consumer participation.

The report includes an in-depth explanation of the methodology, and references at the end to answer any questions related to the project. Our goal was to utilize a new tool to help insights professionals better understand the state of respondent participation globally. We are grateful to our partners at RIWI for volunteering their time and technology to help us reach this goal. Without them it would have been impossible due to time, expense, and complexity.

In keeping with the expectations established by the annual GRIT Report, all of the data is available via an online dashboard generously provided by our partners at Dapresy.

You can access that here: www.GreenBook.org/GRITCPR

Connected to the dashboard is an online visual summary; we encourage you to visit it and share it with your colleagues.

As always, we couldn't have accomplished anything without the contributions of many. Thanks go to the entire RIWI team, especially Neil Seeman, CEO, and Alton Ing, CTO, for their tremendous efforts here. Thanks also to Rudy Nadilo, President of Dapresy North America, and his team, for their generous support in developing the dashboard and summary. Heartfelt appreciation goes out to our sponsors and partners who support our industry initiatives again and again.

GRIT CPR is a tool that is long overdue and we hope you find much value in it.

Best wishes,

Leonard Murphy

Editor-In-Chief, GreenBook

Expand Your Reach Through Civicom's Global Expertise

Civicom®

Ask Why We're Better

Innovative Technology Solutions Marketing Research Support Worldwide



CONSUMER PANEL in the U.S.

Research on CPGs using respondents from panel of innovative professional women.



TEXT ANALYTICS in the U.S.

Data extraction and text analytics from transcriptions of focus groups.



IDs in Africa

Moderator conducts phone interviews to get better insights on the right target market in Africa.



WEBCAM INTERVIEWS in the U.K.

Moderator conducts webcam interviews with Moms and kids.



YOUR CLIENTS in Canada

Observing your live online focus groups.



SIMULTANEOUS TRANSLATION in Western Europe

Research on global advertising in five countries via webcam interviews.



ONLINE COMMUNITY in China

Teens from China join an online community to evaluate online courses.



RETAIL ANALYSIS in Asia

Working with mobile phones as the main means for sending data. Respondents receive and send text messages with ease using their own mobile devices.

DOCTORS IN RESEARCH in North America

Doctors participate in studies on the phone at their convenience.



MODERATORS in North America

Conducting IDs from their U.S. Office with respondents in Russia.



FOCUS GROUPS in the Middle East

Overcoming cultural barriers to involve female respondents from the Middle East. Moderators get the women online and on the phone for a focus group.



MOBILE QUALITATIVE in South America

Following day-to-day activities of the Hispanic community.



PRODUCT TESTING in Australia

Male respondents from Australia go online for a bulletin board focus group where they post videos of themselves testing the product and sharing their thoughts with the group.

About the GRIT Consumer Participation in Research Report

How do everyday consumers from every country around the world – that is, average, *non-incented* Internet users, not just members of a panel or a recruited online community – feel about the experience of being surveyed? Do they prefer being paid or not? Will they even answer if they are non-incented? How frequently do they answer surveys? From what kind of device – desktop or mobile – do these respondents prefer to answer? What browsers and operating systems do they use? How do over 200 different countries and regions of the world compare on these metrics? Are there similarities or differences in response rates? Is there a statistical inter-country concordance for one’s desire to take paid surveys? Is there a different frequency of survey participation? Is there a difference in mobile vs. desktop response? How much? Is there variation in mobile browser vs. desktop browser or OS?

If we look at sub-groups of countries, such as the major English-speaking countries, is there a statistical correlation between those exposed to Internet-based GRIT CPR questions vs. the numbers of people who answer those non-incented questions? What do we know about the attributes of the people who respond and what can be known, in a privacy-compliant manner (using proprietary inferred meta-analytics); about attributes such as education level, ethnicity, and on whether or not they have children? Is there geographic representativeness of the data in selected countries? To what extent do the findings correspond with external, third party information?

One goal of the report was to answer these questions for the world (presenting limitations to the best of our ability), while offering researchers the opportunity to supplement the data with their own insights and complementary data – so-called ‘data philanthropy’. This goal also includes suggesting norms against which researchers can assess their own data sets in different geographies. For example, researchers can understand their data sets better in the context of data presented here concerning mobile, desktop, and browser usage among non-incented, country-specific respondents.

GreenBook and The RIWI Corporation, in this first annual GRIT Consumer Participation in Research (CPR) report, sought to answer these questions, and present their findings to market researchers, insights professionals, and data scientists around the world.

Sample question exposed to a non-incented respondent in the GRIT CPR:

Exit Now

Prior to this survey, when was the last time you answered survey questions?

Past day

Past week

Past month

Over a month ago

Never

How frequently do they answer surveys?



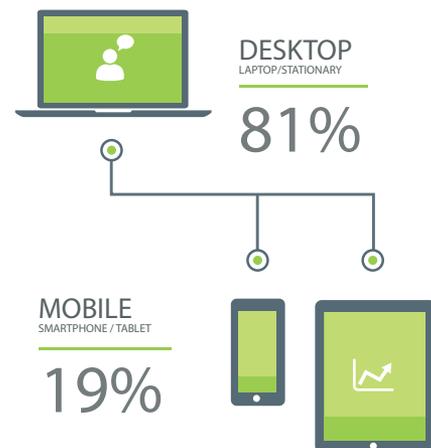
Would you like to get PAID for taking surveys?



YES
I WOULD LIKE TO GET PAID

52%

What kind of device do they answer surveys from?



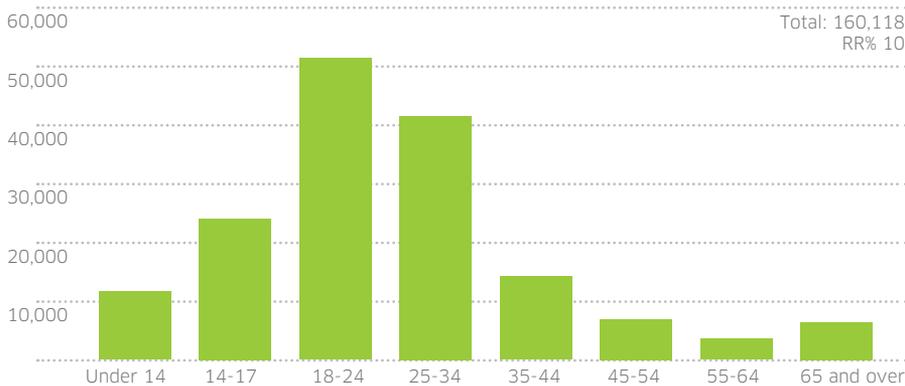
CPR Findings

Age and gender breakdown of respondents

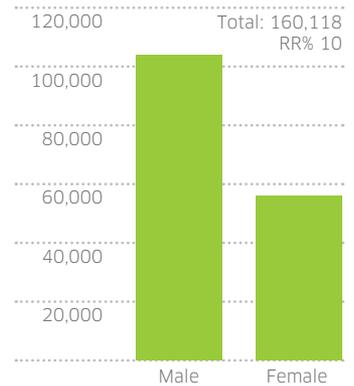
Between July 1 and August 18, 2013 1,587,281 respondents on every Internet browser and every type of Web-enabled device were exposed to a short suite of non-incented questions about the survey-taking experience using a micro-survey model delivered via random domain intercept technology (RDIT™).

GRIT CPR received a 10% response rate, for a total of 160,118 people, who described their age and gender in 231 regions and countries of the world. The highlighted global results and major findings in topical section are presented here below, skewed toward the 34 and under population, who are more intensely Web-enabled and spend much more time on the Internet. Weighted results (weighted to the most current Census data per country) are also provided in the bar charts that follow.

Q1a - What is your age group?



Q1b - What is your gender?



Items of significance about the GRIT CPR age and gender findings:

» 81% of the un-weighted age data show the response population to be 34 years of age and younger. The data presented here therefore suggests that 54% of the Internet population, as opposed to 45% estimated by the ITU Telecom World based on 2011 statistics, is skewed young (ITU data refer to the 25 and under population). CPR data for this age group is high since they take into account not only Web access, but also the frequency of Web usage by this younger age cohort. There are likely other differences for this gap. Firstly, the ITU data looked at usage in 34 countries, while RDIT data examined data in every country and region of the world. Further, the two-year time differential is substantial, given the dramatic rise in tablet, smartphone, broadband, and mobile usage.

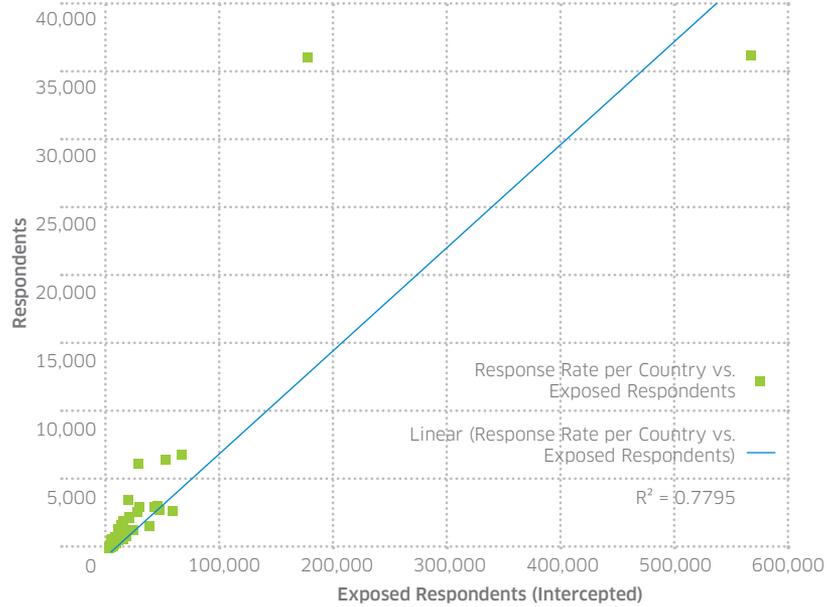
» Almost double (65% vs. 35%) of respondents self-identified as male. The non-incented nature of the RDIT approach captures the fact that men are heavier users of the Internet, and greater users of a greater variety of sites^{###}, and are thusly more likely than women, statistically, to 'stumble upon' a survey landing page. Further, this being a global survey, it is important to recognize that in most countries, men have substantially higher socio-economic status, and are more likely employed in workplace situations where Internet is accessible.

» The response rate to this age/gender question (10% at a global level) is very high by comparable standards to 'best-in-class' Random Digit Dialing or even panelized respondents. This suggests that non-incented respondent Web intercept is a powerful tool to reach a large number of new and potentially less 'jaded' respondents.

GRIT CPR received a 10% response rate, for a total of 160,118 people, who described their age and gender in 231 regions and countries

» There is a relatively strong correlation (0.78/1) between the visitors intercepted on this opening 'engagement question' and the people answering it per country/region (with two outlier countries, each diamond representing one of 231 regions or countries). This suggests statistical consistency on an inter-country level in the methodology.

Visitors to the RIWI landing page on age and gender questions



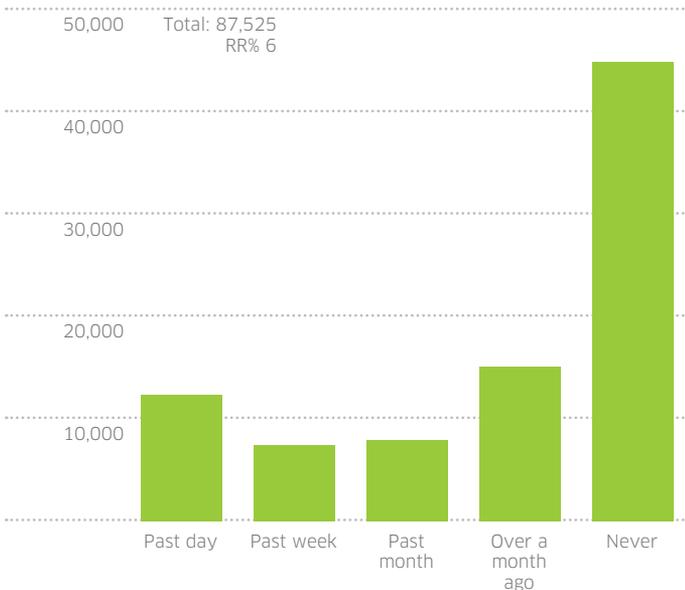
Almost double of respondents self-identified as male

Respondent “freshness”: How often have respondents answered surveys?

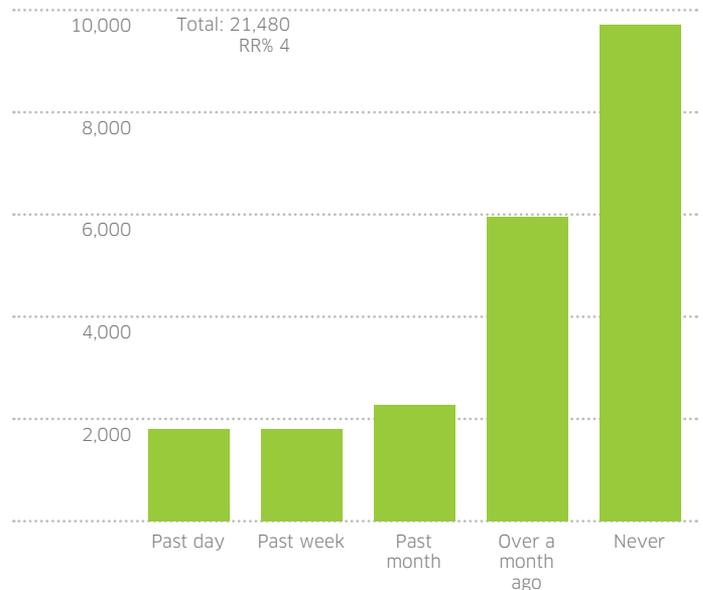
Scholarly literature on the topic concurs that first-time responders (what we refer to as “fresh responders”) provide more reliable replies to surveys than practiced responders do. In addition to this, work by Toepoel et al¹⁴⁴, (2008, 2009) has shown that trained responders tend to take shortcuts when it comes to answering surveys, as well as pay less attention to the questions being asked of them.

Q2 - Prior to this survey, when was the last time you answered survey questions?

Global data



US data



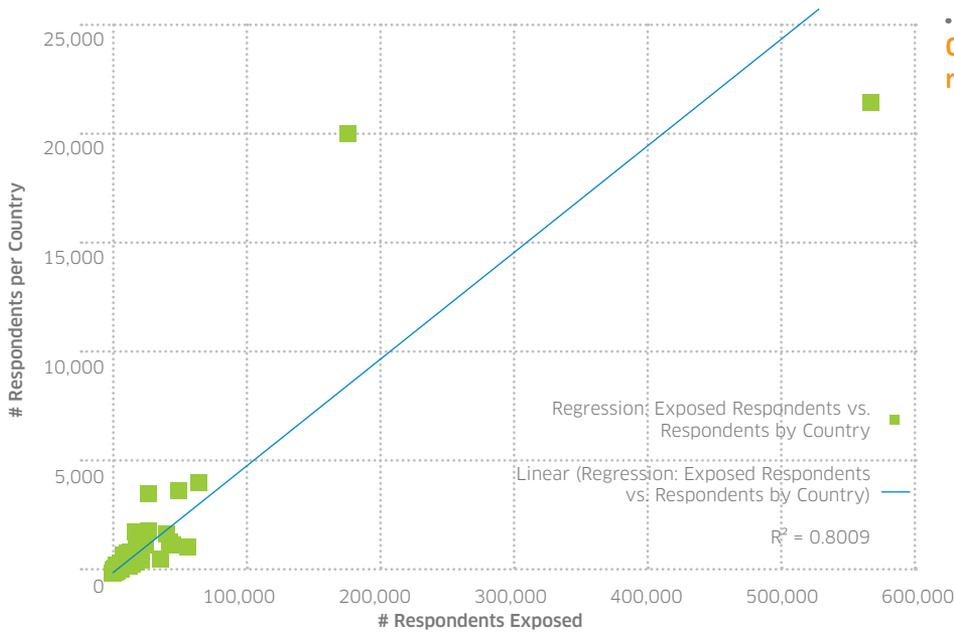
Items of significance about the CPR respondent “freshness” findings:

» CPR examined 231 countries or regions of the world (a region is an area of the world that does not qualify for United Nations designation as a nation state) to examine whether the respondent had answered his or her last survey (i) in the past day; (ii) in the past week; (iii) in the past month; (iv) over one month ago; or (v) never. The total response size to this question globally was 87,525, a response rate of 6%.

» We determined that a “fresh respondent” was someone, out of a total of 87,525 respondents to this question, who answered “over one month ago” or “never”; by this definition, on a global scale, 68% of CPR respondents were “fresh”.

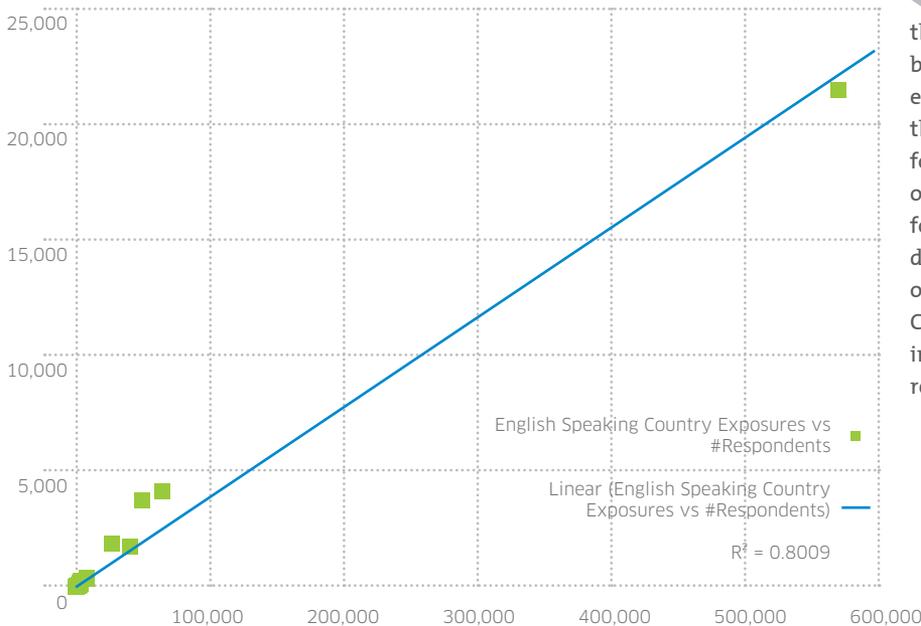
» In the United States, looking at a sample of 21,480 respondents, 73% of respondents qualified as “fresh”, a statistically significant difference.

» In terms of statistical power, there was a strong correlation between the percentage of respondents in each country or region and the percentage that answered. The average response rate for the 87,525 respondents was 6%. In terms of an R-squared correlation, the following figure illustrates that it was strong:



On a global scale, 68% of CPR respondents were “fresh”.

English Speaking Country Exposures vs # Respondents



» When looking at just the English-speaking nations and regions (N=42), there was an even stronger correlation between the percentage of respondents in each country or region and the percentage that answered. The average response rate for the 35,496 respondents was 6%. In terms of an R-squared correlation, the figure following illustrates that it was 0.99/1. The definition of an “English-speaking country or region” was one where the most recent Census data available from that area indicated that 70% or more of the country/region spoke English comfortably.”

Respondent Compensation: Some Surprising Findings

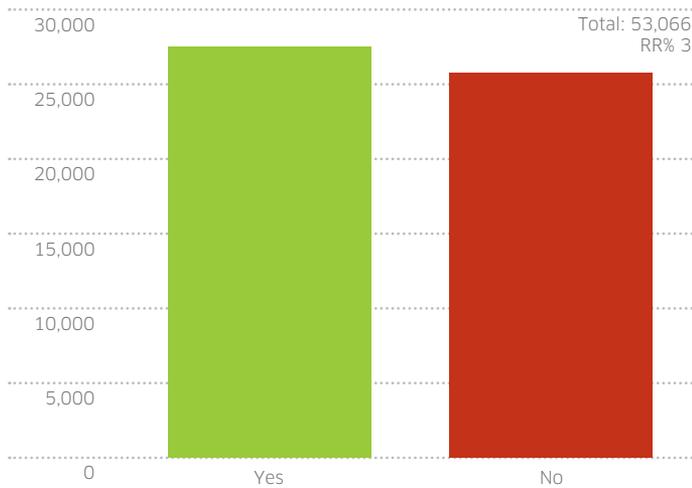
It is a common assumption in data collection that people prefer to be compensated than not when taking surveys. This assumption is built on another assumption that survey-takers will be undertaking long, onerous 25 minute or longer surveys. The GRIT CPR findings compel researchers to re-examine this assumption, at least in the context of short surveys for non-incented respondents, and to look at this issue in a gendered, geo-specific, and age-related context.

48% of respondents said they preferred not to be paid to take surveys

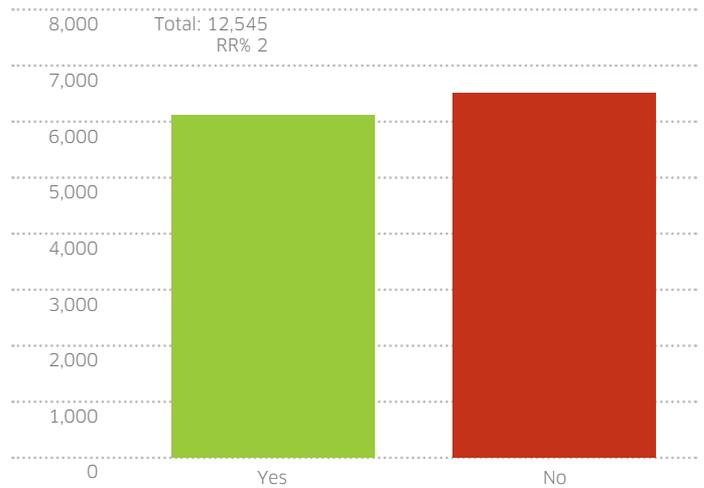
48% of respondents said they preferred not to be paid to take surveys

Q3 - Would you like to get PAID for taking surveys?

Global un-weighted data

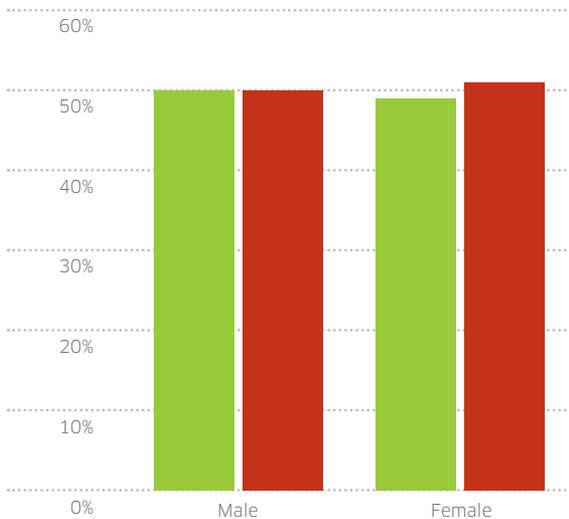


US un-weighted data



Q3 - Would you like to get PAID for taking surveys?

Global Results by Gender



Global Results by Age Group



Items of significance about the CPR respondent compensation data:

At an aggregate level, with a total sample of 53,066 respondents to this question, 48% of respondents said they preferred not to be paid to take surveys. Perhaps the salience of the questions was what mattered to respondents. Our other hypothesis was that, since this was a global survey, it was possible that many respondents were unfamiliar with the standard compensation approaches available to survey respondents in the United States, Canada, or the UK. The margin of error was 0.4%. However, this latter interpretation did not make complete sense to us, since, even in the United States (N=12,545) where paid compensation of some form or another is commonplace for surveys, a slightly higher proportion, 52% said they preferred to give their opinions for no money (un-weighted).

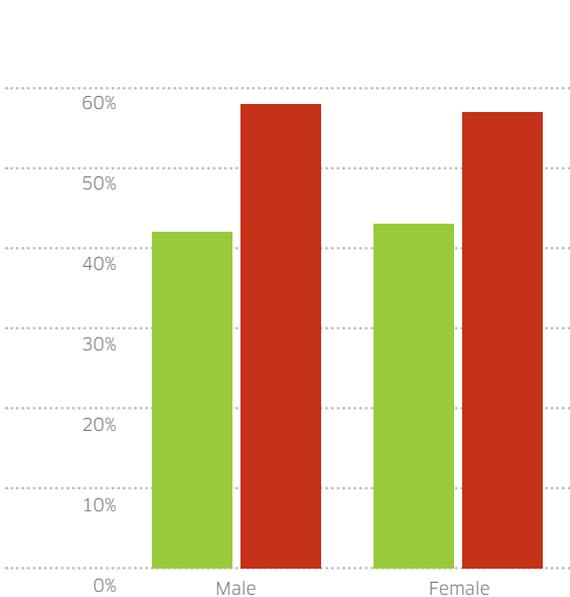
Looking at the weighted data (adjusted to Census age groups), social desirability bias is not the likely explanation for the above finding. More likely, this issue requires a nuanced understanding of how different genders and age groups respond to this issue in specific countries. On the aggregate level, there is no statistical difference between whether men and women want to be paid for survey taking; 50% do want to be paid; 50% do not. However, in the United States, a statistical majority of potential respondents using the RDIT™ methodology do not want to be paid. This is consistent for both men and women; 57% of women say they do not want to be paid; 58% of men report they do not want to be paid (weighted data). This was deduced from a respondent sample of 12,545 US-based respondents. In the United States, there was only one age group – 18-24 – where the weighted data suggested that there was an even split between survey respondents who preferred payment vs. those who did not. The most significant difference was in the 65 and over cohort, where 64% of respondents did not want to be paid.

57% of women say they do not want to be paid; 58% of men report they do not want to be paid

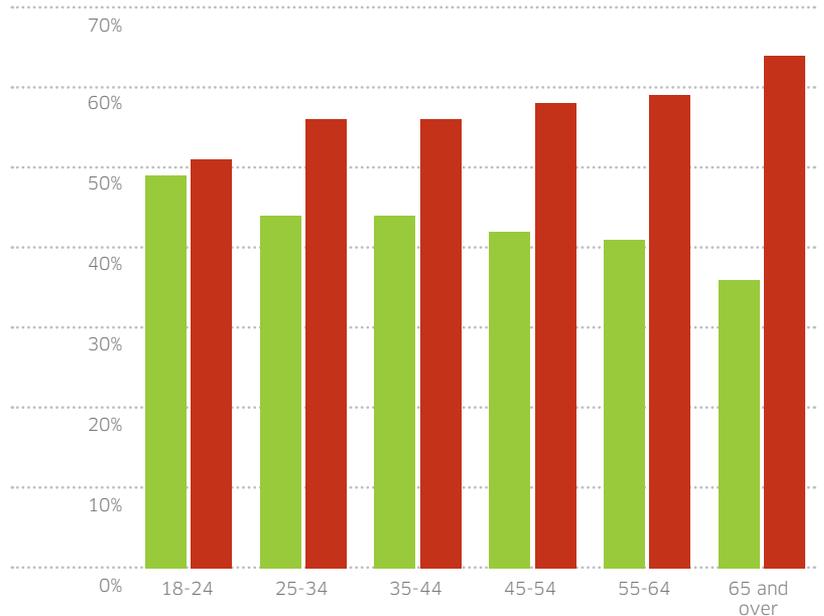
57% of women say they do not want to be paid; 58% of men report they do not want to be paid

Q3 - Would you like to get PAID for taking surveys?

US Case Study Results by Gender



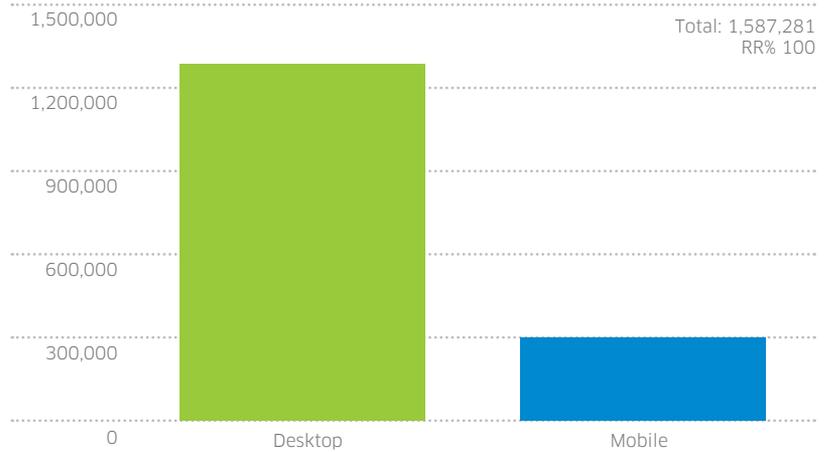
US Case Study Results by Age group



Operating systems, desktop and mobile status of GRIT CPR respondents

Of the 1,587,281 exposed respondents for whom we could gather such information reliably using meta-data analytics, 19% of GRIT CPR respondents globally were using a mobile phone. This suggests, and we believe it to be the case, that current mobile usage (combining smartphones and tablets) vs. desktop usage is significantly higher than previous reports. Those reports have put it at 13.2%^{vii} as of February, 2013, but the latter statistics were limited to websites monitored by a company for its specific clients, not for the general global Web population – the focus of the GRIT CPR. Our findings make the case that mobile survey-based applications, especially those like the Nano-Survey™ RIWI system applied for the GRIT CPR, make for a better respondent experience. More people will respond to a short survey on a mobile device if the survey interface renders large. The bar charts below describing operating systems – on mobile and desktop – may be compared to more dated (2011 and 2012) statistics provided elsewhere^{viii}.

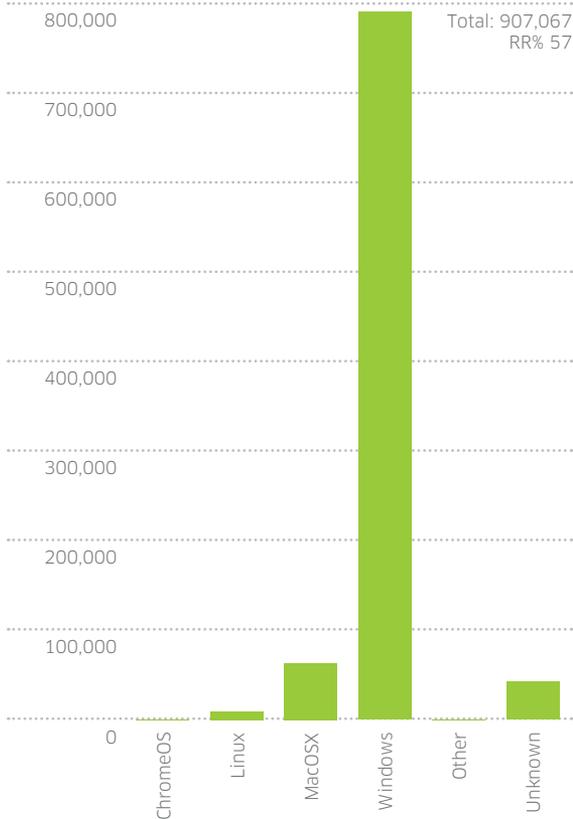
Types of Devices Used



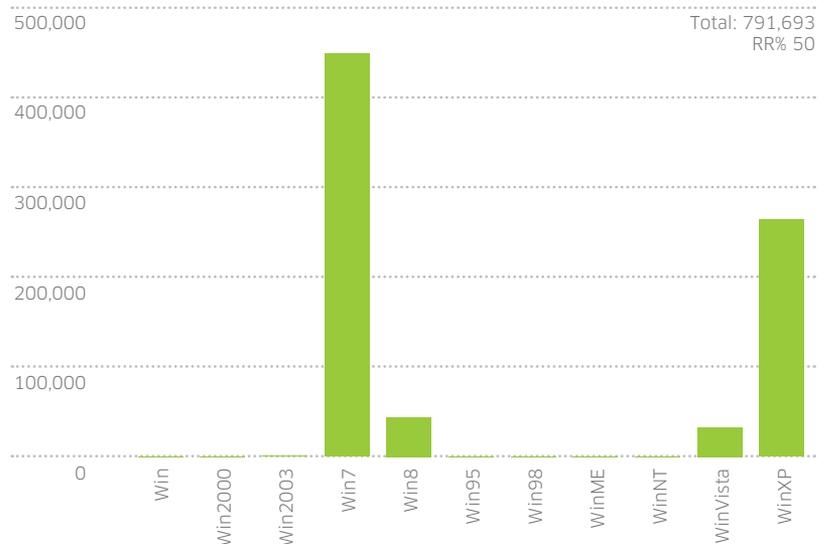
19% of GRIT CPR respondents globally were using a mobile phone

More people will respond to a short survey on a mobile device if the survey interface renders large

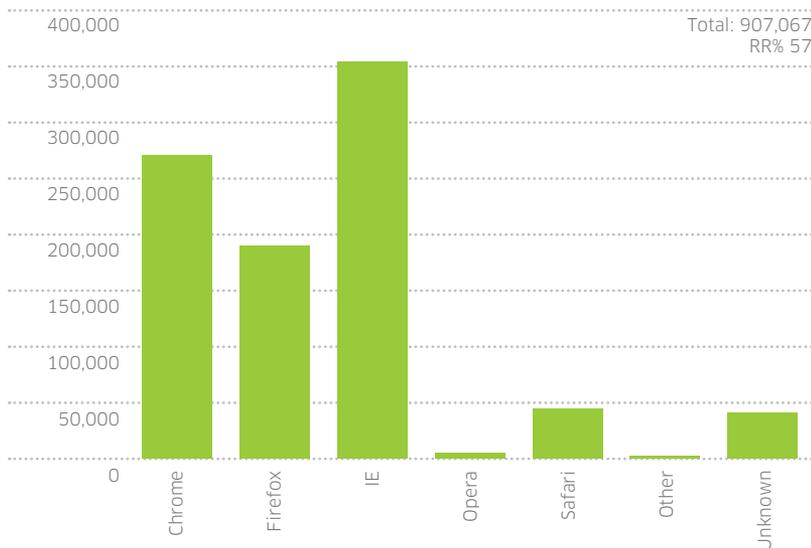
Types of Desktop Operating System Used



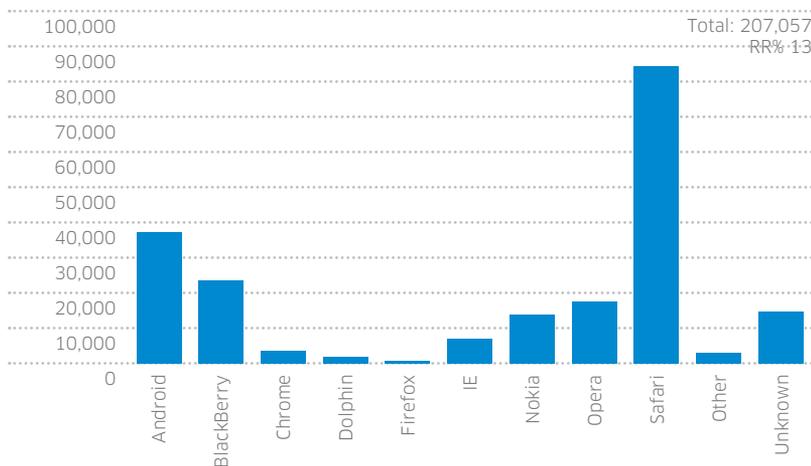
Types of Desktop Windows Versions Used



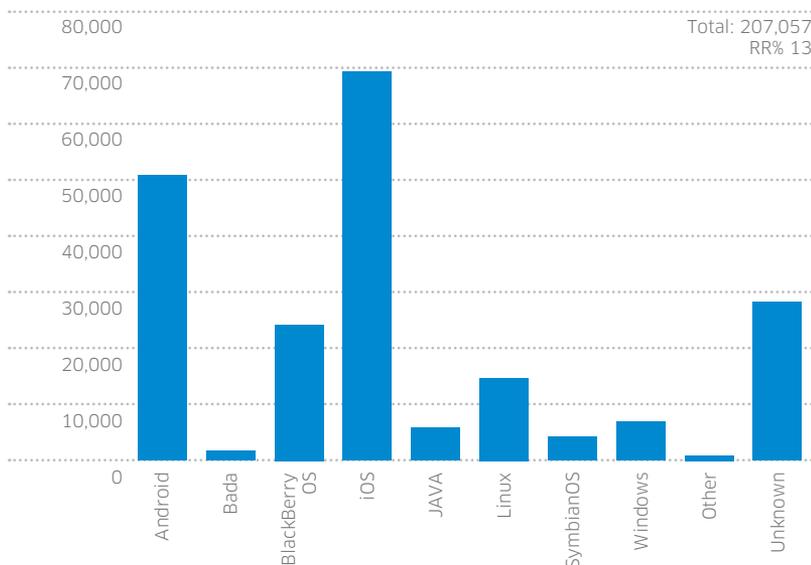
Types of Desktop Browsers Used



Types of Mobile Browsers Used



Types of Mobile Operating Systems Used



Items of significance about the CPR respondent OS, browser data:

» The world results for the CPR show that Windows (87%) is by far the most popular desktop operating system for CPR respondents, in concordance with other statistics³⁴. Our GRIT CPR data also correspond with other external market share data showing that, in order of popularity, CPR respondents around the world use Windows 7 (57%), followed next by Windows XP (33%), then Windows 8 (6%); and lastly, Windows Vista (4%)³⁵.

» Looking at the global CPR respondents, the leading desktop browser is Internet Explorer (39%), followed by Chrome (30%), and then Firefox (21%). Perhaps most surprising is that Firefox users, despite the attempt by Mozilla to downsize the use of the direct navigation bar, are in fact, slightly overrepresented in CPR as compared to Firefox's global share of the browser market (18% as of the time of writing).

» The US respondent results for the CPR show that Windows (80%) is by far the most popular desktop operating system for CPR respondents, while the Mac OS system is next with 14%.

» For CPR respondents in the US on desktop browsers, Internet Explorer leads (47%), followed by Chrome (22%), Firefox (16%), and Safari (9%).

» In terms of mobile respondents, they constituted 19% of our global CPR respondents, as compared to 18% of US-based CPR respondents. In the US, the leading mobile OS was iOS (51%), followed by Android (25%), then by Linux (11%). Therefore, iOS was overrepresented, while Android was underrepresented, in the CPR respondent pool, as compared to their US market share³⁶ (most recent comScore data, 2013). This compared, globally, to iOS being the leading mobile platform (33%), followed by Android (25%), then by BlackBerry (12%). Here again, iOS was heavily overrepresented, and Android dramatically underrepresented, compared to third party data (International Data Corporation, 2013)³⁷.

» Across all global respondents, the leading mobile browser was Safari (41%); followed by Android (18%); then BlackBerry (11%). This supports the thesis that the iPhone or iPod running Safari may be more conducive to browsing than other devices. In the US, the BlackBerry system stands at a low 3% of CPR respondents³⁸. Mobile-based CPR respondents in the US were using the Safari browser (62%), Android (22%), while BlackBerry and Internet Explorer were in a distant tie for third (at 3%) in terms of Internet browser.

Major limitations of the GRIT CPR study

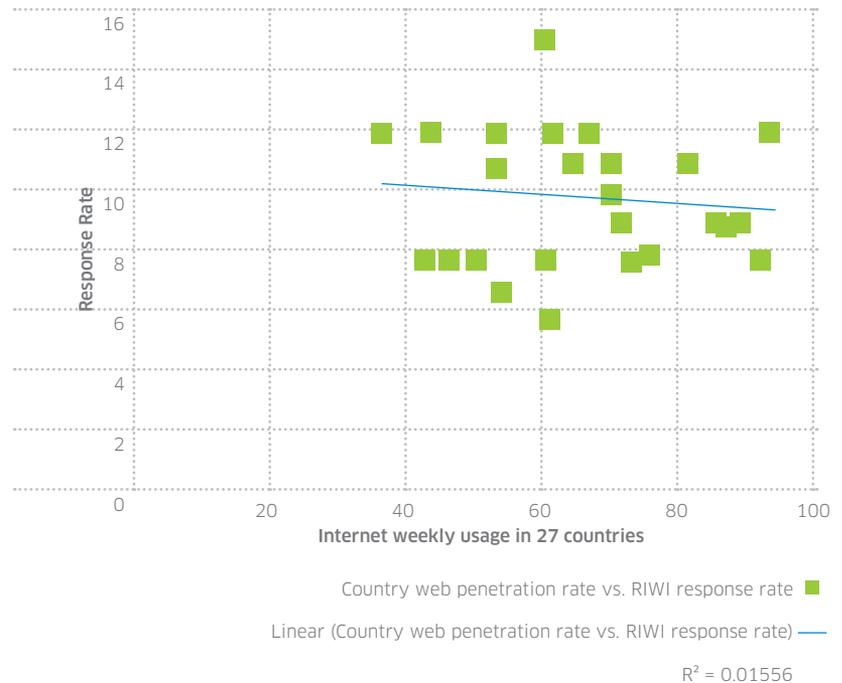
Limitations include the younger skew of the Internet user population, and the relative intensity of Internet usage that varies among countries. Both of these limitations can be mitigated via respondent population re-weighting to Census data and through large samples of all age categories exposed to the questions. It is important to note, perhaps surprisingly, that there is no statistically significant positive co-linear relationship (based on the most recent comparable 2012 statistics available for a very small subset of countries) between weekly Internet usage for the 16-74 population and the CPR response country per country (0.015 R-squared). One possible explanation is that Internet penetration does not, in fact, affect response rate and is the wrong variable of significance. Indeed, there are differences in response rates that vary among countries for cultural and topic-related reasons. In many less developed countries, potential respondents are less used to answering questions and therefore do so happily if the surveys are simple and engaging. The response rates are also higher when the survey process is devoid of malware warnings and the survey is very fast and intuitive for the respondent.

The figure that follows shows that Internet penetration and usage does not appear to affect response rate for the CPR, but it does show that response rates on the opening engagement question are strong and of very limited variability in the 27 countries – ranging from 8 to 15%. These countries enjoy comparable statistics on weekly Internet usage (for both sexes) for the 16-74 year old population.

It is important to note that the methodology for collecting the comparable UNECE data²⁴ are self-reported, panel-based (and therefore incentivized), and not based on real-time actual Internet activity for all ages as measured by the RIWI system. Like television viewing, there is a general social desirability bias in underreporting use of the Internet in our opinion. Note that the UNECE Statistical Database includes 27 countries with a mix of developed and developing nations in terms of broadband take-up and socio-economic composition²⁵.

The largest methodological consideration for this study was that it was conducted in English only.

Country web penetration rate (UNECE data for 27 countries) vs. RIWI response rate



.....
Internet penetration does not affect response rate and is the wrong variable of significance

One rationale for this English-only approach was to ensure methodological consistency across countries, and we highlight the 42 countries/regions in the world where, according to Census data, over 70% of the population speaks English comfortably. Future versions of this report will offer multi-lingual options for respondents.

Another rationale for the current approach was that inferential analytics on the nature of the people exposed to the questions was only possible if the global survey was completed on a domain portfolio slightly more likely to reach English respondents. In sum, comparability on the consumer experience at the time of writing is only methodologically and statistically valid across the English-speaking world given current browser variation and inferential analytics software limitations, and due to the inherent English-language bias of all surveys conducted on the Web.

Further, commonly used software tools based on IP geo-location are limited. There are also limitations on the meta-data inferential analytics presented here, given current software limitations relating to plug-ins.

In conclusion

GRIT CPR brings up some interesting challenges to conventional research wisdom.

First, the migration to mobile devices as a primary means of accessing the Web continues to grow. Surveys not designed for a Web experience will increasingly foster disengagement with participants.

.....
Migration to mobile devices as a primary means of accessing the Web continues to grow

.....
Incentives are not always necessary for participation in research

.....
Use of meta-data for insight generation is a viable and perhaps even vital tool in the research toolkit

Second, incentives are not always necessary for participation in research. If the survey is short and device-agnostic, a significant portion of the audience will participate in research. That does not mean incentives are not necessary in all, or even most, research projects, but it certainly indicates that awards are not the primary driver of engagement.

Finally, the use of meta-data for insight generation is a viable and perhaps even vital tool in the research toolkit. We simply do not have to ask every question; often that data can be captured with a high degree of accuracy via the underlying data that drives all online usage.

We will continue to monitor these trends and report any shifts we see occurring while exploring new ways to conduct this important research.

 **insight
innovation** **exchange** LATAM

Imagine it. See it. **Do it.**

Santiago de Chile
April 8th - 9th 2014

Learn more at www.iiex-latam.org

Toward the Future

The experience of the GRIT CPR study has suggested some interesting innovations to the process going forward. In the spirit of transparency, here are our thoughts on new approaches that could go far to improving both impact and quality of the GRIT CPR study and of other studies using similar approaches.

Meta-data inferential analytics on audience demographics

Imagine – without using any beacons or ‘cookies’ or browsing history data – if the CPR knew more characteristics about the almost 1.6 million people exposed to the CPR questions, but did not need to ask such questions?

One way to do this is to examine the audience demographics of the hundreds of different portfolios of domains, including every domain name that used the RIWI system for the GRIT CPR.

Knowing the characteristics of the people exposed to the question would provide insight to the extent to which the exposed respondents were insufficiently representing certain groups of value to the researcher (such as college-educated respondents). The researcher could also map the percentage of college-age educated respondents who ultimately completed the study to determine if there was concordance between the exposed sub-population of interest and the sub-population ‘completes’ (i.e. those who answered the survey from within this group).

Amalgamating all sites used for this study during the duration of this study, RIWI concluded, using the Alexa audience demographics^{aw} on the sites used to deploy the CPR survey, that the following were characteristics of the nearly 1.6 million exposed respondents. In the following figure, “below” refers to below the average general Internet population; “above” means the opposite. In our study, Alexa audience demographics show male exposures were 14% “below”, but, when looking at the final global totals, men were 65% of the total respondents. This means that a researcher may wish to consider that the already significant over-representation of men in our CPR survey may, in fact, have been higher than the 65-35 split.

Why might this be the case?

It is possible that the nature of the opening engagement question – asking people to self-identify their gender and ages – was significantly less appealing to women than to men ‘stumbling upon’ the CPR question. The same adjustment technique could be done by looking at our age groups vs. the ones in the table below, i.e., comparing the age distribution of exposure below to actual respondents to determine the degree of representativeness the latter reveals, and, possibly, to reweight the data accordingly.

We suggest this as a research mechanism for the future as technology evolves in the ways described in the following pages; the meta-inferential analytics approach we suggest is currently limited by the nature of domain information for reasons explained beneath the Table. We do, however, foresee a day in the near future when these inferential limitations will melt away.

Demographic of Interest	Similarity	Demographic of Interest	Similarity	Demographic of Interest	Similarity
Gender:		Education:		Ethnicity	
Men	14% below	No college	22% below	African	20% below
Women	8% above	Some college	2% below	African American	58% above
Age:		Graduate School	6% below	Asian	16% below
18-24	22% above	College	2% above	Caucasian	36% above
25-34	40% above	Income:		Hispanic	4% above
35-44	12% below	0-\$30K	2% below	Middle Eastern	40% below
45-54	44% below	\$30K-\$60K	22% below	Other	36% below
55-64	58% below	\$60K-\$100K	14% above		
65+	56% below	\$100K+	18% above		
Has Children:		Browsing location:			
Yes	24% below	Home	26% below		
No	22% above	School	58% below		
		Work	36% above		

% above or % below the general internet population
Error is ±2%

Limitations of the meta-inferential analytics approach on exposed respondents

At the time of writing, Alexa only gathers audience demographics on sites based on the response of users with the Alexa toolbar (ATB) plug-in. There is thus a natural bias existing which excludes audience subgroups that do not take part in these rankings. Those users with the ATB may not accurately represent typical Internet browsing behavior.

People who install the ATB provide data about demographics. Alexa would not provide the authors with the percentage of users who have downloaded the ATB. As Google engineer Peter Norvig wrote in a recent blog post, 'Alexa Toolbar and the Problem of Experiment Design,' the data would be good if it truly represented a random sample of internet users, but in fact it only represents those who have installed the Alexa toolbar, and that sample is not random."^{xvii}

In our case the process of amalgamating and averaging the audience demographics on the Websites used for the GRIT CPR, even if each URL has a low ATB plug-in rate, will add to the statistical accuracy of the audience demographics. This lends power to statistical inference, the process of drawing conclusions from data that are subject to random variation, for example, sampling variation. The large sample size also mitigates the challenge here.

Our vision of the future of privacy compliant meta-analytics is definitely possible, assuming approaches similar to what we used for GRIT CPR could be complemented with a system where each domain a greater than 70% consent-based plug-in rate upon which to compile rich audience demographics. Then asking just 3 questions would be the equivalent of asking 20, 50 or 100 or more demographic questions.

.....

Asking just 3 questions would be the equivalent of asking 20, 50 or 100 or more demographic questions

Improving geo-location

To ensure a respondent comes from a particular country of origin, we, like many data collection companies, employ multiple IP-mapping software tools, such as MaxMind.™ Using multiple mapping tools enables data collection platforms to provide strong geo-location, identifying respondents' locations (down to the state/province/region level) by their IP addresses. Other players within this niche industry include geoPlugin, IP2Location, Quova, Digital Envoy, NetGeo, Cyscape, CountryHawk, Digital Element, and IPIntelligence.

It is widely understood in the Internet industry, however, that the more granular the data required (say, within a city) the harder it is to be consistent, and city-based data quality can vary from country to country, especially in developing countries. We are not at a point now where the CPR provides city-based data reliably for large numbers of respondents. However, this could change with the approach we suggest here for future research.

Small sample sizes at the city level are usually problematic in Internet surveys. A country's population is typically divided and 'diluted' across many cities. Many cities are members of a 'greater metropolitan area', which can be aggregated together to improve sample size. The challenge is the lack of a unified system or database that maps city names into metropolitan areas. To research and map cities to metropolitan areas can be manually intensive and very time consuming.

For the purpose of population segmentation by cities, we suggest another approach that can be computationally efficient. Using the Haversine formula (http://en.wikipedia.org/wiki/Haversine_formula), the distance between the target city center and an observation (respondent) can be calculated from their latitudes and longitudes. Latitude and longitude data are readily available from the aforementioned geo-location database providers.

Each observation (respondent) can be classified to a target "city center" by determining if distance to the target is within the approximate radius of that metropolitan area. Repeating the classification for each observation for every target "city center" is not suitable for real-time processing. Instead, these full table scans are better handled by post-processing. Since many statistics software packages have geo-location integration, "city center" classification can also be deferred to the analysis phase.

This classification technique eventually clusters observations into "city centers". For example, a 50 km radius from the city center captures, in our experience, two times the respondent observations in Toronto and 5 times in New York. There are some cities such as Brussels where the radius must be reduced to exclude adjacent city centers like Antwerp.

There are some possible limitations and optimizations to this approach.

1. "City center" classification suffers the same accuracy issues as IP addresses since latitude and longitude data are derived directly from IP addresses using geo-location databases.
2. "City center" classification is better suited for major metropolitan areas where sample sizes accumulate substantially.
3. Prior knowledge of target "city centers," their latitude/longitude and approximate radius are required.
4. The resulting dataset will likely be sparsely populated after classification is complete, especially in countries with high rural populations.
5. If the Haversine formula is computationally expensive, we can substitute the Equirectangular approximation (http://en.wikipedia.org/wiki/Equirectangular_projection), which is sufficient in practice since we only require relative thresholds for "city center" classification.
6. If full table scans are too computationally expensive (especially where many cities in many countries require classification), we can apply a 'decision tree' approach that is partitioned by country and region to reduce the number nodes (possible "city centers") traversed. The 'decision tree' requires more code and prior knowledge for implementation, but can be applied incrementally.

.....
Small sample sizes at the city level are usually problematic in Internet surveys

Complete GRIT CPR data sets for 200+ countries and regions are available at www.GreenBook.org/GRITCPR

Explore, add comments, analyze or benchmark your own data sets. Enjoy!

Appendix

About The RIWI Corporation

RIWI offers a patented, exclusive random domain intercept technology (RDIT™) as a gateway to a non-incented global stream of respondents and data. Peer-reviewed and award-winning (Insight Innovation Competition in NA, 2013), it is unique technology for ongoing data capture. RIWI works with global partners to identify new product opportunities, monitor and improve customer experiences in 190+ countries, identify competitors' strategies, track brand preferences, or monitor product performance.

RIWI methodology

Random domain intercept technology repurposes the URL (direct navigation) bar on any browser on any Web-enabled device to create an entirely new random contact point for a non-incented respondent. As Internet users input addresses into URL^{xviii} bar they can miss-type, miss-guess, or simply have a 'fat fingers' error on a small device^{xix}. It is now becoming best practice, at companies such as Twitter^{xx} and LinkedIn, to ask customers via email to type in their browser to reset to the correct login URL as a safety precaution measure against malware. This GRIT CPR Report proves, using the most current data available, that direct navigation is so frequent as to be effectively universal among Web users. It is almost inevitable that a Web user ends up on an unintended website from time to time. This study is the first to show that the phenomenon is *equally common in every country of the world*, and not just on popular websites we are all familiar with such as Twitter or Tumblr^{xxi}, but also on non-trademarked 'nonsense' domains.

Many mobile users *prefer* direct navigation URL bar search to using the search engines. A survey in 2007 from Opinion Research Corporation found that Internet users prefer direct navigation more than had been previously reported. 64% of users said they bypass search engines and type keywords into their browsers attempting to be directed to the right URL. Other reports have found that around 17% of Web searchers use direct navigation instead of search engines. The survey also revealed that 26% of users say they bypass search engines at least once a day, and 38% said they bypass search engines weekly or monthly. 53% of users surveyed said they would always or sometimes bypass search engines if they could find

what they are searching for by typing the category name into their browser. The reason is this: the URL bar is bigger, and mobile phone usage is rising. Comparing across categories, the results are clear:

	Travel:	Autos:	Restaurants:
Direct navigation:	43%	46%	44%
Search engine:	24%	37%	33%

Study by Marketing Land, 2012

As we report in our CPR findings, more than 1.5 million Web users around the world – from July 1, 2013 to August 18, 2013 – were exposed to the CPR questions through the patented and peer-reviewed RDIT approach.

RIWI questions for the CPR can happen on any Web-enabled device, on a 'page does not exist' error or a domain the user did not intend to reach that otherwise serves up ads. These are not 'pop-ups' or 'pop-underers'. They are real 'run-of-site' URL pages devoid of malware or beacons. These URL pages were aggregated and fed into RIWI's RDIT™ engine where, instead of errors, or ads, the potential respondent was offered up a RIWI Nano-Survey™ (a short survey) on their experience with surveys.

To acquire the sample for these short surveys (i.e. full-page 'run-of-site' surveys that render on any web-enabled device), RIWI has access, through partnerships built since the late 1990s and via proprietary algorithms, to hundreds of thousands of appropriate, non-trademarked URLs from a large number of portfolios of domains, rotating in real-time (via DNS^{xxii} redirects), and through multiple geo-location software algorithms, to ensure, for example, that respondents are not using a proxy server in a location from other than where they appear to respond. Respondents are only able to answer any question from a particular IP address once.

The RIWI process creates a 'statistical router' of highly randomized representative online Web users in every country and in every micro-region of the world. Proprietary code ensures traffic is randomized, 'bot'-free, smart-phone and tablet-optimal, geo-representative, and quality-controlled. It enables 24/7 real-time data capture, simultaneously captured in all geographies, and it leverages neuro-associational snap short surveys for data purity and elimination of coverage bias^{xxiii}.

Footnotes

- i. Operating System, such as Windows (with multiple versions).
- ii. See: <http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2011.pdf>
- iii. See, for example, Beth Synder Bulik, (2011), "Always on Women: A Survey of how Women are Using Technology Today" (Advertising Age). See: <http://gaia.adage.com/images/bin/pdf/1114WP.pdf>
- iv. Toepoel, V., Das, M., Van Soest, A., 2008. Effects of Design in Web Surveys. Comparing Trained and Fresh Respondents. *Public Opinion Quarterly*
- v. Toepoel, V., Das, M., Van Soest, A., 2009. Relating Question Type to Panel Conditioning: Comparing Trained and Fresh Respondents. *Survey Research Methods* 3(2):73-80.
- vi. http://en.wikipedia.org/wiki/List_of_countries_by_English-speaking_population
- vii. http://www.computerworld.com/s/article/9237855/Mobile_s_browser_usage_share_jumps_26_in_three_months
- viii. See: http://en.wikipedia.org/wiki/Usage_share_of_operating_systems#Estimates_for_2011; http://en.wikipedia.org/wiki/Usage_share_of_web_browsers
- ix. If one adds up all of the Windows OS versions cited at this definitive source of running statistics, the sum total is 83% as of July, 2013, the most current external source of statistics at the time of writing. See: http://www.w3schools.com/browsers/browsers_os.asp
- x. Ibid. Our statistics indicate that XP and Vista Windows are statistically overrepresented in CPR global results.
- xi. comScore Reports April 2013 U.S. Smartphone Subscriber Market Share - comScore, Inc. Comscore.com. 2013-06-04. Retrieved 2013-07-20.
- xii. Top Five Smartphone Operating Systems, Shipments, and Market Share, 1Q13 (Units in Millions).
- xiii. See: http://gs.statcounter.com/#mobile_browser-ww-monthly-201207-201307-bar
- xiv. Source: UNECE Statistical Database, compiled from national and international (Eurostat) official sources.
- xv. Definition: Use of Internet includes any kind of use, whether at home, at work or from anywhere else, for private or professional purposes, using a computer or any other means.
- xvi. http://w3.unece.org/pxweb/Dialog/varval.asp?ma=02_GEICT_InternetUse_r&path=../database/STAT/30-GE/09-Science ICT/&lang=1
- xvii. Alexa.com
- xviii. <http://norvig.com/logs-alexa.html>
- xix. The Uniform Resource Locator (URL) was created in 1994 by Sir Tim Berners Lee, and the URI working group of the Internet Engineering Task Force (IETF), the brainchild of collaboration launched at the IETF Living Documents "Birds of a Feather" event in 1992. The format combines the pre-existing system of domain names (created in 1985) with file path syntax, where slashes are used to separate directory and file names. The purpose of this common vernacular (much like telephone numbers), the lingua franca of the Web - the URL - was to easily enable companies and countries to 'talk to one another' using the same Internet syntax. This is why the number of URLs started to flourish in 1993-1995.
- xx. Theoretically, app-based URL linkage may ultimately one day be more popular on tablet for premium high-usage sites like LinkedIn, but users limit the number of apps on tablet for load (and physical space) reasons, and, just like 'favorites' on your browser, early data suggest this will be highly limited and fleeting in take-up.
- xxi. e.g. "Log in directly at Twitter.com if you're unsure: if you're ever uncertain of a website, just type Twitter.com into your browser bar, hit enter, and log in directly from our homepage."
- xxii. Flickr.com used to be an independent site. Then Yahoo, which owns Flickr, the photo-sharing site, took control over it. Fairwinds, which advises companies to protect their brands, calculated that much of Flickr.com's traffic was manual type in. That spelling had 3.6 million unique visitors in one year prior to the Yahoo acquisition of the site. In late June 2010, Facebook recovered many trademarked domains, including Facebok.com, after a lengthy lawsuit. According to Facebook's own SEO team, that Facebok.com domain alone received 250 million unique visitors per year.
- xxiii. DNS is the naming system for computers, services, or any resource connected to the Internet or a private network. It associates various pieces of information with domain names assigned to each of the linked entities.
- xxiv. See article by Spero, Lawrence. ESOMAR Journal, September/October Issue, 2013.



Imagine it. See it.
Do it.

Learn more at www.iiex-eu.org



The best thing to happen to your pocket since the pocket protector.

This app is research whiz certified.

- **3 statistical calculators:** Sample size, confidence interval & margin of error
- **Data on 38 countries:** Basic census data, online penetration rates, socio-economic levels & number of mobile users
- **Calendars:** Conferences & holidays by country



For more about *Research Now* or for a custom quote, contact us at info@researchnow.com.

© 2013 e-Rewards, Inc. All rights reserved.



**DOWNLOAD
FREE
TODAY**