Incentives Research -on-Research

An evaluation of survey data quality under various conditions.

DECEMBER 10, 2015

tap RESEARCH DM2 digital marketing & measurement

Abstract

In late 2015, TapResearch commissioned DM2, a digital marketing & measurement company, to conduct research-on-research to assess if there were any discernable differences in sample characteristics and response quality among sources utilizing different respondent incentives. DM2 utilized a multi-faceted survey with quality scoring algorithms to compare results for the different sample types. Sample was acquired in partnership with Lucid, who selected providers that met the defined characteristics of the study – while keeping sourcing companies blind from DM2 and Tap. Key findings obtained from the study were many, including indication that respondents receiving and preferring cash-based incentives may be more prone to provide lower quality survey data, and that there are some fundamental differences in exposure to and participation in research based on the platform from which respondents engage surveys.

Objectives

The primary objective of the research was to determine which recruitment sources and incentive mechanisms (if any) might have systemic influences on survey data, and under which conditions. Some specific objectives included:

1. Study a combination of sample sources, sample types (panel/river), and incentives
2. Assess ability to recruit and engage respondents using various combinations (where possible)
3. Examine the respondent characteristics of the various combinations
4. Examine survey response quality and survey-taking behaviors under these various combinations
5. Examine attitudes and preferences around survey incentives, and how those might vary by type

Methodology

DM2 deployed a 12-minute survey that explores a variety of subjects, examining many consumer attitudes and behaviors. In the process of profiling respondents on multiple subjects, survey response quality and in-survey behavior was assessed in the background. Sample from multiple providers was acquired to fill-out three types of platforms: opt-in research panels, multi-rewards programs (in which surveys are one way of earning), and real-time recruitment using virtual incentives. In addition, four types of survey incentives were delivered from these platforms: cash, reward points, virtual currencies, and charitable contribution. In order to cross-compare characteristics, a nested Age/Gender quota structure was put in place. In post processing, minimal weights were applied to the quota cells to further refine the comparisons. Statistical analyses were conducted (at a 95% confidence level) on the sample subgroups containing sample sizes of n=400 or greater. There were 1895 total study completes.

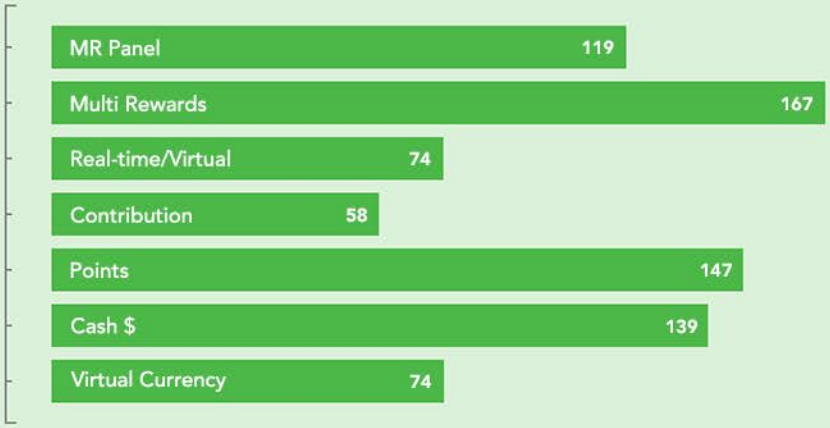
Key Findings

The survey provided a large amount of data which was analyzed in multiple ways. A high priority was to profile and assess differences among the subgroups. The primary differences and key findings are noted below.

The survey had an average interview length of just over 14 minutes, and a median length of 11.5. Past research-on-research by many practitioners (Miller and Menig, 2014, et al) has shown a high degree of correlation between survey speeding and poor response. For this survey, the fastest 10% completed the survey in less than 6.5 minutes – a threshold under which survey engagement might be questionable.

- The proportions of respondents falling into this questionable category varied widely by sample source and type, ranging between 4.7% and 16.4%
- When assessing how sample types performed in comparison to one another, Real-time/Virtual currency and Charitable contribution sources had the fewest questionable respondents – noted below as indices where 100 is the total sample average

Index of Questionable Respondents by Sample Type



FINDING:
% of speeding/questionable respondents varied widely by sample type

Key Findings

A method to better quantify the existence of poor quality responses involves assessing answers within the survey and classifying certain types of response as issues. Tests of response included:

- Consistency of stated age, numeric entry at the front and age range categories in the back
- Consistency on attributes conveying direct opposites
- Open end response quality
- Reasonable statements regarding multiple low incidence items
- Reasonable completion times for various question sets within the survey

In terms of issues identified, Panelists receiving direct cash compensation, either in the form of a check or Paypal, had a statistically higher number of errors/data inconsistencies than those receiving points or virtual currencies.

Response Quality Issues by Incentive Type

FINDING:
Cash had the most issues/lowest quality

Key Findings

Frequent survey participation has been shown to impact survey response, primarily through conditioning of respondents to answer in certain ways. There were multiple differences in survey participation between respondents sourced in real-time (receiving virtual currency incentives) and the more traditional online survey-takers, including:

- Real-time/virtual currency respondents had a directionally lower survey tenure of 2.7 years, versus ~4.5 years for each of the traditional panel and multi-reward program subgroups
- The number of companies for which real-time/virtual respondents complete surveys is lower, averaging 3.0 versus 5.4 for the other subgroups
- Past 30 day survey completion is far lower for the real-time/virtual group at 12.9, versus 26.4 for traditional panelists and 31.2 for multi-rewards respondents.

Response Quality Issues by Incentive Type

Survey Participation	MR Panel	Multi Rewards	Real-Time/Virtual
Years taking online surveys	**4.3**	**4.6**	**2.7**
Number of Companies/Sites registered with for surveys	**5.7**	**5.2**	**3.0**
Self-stated surveys completed Past 30 days	**26.4**	**31.2**	**12.9**

RED denotes statistically lower than the other subgroups @ 95% confidence level

FINDING:
Virtual took the fewest surveys – important to those concerned about "professional" respondents and the implications of conditioning on heavy survey-takers.

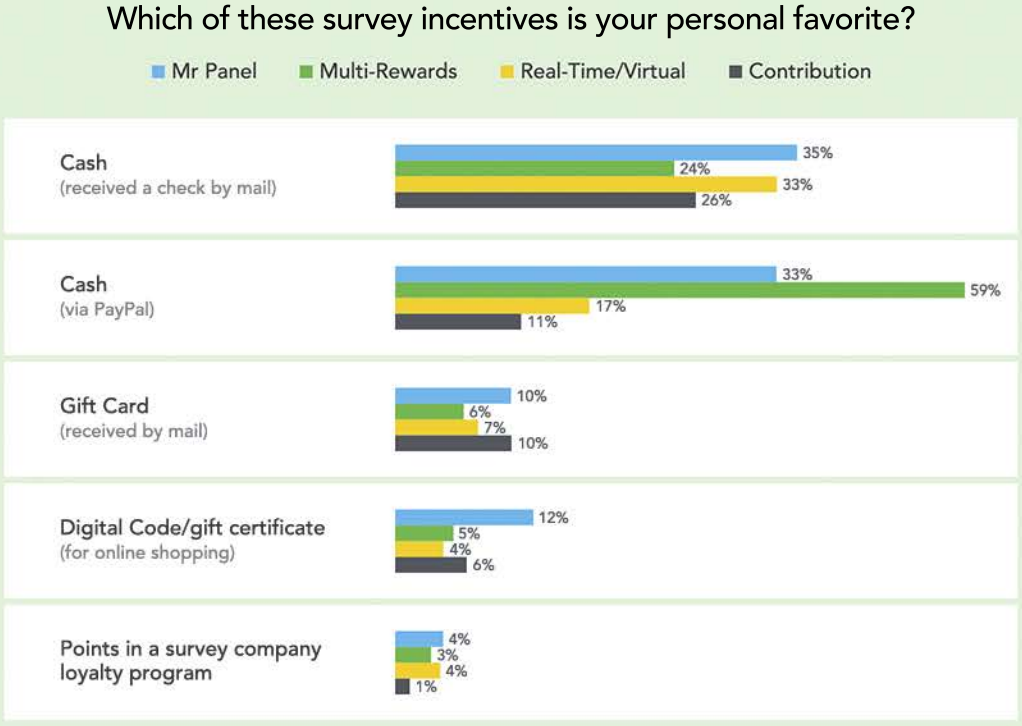
Key Findings

Motivation for taking surveys was similar across all groups, with the exception of the charitable contribution sample

- That sample was half as likely to say "earning rewards" was their primary motivator (25% vs 57%) than the other reward type samples

The preferred incentive was reasonably similar across all groups, with some exceptions:

- 41% of the charitable contribution sample stated that was their preferred incentive, versus ~2% for the other groups
- 28% of those receiving virtual currencies stated that was their preference, versus ~2% for the other groups
- Not surprisingly, cash was highly preferred, the most by the sample that was receiving it (86%), versus cash being the preference for points respondents (64%) and virtual currency respondents (50%).
  - All groups preferred Receiving a Check to Paypal, with the exception of Multi-rewards (who strongly preferred Paypal)
  - Some form of cash was the preference for 37% of the charitable contribution respondents



Key Findings

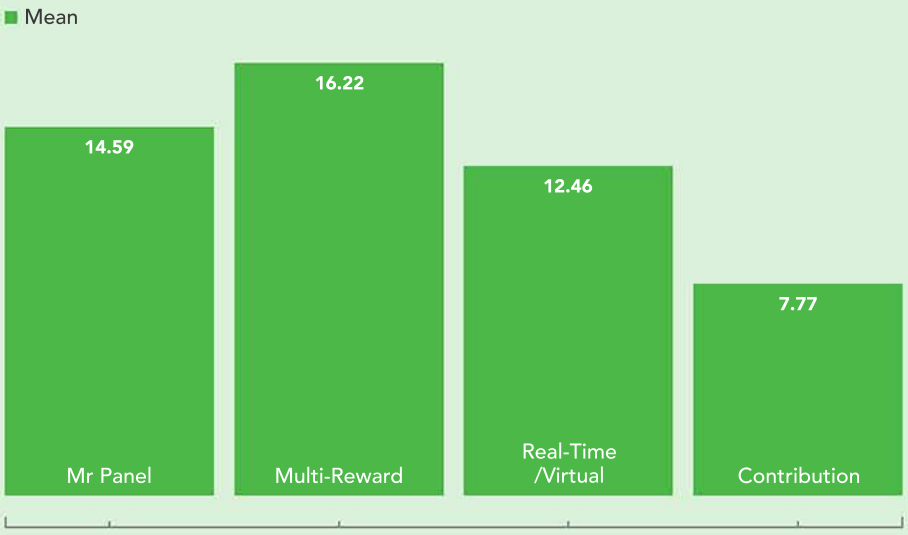
Not surprisingly, respondents most tolerant of longer surveys appear to be the most traditional, from panels and multi-rewards platforms. They stated a "good" survey length where they could maintain attention would be 15.4 minutes on average.

- This aligns with the higher survey participation metrics, indicating those two traditional platforms have the most seasoned survey takers
- Real-time/virtual respondents stated a good length was 12.5 minutes
- Charitable contribution respondents were least tolerant, stating 7.8 minutes was good

FINDING:
Points respondents appear most conditioned to longer surveys, virtual is among less tolerant, and contribution-only sample appears not appropriate for most commercial surveys.

Key Findings

And about how long would be a good length (in minutes) for you where you could maintain your attention all the way through?



In general, the attitudinal and behavioral characteristics of the samples were similar, with the primary exception being the Charitable Contribution sample skewed much more educated and affluent. This could have a significant impact on data outcomes for a variety of research topics (e.g., higher-end products and services) and certain research objectives (e.g., pricing research). Two key differences:

- Bachelor's degree or more (58% vs 30%)
- HH income >$100K (23% vs 9%)

Other areas in which charitable contribution sample skewed statistically different from the others include:

- Higher P30Day alcohol consumption versus the others (62% vs 48%)
- Much lower identification as "current smoker" (11% vs 28%)
- Directionally identifies as more liberal on a political views scale
- More likely to be a registered voter (87% vs 79%)

Key Findings

- Less price conscious when shopping (5pt scale mean 3.02 vs 3.37)
- Less brand loyal when shopping (5pt scale mean 3.20 vs 3.57)
- More likely to own an iPhone (45% vs 29%)
- More likely to own higher end electronics (Tablet, Streaming TV, Car with NAV)
- Higher P12Mo viewing of movie in a theater (68% vs 54%)

FINDING:
Significant attitudinal and behavioral differences with contribution-only sample. Skews could be an issue for many survey topics.

For the most part, those incentivized with cash looked similar to the other sample types. One key distinction of cash respondents was:

- Statistically more likely to Completely Agree that Price is most important consideration while shopping (19% vs 14%)

While most shopping attitudes looked similar across sources, researchers should carefully consider sample sourcing and incentives for pricing studies.

FINDING:
Some concern with engaging cash-only respondents for pricing and branding surveys

- Social media
- Coupons
- Loyalty program use/redemption (which is somewhat counter-intuitive)

Key Findings

Researchers should keep these elements in-mind when commissioning work around these topics.

In general, the vast majority of measures showed similar attitudes and behaviors across sample sources and incentive platforms. This leads us to believe any differences among subgroups are more nuanced and related to items noted above, and not systemic in nature (with the exception of Contribution sample, as noted above). This is a fairly positive finding for purchasers of sample who may have concerns about the presence of latent characteristics in particular subgroups. Researchers should be aware, however, that an individual provider may or may not align with their peers in any particular subgroup.

Items that look generally the same across all sample and incentive types include:

- Outlook, Ailments and Health
- Political views and voter registration
- Channels shopped for a variety of products
- Shopping attitudes (with noted exceptions around importance of Price)
- Tech product ownership
- Media consumption, including movie/music usage and purchase

See graph on page 13

Response Quality Issues by Incentive Type

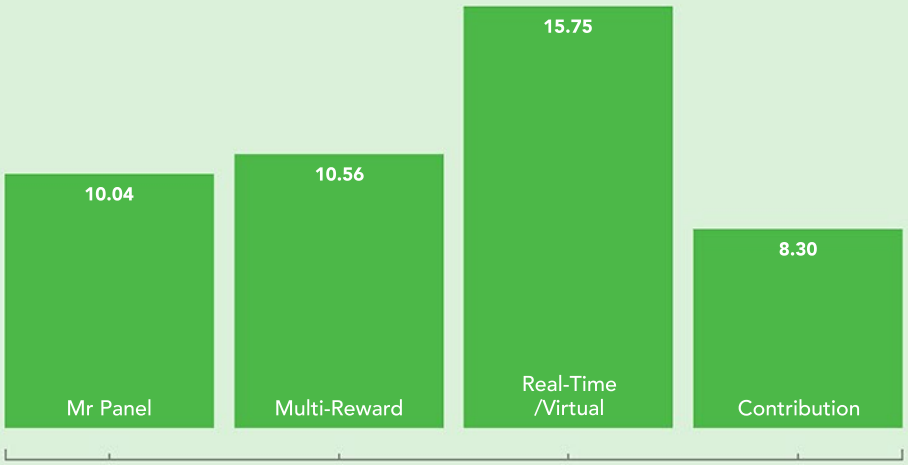
Example Measures Mean scores / %	Points (A)	Cash (B)	Virtual Currency (C)	MR Panel (D)	Multi Rewards (E)	Real-Time /Virtual (F)	Contribution (G)
Job Satsifaction (7=Completely Satisfied)	4.97	5.20	5.10	4.96	5.20	5.10	5.03
Happiness (7=Completely Satisfied)	**5.02**	**5.25 A**	**5.21 A**	5.06	5.21	5.21	5.22
Life Satisfaction (7=Completely Satisfied)	**5.27**	**5.48 A**	**5.48 A**	5.29	5.47	5.48	5.34
Health in General (1=Excellent)	2.19	2.15	2.19	**2.20 G**	2.14	**2.19 G**	**2.04**
Political Leaning (1=Extremely Liberal)	3.90	3.92	4.01	3.95	3.87	4.01	3.71
Concerns about Privacy (5=Extremely Concerned)	3.15	3.12	3.25	3.17	3.09	3.25	3.18
Use generics more than brands (7=Completely Satisfied)	3.18	3.33	3.27	3.24	3.27	3.27	3.16
Price is most important (5=Completely Agree)	3.36	**3.48 C**	**3.27**	**3.42 FG**	**3.43 G**	**3.26 G**	**3.02**
Have Cable or Satellite TV in Household	74%	70%	73%	73%	72%	73%	70%
Own a Tablet in Household	57%	54%	60%	**53%**	**58%**	**60%**	**69% DEF**
Watched a Movie in a Theater P12mos	54%	52%	56%	**50%**	**56%**	**56%**	**68% DEF**
Hours per Week use Internet for Personal Use	24.0	23.3	24.5	22.5	**24.8 G**	**24.5 G**	**19.8**

Cell shading and column designations denote statistically significant differences @ 95%

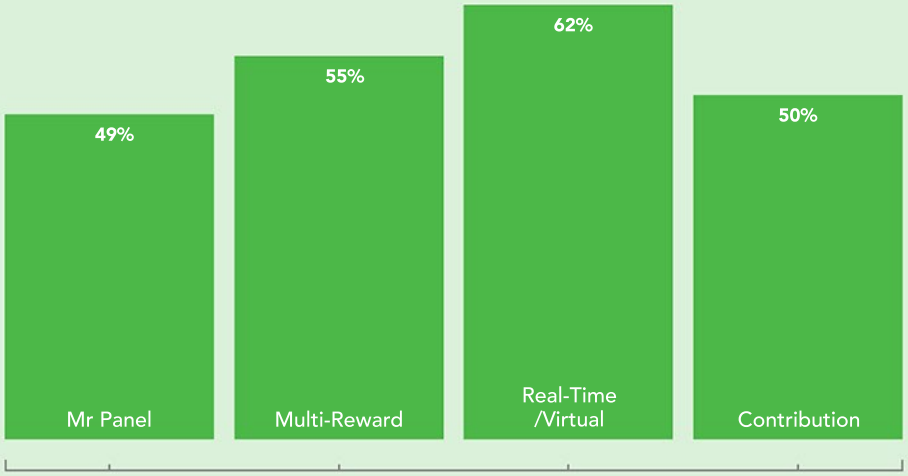
Key Findings

When looking at virtual currency respondents, one key difference related to gaming. Game usage and ownership were both higher among real-time/virtual currency respondents. This may need consideration for certain survey subject matter.

Weekly Hours Games on Android



HH Products Own - Game Console



Summary

In general there were only marginal differences across sample sourcing and incentive types, with the exception of contribution-only sample. The key differences identified above should be noted and considered when selecting sample given the research objectives and subject matter. For example, given its affluence skews, contribution-only sample would not be a good choice for estimating purchase volume of a new high-priced, high tech product.

One of the newest sources of sample for our industry is real-time, virtual currency respondents. For many, little is known about them. When looking at these respondents, a number of characteristics suggest they are highly suitable for many types of research:

- They are among the sources with the fewest speeders and fewest survey issues
- They appear to be the least conditioned respondents, having the lowest survey participation and tenure
- They appear to provide thoughtful, quality response including robust open-ends
- The only notable characteristic on which they differ widely is gaming

While this research uncovered some nuanced differences among the various sample sourcing and incentive options available to researchers, with careful consideration on a study-by-study basis sample buyers can select sample that fulfills all objectives and delivers the best results to clients.