



empanel
ONLINE

**DIGITAL FINGERPRINTING PRODUCT EVALUATION:
COMPARING SURVALIDATE AND RELEVANTID**

INTRODUCTION

Incentives, whether points, gifts or cash, are ubiquitous when it comes to online market research. These incentives range from a few cents to hundreds of dollars. Incentives are the primary tool used to engage online survey respondents. While incentives are effective at motivating respondents to fill out online surveys, there is a dark side. With the potential to earn quick cash, there arises the opportunity for the less honest among us to take advantage of this windfall. The three most common avenues for these thieves are Dupes, Bots and Fraud.

How are we, in the online survey business, to protect ourselves against these virtual bandits? Enter the Digital Fingerprinting tools. These high-tech tools are able to instantly and automatically detect these would-be robbers and block them from polluting your online survey data.

This study compares the two biggest Digital Fingerprinting tools in the market research industry: SurValidate by Bizpinion (a subsidiary of Empanel Online) and RelevantID by Imperium.

WHAT IS DIGITAL FINGERPRINTING?

Most online survey software packages utilize at least some sort of IP check to help prevent ballot-stuffing. Digital Fingerprinting software employs these IP checks as well as many additional variables to create a much more robust system to identify and eliminate these bad respondents. These additional variables may include things such as system settings, browser signatures, language setting, etc.

The two specific Digital Fingerprinting tools evaluated in this study are SurValidate and RelevantID. It was our expressed goal to better understand how effective each tool is in protecting your survey from duplicates and fraud.

More information about each software package can be found online at the following websites:

SurValidate: <http://bizpinion.com/survalidate/>

RelevantID: <http://www.imperium.com/services/relevantid/>

SURVEY SPECS AND OVERVIEW

Sample Size (N) = 250

Length of Interview (LOI): <3 min

Incidence Rate: 100%

Geography: US (*respondent zip code list available in Appendix B)

Target: General Consumers

This test was set up to see how accurately each tool could identify duplicate respondents. To accomplish this, respondents to the study were taken through both SurValidate and RelevantID and then into the online survey. Once in the study, respondents were programmatically taken through the Digital Fingerprinting tools a second time. By doing this each respondent was “forced” into being a duplicate respondent.

How well did each tool identify each of these duplicates?

METHODOLOGY

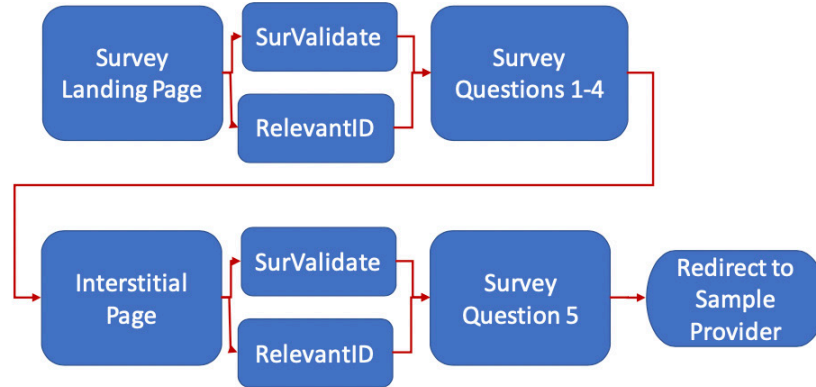
We employed a sample of 250 randomly selected online respondents to take a 2-3 minute survey (*questionnaire available in Appendix A). To eschew any potential bias, internal sample was not used and an independent programmer was used to program and host the survey. The online sample was provided by Branded Research while the online programming and hosting was done by Amaiten.

As customary, respondents received an email invite and a URL to the online survey. A standard, nominal incentive was given to the respondents for their participation.

The default tolerance levels were used for both SurValidate and RelevantID.

As the respondents entered the study, they were passed instantaneously through both SurValidate and RelevantID and then into the survey. The respondent followed the survey through 4 short questions to an interstitial page where they were again sent through SurValidate and RelevantID and then back into the same study at Q5 where the respondent then completed the survey.

DIGITAL FINGERPRINTING PRODUCT EVALUATION: COMPARING SURVALIDATE AND RELEVANTID



This second pass through SurValidate and RelevantID was instantaneous and unknown to the respondent. In doing this each respondent was unwittingly “forced” into being a duplicate respondent. As each of the 250 respondents entered the study twice, the final sample size for the study was 500. The goal here was simple: to understand how well each software tool identified each of these duplicate respondents.

RESULTS

Upon review of the data, it was noted that 12 of the respondents had not fully completed the study. These respondents were eliminated which left 238 unique respondents and 476 completes for the final data set.

PASS 1 – INITIAL RESPONDENT ENTRY INTO THE SURVEY

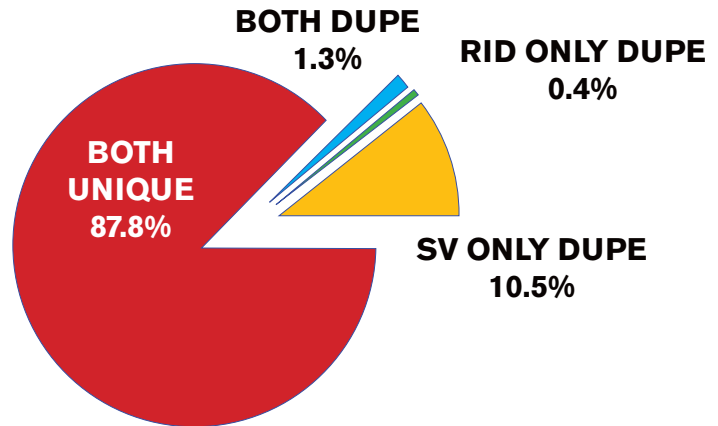
As indicated, digital fingerprinting was employed to validate each respondent as they initially entered the study. Both RelevantID and SurValidate detected potential duplicates in this first pass. This table shows how each respondent was flagged by SurValidate and RelevantID during this first pass.

PRODUCT	# OF UNIQUE UNIQUE RESPONDENTS	# OF DUPE RESPONDENTS	TOTAL
SurValidate	210	28	238
RelevantID	235	3	238

DIGITAL FINGERPRINTING PRODUCT EVALUATION: COMPARING SURVALIDATE AND RELEVANTID

A comparison of the initial pass into the survey reveals that both digital fingerprinting tools agreed 89% of the time with 209 of 238 respondents labeled as unique and 3 respondents labeled as dupes by both SurValidate and RelevantID. There was 1 case where RelevantID identified a duplicate which SurValidate labeled as unique while there were 25 cases where SurValidate identified duplicates while RelevantID labeled these as unique.

PASS 1 FLAG COMPARISON



CATEGORY	COUNTS	%
Both Unique	209	87.8%
Both Dupe	3	1.3%
RID Only Dupe	1	0.4%
SV Only Dupe	25	10.5%
Totals	238	100.0%

PASS 2 – FORCED DUPLICATES

After each respondent finished their fourth question, all 238 respondents were then passed through both digital fingerprinting tools a second time. This table shows how each respondent was flagged.

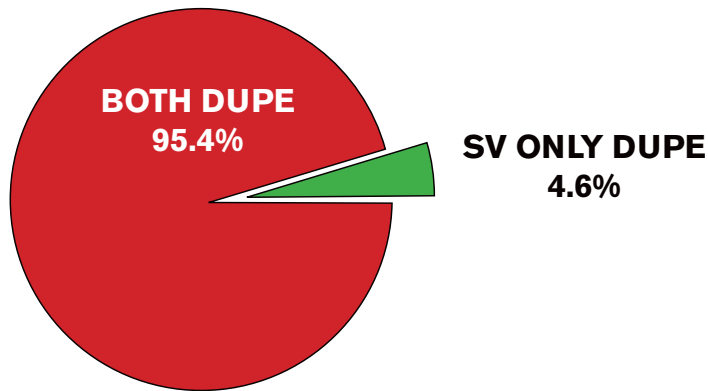
PRODUCT	FLAGGED AS UNIQUE	FLAGGED AS DUPLICATE	TOTAL
SurValidate	0	238	238
RelevantID	11	227	238

DIGITAL FINGERPRINTING PRODUCT EVALUATION: COMPARING SURVALIDATE AND RELEVANTID

After Q4 of the survey, respondents were programmatically (and seamlessly) redirected out of the study, back through both Digital Fingerprinting tools and then back into the survey. (See flow chart above.) This 2nd pass through SurValidate and RelevantID had the effect of artificially creating 238 duplicate respondents. Given this process, both digital fingerprinting tools should have identified 100% (238n) of the respondents as dupes. This was not the case.

Nearly 5% of the time, RelevantID failed to identify the forced duplicate. SurValidate correctly identified all 238 duplicates.

PASS 2 FLAG COMPARISON



CATEGORY	COUNTS	%
Both Unique	0	0.0%
Both Dupe	227	95.4%
RID Only Dupe	0	0.0%
SV Only Dupe	11	4.6%
Totals	238	100.0%



CONCLUSION

Having duplicate respondents within a study increases sample costs, degrades data quality and can lead the users of that data to make incorrect and costly decisions.

Unfortunately, the potential for duplicate respondents is quite high for any online study. Whether that duplicate respondent is due to honest mistakes from things such as cross-panel duplication where a respondent is a legitimate panelist on multiple panels or whether the duplicate respondent is simply a professional survey taker trying to finagle another incentive, we must all be vigilant in keeping these dupes out of our studies.

On the respondents' initial entry into our survey, SurValidate identified 28 respondents (12%) as duplicates while RelevantID found only 4. Admittedly, we cannot be certain how many of these first pass respondents were true duplicates but based on IP addresses, machine information and the rest of the respondents' digital fingerprints, there is an extremely strong chance that the duplicate classification is accurate and not merely a false positive. The wise researcher would eliminate these 28n from their study.

What is certain is that 100% of the respondents were duplicates on their programmatic "forced dupe" second pass into the study. While both SurValidate and RelevantID were proficient at identifying these duplicates on this second pass, this test clearly showed that SurValidate outperformed RelevantID in accurately flagging these duplicate respondents. SurValidate detected 100% of known duplicates while RelevantID detected 95% of those known duplicates.

The cost of duplicate respondents will vary depending on the study. Having dupes in a study about favorite celebrities may have fewer implications than having dupes in a study about a prescription medication; however, a survey worth doing is worth doing properly and without dupes. Digital Fingerprinting tools won't help you write a proper survey, but the right tool will assure you that you have unique respondents. SurValidate has proven itself to be the superior software to eradicate duplicate respondents.

STUDY EXTENSIONS

This initial study was very narrow in scope; we simply looked at a random sample of US consumers and tested to see which software was more proficient at identifying duplicate respondents.

The current study looked at respondents from a single panel. How do these tools compare when duplicates are a result of a panelist being invited from separate panels (i.e. cross-panel duplication)?

It will also be helpful to test these Digital Fingerprinting tools on their ability to defend against Bots and Fraud. Is someone from Russia trying to play havoc with my US political poll?

Additionally, how would the results be affected if International Sample was used instead of US only?

Are these tools as effective if we have duplicate respondents change locations between entering the study?

And finally, any software can label you as a duplicate and kick you out of the study, but how can we be sure these aren't false positives?

So, there are more tests to do on these Digital Fingerprinting tools. Stay tuned!

APPENDIX A: Survey Questions:

Q1: Where are you completing this survey?

- ☐ Home
- ☐ Work
- ☐ Mobile (car, train, plane, etc)
- ☐ Public space (i.e., coffee shop, mall, etc)
- ☐ Other (specify)

Q1A: In what zip code are you currently located?

Enter 5 digit zip code here _____

Q2: How are you currently connected to the internet?

- ☐ Hard-wired (i.e., Cat5/6)
- ☐ Wireless
- ☐ Mobile Data
- ☐ Other (specify)

Q3: What type of device are you using to complete this survey?

- ☐ Computer - desktop
- ☐ Computer - laptop/notebook
- ☐ Tablet - Windows
- ☐ Tablet - IOS/Apple
- ☐ Tablet - Android
- ☐ Tablet -
- ☐ Smartphone - IOS/Apple
- ☐ Smartphone - Android
- ☐ Smartphone - Windows/Other
- ☐ Any other phone
- ☐ Other (specify)

Q3B: What operating system are you using right now?

- ☐ Windows
- ☐ IOS - Phone/Tablet
- ☐ IOS - PC/MAC
- ☐ Chrome OS
- ☐ Other

Q3C: What browser are you using to access this survey?

- ☐ Firefox
- ☐ Chrome
- ☐ Safari
- ☐ Internet Explorer
- ☐ Edge
- ☐ Other

Q4: Are you currently using a VPN or Tor to connect to the internet?

- ☐ Yes
- ☐ No
- ☐ I don't know

CLICK BELOW TO PROCEED

(If Stage=1, assign Stage=2 and (loop to beginning))

Q5: (ask only if Stage=2) About how often do you participate in online surveys?

- ☐ One or more each day
- ☐ Multiple times a week
- ☐ Once a week
- ☐ Several times a month
- ☐ Once a month
- ☐ Less than once a month
- ☐ Don't know

APPENDIX B: RESPONDENT ZIP CODE LIST

00685	11367	17201	28904	33473	42748	57626	67062	80909	92114
01510	11779	18038	29006	33558	44202	59870	67337	84097	92277
01982	11791	19013	29138	33612	44224	60002	70006	85001	92314
02148	11967	19064	29423	33614	44446	60016	70131	85014	92507
02905	12010	19111	29527	33625	44455	60120	70810	85142	92543
04979	12117	19365	29644	33702	44875	60145	72843	85308	92703
05156	12601	19446	29681	33770	45039	60423	73568	85345	93003
06239	13155	19802	30030	33771	45239	60435	73669	85635	93646
07040	13203	20019	30548	33990	46268	60445	75033	85749	94002
07202	13205	20147	30755	34105	46350	60462	75074	87942	94107
07446	13207	21740	30813	34203	47803	60475	75077	89103	94112
07735	14052	22150	31027	34285	48066	60617	75214	89123	94122
08028	14063	22401	31047	34293	48239	61326	75232	89131	94571
08094	14086	23222	32065	34472	51054	61434	76180	89147	95054
08530	14202	24794	32091	34654	51632	61723	77023	89178	95118
10002	14210	24882	32141	34986	52403	62812	77095	89701	95124
10009	14214	25530	32455	35581	53919	63021	77574	90003	95337
10011	14615	27502	32763	36507	54017	63122	78410	90014	96818
10016	14623	27511	32812	37620	54481	63143	78415	90024	97394
10038	14757	27834	32817	37774	54568	64063	78602	90046	97420
11102	15221	28613	33023	37857	55350	64835	79109	90210	98270
11104	16602	28630	33024	37876	55421	65806	80031	90744	98405
11109	16801	28650	33067	38125	55426	66048	80231	91709	98445
11234	17109	28681	33176	38701	56520	66210	80517	91780	99207

APPENDIX C: ADDITIONAL STUDY RESULTS

Note: Data are self-reported

LOCATION WHERE SURVEY WAS TAKEN:		
CATEGORY	COUNTS	%
Both Unique	0	0.0%
Both Dupe	227	95.4%
RID Only Dupe	0	0.0%
SV Only Dupe	11	4.6%
Totals	238	100.0%

TYPE OF INTERNET CONNECTION:		
CATEGORY	COUNTS	%
Hard-Wired	42	27.4%
Mobile	25	9.3%
Wireless	164	60.7%
Other	7	2.6%
Totals	238	100.0%

DEVICE USED FOR SURVEY:		
CATEGORY	COUNTS	%
Laptop	89	37.4%
Desktop	63	26.5%
Android Phone	36	15.1%
Apple iPhone	18	7.6%
Apple Tablet	16	6.7%
All Other	16	6.7%
Totals	238	100.0%

USE OF VPN/TOR BROWSER		
CATEGORY	COUNTS	%
Yes	33	13.9
No	152	63.9
Don't Know	53	22.2
Totals	238	100.0%

OPERATING SYSTEM USED:		
CATEGORY	COUNTS	%
Windows	124	52.1%
Android	41	17.2%
Phone.Tablet IOS	23	9.7%
Chrome OS	21	8.8%
Mac OS	13	5.5%
All Other	16	6.7%
Totals	238	100.0%

TYPE OF BROWSER USED:		
CATEGORY	COUNTS	%
Chrome	129	54.2%
Firefox	30	12.6%
Safari	30	12.6%
Internet Explorer	20	8.4%
Don't Know	9	3.8%
All Other	20	8.4%
Totals	238	100.0%

FREQ OF ONLINE SURVEY PARTICIPATION		
CATEGORY	COUNTS	%
One or more per day	159	66.8%
Multiple times per week	47	19.7%
Once a week	12	5.0%
Several times a month	5	2.1%
Once a month	6	2.5%
Less often	9	3.8%
Totals	238	100.0%



empanel
ONLINE

Empanel Online
6017 Catamaran Court,
Flowery Branch, GA 30542
770.965.1867
sales@empanelonline.com