COMPARISON OF TELEPHONE, MAIL, WEB, AND IVR SURVEYS OF DRUG AND ALCOHOL USE AMONG UNIVERSITY OF GEORGIA STUDENTS

James J. Bason, ,University of Georgia
Survey Research Center, 114 Barrow Hall, Athens, GA 30602-2405

Key Words:

During the Spring 2000 academic semester at the University of Georgia, an experiment involving mode effects between four modes of administration of the Core Institute (Core Institute, Southern Illinois University, Carbondale) survey of drug and alcohol use was conducted by the Survey Research Center among currently enrolled students. The purpose of the study was to examine differences in administering the Core survey by telephone, mail, web-based, and IVR (Interactive Voice Recognition) techniques. The study had three specific aims. First, to examine the degree to which each method produced a sample of students representative of the total population of University of Georgia students, second, to examine relative response rates for each method, and third, to examine substantive differences in frequency and prevalence of alcohol and drug use.

Traditionally, telephone surveys have achieved much higher response rates than mail surveys (Kalton, 1983; Weisberg, Krosnick, and Bowen 1996). However, for specialized populations with near universal internet access, web based or internet surveys have achieved response rates comparable to that of telephone and mail surveys (Schaefer and Dillman 1998), with greatly reduced data collection costs over telephone surveys (Dillman 2000). Although Interactive Voice Recognition (IVR) methods of data collection have been utilized for quite some time for a variety of types of studies, there is a paucity of research that compares this technology to that of telephone, mail, and even web-based surveys. IVR-based survey approaches face obstacles unlike those of web surveys, but clearly the possibility of using IVR in a mixed mode survey strategy is not altogether unreasonable.

A random sample of 3,000 University of Georgia students was drawn by the Registrar's office and randomly partitioned into four groups of 750 students each. Group 1 completed the Core Drug and Alcohol interview by telephone, Group 2 completed the instrument by mail, Group 3 by a web-based survey method, and Group 4 through an Interactive Voice Recognition (IVR) method. For the mail, web-based, and IVR modes of data collection, sampled students received an advance letter inviting them to take part in the study using their respective mode of administration. The advance letter explained the purpose of study,

protections of confidentiality, and instructions for completing the survey. The web-based subjects were provided a unique numeric identification number to allow them to access the web survey on the Survey Research Center web site. IVR subjects also received a unique numeric identification number along with a 1-800 number to call and access the survey on the IVR system. No follow-up was conducted among any of these three groups. Students selected for the telephone mode were called a maximum of two attempts to complete the interview in order to allow a fair comparison against the single wave approaches used for the other three modes.

Table 1 displays response rates for each of the four modes of administration. Surprisingly, the mail mode achieved the highest response rate among the four modes with 27.7 percent of students responding. The response rate for the telephone mode was 23.9%, while the IVR method achieved a 17.4% response rate, and the web-based approach a 15.5% response rate. Response rates are calculated using AAPOR standard definitions. The relative difference in response rates is not large, although the web and IVR methods are lower than the mail and telephone methods. The lower response rate achieved by the web and IVR methods may be related to an increased burden of responding through non-traditional methods. The high response rate achieved by the mail method may be in part due to the fact that students were able to choose when they completed the survey, and generally had a longer period of time to complete the survey and return it. The response rate for the telephone method would certainly have been higher if repeated attempts to reach and interview respondents had been made, as evidenced by the small refusal rate (5.9%), and the high proportion of unresolved callbacks (22.1%) and unknown cases (48.1%). Conversely, the telephone method would have been more costly if repeated attempts to reach students were made due to the costs of labor to conduct telephone interviews. Costs for the other three methods are essentially fixed at the cost of the advance letter and supplies necessary to conduct the mail out.

Table 2 examines the degree to which each method produced a representative sample of students. Each of the methods produced samples that appear to be generally representative of the University of Georgia student population. The telephone method came the closest to matching the true distribution of the student

population by gender, while the mail approach produced a higher proportion of female respondents than did the telephone, web, or IVR methods. Both the IVR and web methods slightly over-represented females. Although the telephone method performed the best in matching true population parameters, a four-way ANOVA test showed no statistically significant differences on the proportion of males and females among the four modes of administration.

No statistically significant differences among groups were found with regard to student classification either. Each of the methods produced high proportions of freshman respondents, but proportions of sophomores, juniors, and seniors were very close to actual population proportions. The telephone method did under-represent graduate and professional students, while the IVR method under-represented sophomores.

No statistically significant differences emerged by mode with regard to ethnicity or age. Each method tended to match proportions of white and African-American students fairly well. White respondents were slightly under-represented in the telephone mode and slightly over-represented in the web mode. African-American students were also under-represented in the web mode.

In terms of age, the 20 year old and under group was slightly over-represented in each mode, but generally each sample did a comparable job in representing population demographics. The examination of demographic representation across mode suggests tentatively that each method will produce a sample generally representative of the total population of students.

Table 3 compares frequency of usage of alcohol and other drugs by each mode of administration. The first item asked students to report the number of times they consumed 5 or more drinks at one sitting during the past two weeks, a measure of binge drinking. The proportions of University of Georgia students reporting at least one such occasion were not statistically significant across mode. University of Georgia students, however, were much less likely than the national average (45.6%) to report incidence of binge drinking. The highest reported proportion of binge drinking was among telephone respondents (38.5%), followed by Web respondents (37.1%), IVR respondents (33.1%), and Mail respondents (31.3%). Across individual categories of use, slight differences among mode are detected.

The average number of reported drinks consumed by students was statistically significant (F = 18.52, p $\leq$ .001) across mode, with Web and IVR respondents much less likely to report having consumed alcohol during the previous two weeks. Nearly 84.0% of IVR respondents and 82.0% of web respondents reported none for the item, while the proportion answering none was 31.2% for telephone respondents and 39.4% for mail respondents. Telephone and mail respondents are also more likely to report a higher number of average drinks at the higher categories than are web and IVR respondents. Compared to the national prevalence rate of 34.9% reporting no drinks in the previous two weeks, data for telephone and mail respondents seems much more reasonable than that for the web and IVR approaches.

Observed differences of the frequency of use of alcohol during the last 30 days is much more even across mode, and matches closely the national figures from 1998. Thirty-one percent of mail respondents, 28.8% of telephone respondents, 28.8% of IVR respondents, and 26.4% of web respondents report having consumed 0 drinks during the last 30 days. Proportions across the other categories are similar with slight exceptions, and again match closely the 1998 national prevalence rate of 26.9%. No statistically significant mode differences were observed for frequency of alcohol use in the last 30 days.

Although frequency of tobacco use by mode is not statistically significant at conventional levels (i.e. 95 percent confidence interval), telephone and IVR respondents reported higher prevalence (37.9% and 30.3%) of tobacco use than did mail and web respondents (26.1% and 27.5%). The national prevalence rate in 1998 was 37.4%, which is closer to the prevalence rate of telephone respondents (37.9%).

Mode differences of frequency of marijuana use, cocaine use, use of hallucinogens, and use of designer drugs were not statistically different across mode. In each case, differences in prevalence and usage by mode and in each category were similar, and compared closely to national rates from 1998.

The final section of items tested asked respondents to report the frequency of five types of behavior resulting from use of alcohol or drugs. Responses to these items are summarized in Table 4. Respondents were first asked to report if they had a hangover as a result of alcohol or drug use, and no statistically significant differences by mode were observed. Web respondents reported the lowest proportion of respondents reporting never (31.4%), followed by mail respondents (45.3%), telephone respondents (46.5%), and IVR respondents (46.7%). Proportions across other categories are close, and the national rate is also comparable at 36.4%.

Significantly higher proportions of mail and web respondents reported having done poorly on a test as a result of alcohol or drug use than did telephone (84.9%) and IVR (81.9%) respondents (F = 3.08, p < .03). There were no statically significant differences among groups on having been in trouble as a result of

1075

alcohol or drug use, and proportions for having never been in trouble ranged from 92.3% to 94.3% across all four groups.

Higher proportions of mail (34.0%) and web (42.3%) respondents reported having missed a class as a result of alcohol or drug use than did telephone (29.6%) and IVR (24.3%) respondents, and the difference among groups was statistically significant (F = 2.91, p ≤ .04). Finally, similar proportions of respondents reported having driven while intoxicated (Telephone = 24.1%, Mail = 28.2%, Web = 21.1%, and IVR = 28.2%) and the rate for University of Georgia students was considerably lower than the 34.4% rate nationally. Differences among groups on this item were not statistically significant.

Comparison of administration of the Core Institute Alcohol and Drug use survey among University of Georgia students showed remarkably few differences across mode. Response rates, demographic representativeness, reported frequency of alcohol and other drug use, and behavior resulting from alcohol and drug use were similar across modes of administration, and only in a few instances were statistically significant differences apparent in the data. These results suggest that using web-based and IVR approaches to data collection may be feasible alternatives to more traditional survey techniques. Several caveats apply, however. In the current study, follow-up was not possible, and as Dillman (2000) notes, "just as multiple contacts are the most important determinant of response in face-to-face, telephone, and regular mail surveys, they are essential for e-mail surveys" (p. 367). Attempts to conduct surveys through web-based or IVR approaches need to utilize vigorous follow-up to be successful.

Secondly, results from the University of Georgia survey reveal that the length of the survey is a limiting factor. The Core Institute survey used in the University of Georgia Drug and Alcohol survey was the short version, but still contained over 100 items and took approximately 25 – 30 minutes to complete. The length of the survey instrument in this case resulted in a loss of approximately 30 respondents for the web version and 40 respondents for the IVR version by the end of the survey. Clearly, the telephone and mail modes were superior in this regard. Finally, data produced in this study showed few statistically significant differences on substantive items in the survey, but additional research is needed to determine if mode differences might exist in different settings or among different populations.

Given the continued decline in response rates in the survey industry, it is clear that new methodological approaches to combat non-response will be needed in the future, and the possibility of using web-based and IVR technologies toward that end seem promising. It is even more likely that these approaches may be used in combination with more traditional approaches in a mixed mode strategy. The results of the current study suggest that such an approach may make sense both in terms of an overall strategy designed to reduce non-response as well as from a cost perspective.

References

Dillman, Don A. 2000. Mail and Internet Surveys: The Tailored Design Method, 2nd ed. New York: John Wiley and Sons.

Kalton, Graham. 1983. *Introduction to Survey Sampling.* Sage University Paper Series on Quantitative Applications in the Social Sciences, No. 07-035. Beverly Hills, CA: Sage.

Schaefer, David R, and Don A. Dillman. 1998. *Development of a Standard E-Mail Methodology.* Public Opinion Quarterly. 62: 378 – 397.

Weisberg, Herbert F., Jon A. Krosnick, and Bruce D. Bowen. 1996. An Introduction to Survey Research, Polling, and Data Analysis. Thousand Oaks, CA: Sage.

Table 1: Response Rate by Mode

Mode:	Telephone (n = 161)		Mail (n = 204)		Web (n = 115)		IVR (n = 128)	
Eligible:	$\underline{n}$	$\underline{\%}$	$\underline{n}$	$\underline{\%}$	$\underline{n}$	$\underline{\%}$	$\underline{n}$	$\underline{\%}$
Complete	161	23.9	204	27.7	115	15.5	128	17.4
Refusal	40	5.9	---	---	---	---	---	---
Callbacks	149	22.1	---	---	---	---	---	---
No Answer	324	48.1	---	---	---	---	---	---
Unknown	---	---	531	72.2	627	84.5	608	82.6
TOTAL	674	100.0	735	99.9	742	100.0	736	100.0
Not Eligible:								
Non-Working/ Wrong Number	76	100.0	---	---	---	---	---	---
Undeliverable	---	---	15	100.0	8	100.0	14	100.0

Table 2: Demographic Characteristics of Respondents

Mode:	Telephone (n = 161)		Mail (n = 204)		Web (n = 115)		IVR (n = 128)		% UGA 1999
	$\underline{n}$	$\underline{\%}$	$\underline{n}$	$\underline{\%}$	$\underline{n}$	$\underline{\%}$	$\underline{n}$	$\underline{\%}$	
Gender:									
Male	67	42.4	59	30.9	43	39.4	51	39.8	44.5
Female	91	57.6	132	69.1	66	60.6	77	60.2	55.5
	$F = 1.81$, Not Significant								
Classification:									
Freshman	42	26.1	50	24.5	28	24.3	40	31.3	19.8
Sophomore	32	19.9	39	19.1	23	20.0	16	12.5	19.0
Junior	31	19.3	32	15.7	18	15.7	20	15.6	18.2
Senior	33	20.5	39	19.1	21	18.3	20	15.6	20.5
Grad./Prof.	23	14.3	44	21.6	25	21.8	32	25.0	22.5
	$F = 0.26$, Not Significant								
Ethnicity:									
White	130	80.7	172	86.0	103	92.7	111	86.7	84.9
African-American	14	8.7	14	7.0	2	1.8	6	4.7	5.7
Asian/Pacific Islander	7	4.3	5	2.5	3	2.7	5	3.9	2.9
Hispanic	3	1.9	2	1.0	1	0.9	0	0.0	1.2
Native American	1	0.6	2	1.0	0	0.0	3	2.3	0.2
Other	6	3.7	5	2.6	2	1.8	3	2.4	5.0
	$F = 0.27$, Not Significant								
Age:									
20 or less	84	52.2	101	49.7	57	50.9	63	49.2	40.4
21 – 24	55	34.2	62	30.5	36	32.1	38	29.7	38.4
25 – 29	11	6.8	20	9.9	9	8.0	14	10.9	10.9
30 and older	11	6.8	20	9.9	10	8.9	13	10.2	10.3
	$F = 1.07$, Not Significant								

Table 1: Response Rate by Mode

Mode:	Telephone (n = 161)		Mail (n = 204)		Web (n = 115)		IVR (n = 128)	
Eligible:	n	%	n	%	n	%	n	%
Complete	161	23.9	204	27.7	115	15.5	128	17.4
Refusal	40	5.9	---	---	---	---	---	---
Callbacks	149	22.1	---	---	---	---	---	---
No Answer	324	48.1	---	---	---	---	---	---
Unknown	---	---	531	72.2	627	84.5	608	82.6
TOTAL	674	100.0	735	99.9	742	100.0	736	100.0
Not Eligible:								
Non-Working/ Wrong Number	76	100.0	---	---	---	---	---	---
Undeliverable	---		15	100.0	8	100.0	14	100.0

Table 2: Demographic Characteristics of Respondents

Mode:	Telephone (n = 161)		Mail (n = 204)		Web (n = 115)		IVR (n = 128)		% UGA 1999
	n	%	n	%	n	%	n	%	
Gender:									
Male	67	42.4	59	30.9	43	39.4	51	39.8	44.5
Female	91	57.6	132	69.1	66	60.6	77	60.2	55.5
	$F = 1.81$, Not Significant								
Classification:									
Freshman	42	26.1	50	24.5	28	24.3	40	31.3	19.8
Sophomore	32	19.9	39	19.1	23	20.0	16	12.5	19.0
Junior	31	19.3	32	15.7	18	15.7	20	15.6	18.2
Senior	33	20.5	39	19.1	21	18.3	20	15.6	20.5
Grad./Prof.	23	14.3	44	21.6	25	21.8	32	25.0	22.5
	$F = 0.26$, Not Significant								
Ethnicity:									
White	130	80.7	172	86.0	103	92.7	111	86.7	84.9
African-American	14	8.7	14	7.0	2	1.8	6	4.7	5.7
Asian/Pacific Islander	7	4.3	5	2.5	3	2.7	5	3.9	2.9
Hispanic	3	1.9	2	1.0	1	0.9	0	0.0	1.2
Native American	1	0.6	2	1.0	0	0.0	3	2.3	0.2
Other	6	3.7	5	2.6	2	1.8	3	2.4	5.0
	$F = 0.27$, Not Significant								
Age:									
20 or less	84	52.2	101	49.7	57	50.9	63	49.2	40.4
21 – 24	55	34.2	62	30.5	36	32.1	38	29.7	38.4
25 – 29	11	6.8	20	9.9	9	8.0	14	10.9	10.9
30 and older	11	6.8	20	9.9	10	8.9	13	10.2	10.3
	$F = 1.07$, Not Significant								

Table 3: Frequency of Usage: Alcohol, Other Substances

Mode:	Telephone (n = 161) n	%	Mail (n = 204) n	%	Web (n = 115) n	%	IVR (n = 128) n	%	Nat'l.
ALCOHOL:									
# Times 5+ Drinks in Previous Week		38.5		31.3		37.1		33.1	45.6
0	99	61.5	138	68.7	61	62.9	83	66.9	55.1
1	26	16.1	20	10.0	13	13.4	14	11.3	14.0
2	12	7.5	16	8.0	9	9.3	10	8.1	10.7
3 – 5	17	10.6	15	7.5	8	8.2	8	6.5	13.8
6 – 9	6	3.7	11	5.5	4	4.1	8	6.5	4.4
10+	1	0.6	1	0.5	2	2.1	1	0.8	2.0
F = 0.21, Not Significant									
Average Drinks Per Week		68.7		60.6		18.1		16.2	65.1
0	50	31.2	80	39.4	77	81.9	103	83.7	34.9
1 – 3	50	31.2	60	29.6	17	18.1	18	14.6	25.4
4 – 9	36	22.5	33	16.2	0	0.0	2	1.6	17.4
10 – 15	15	9.4	15	7.4	0	0.0	0	0.0	12.7
16+	9	5.6	15	7.4	0	0.0	0	0.0	9.6
F = 18.52, p ≤ .0001									
Frequency of Use in Last 30 Days		71.2		68.5		73.6		71.2	73.1
0	46	28.8	64	31.5	24	26.4	34	28.8	27.1
1 – 2	32	20.0	46	22.7	25	27.5	24	20.3	19.8
3 – 5	32	20.0	24	11.8	13	14.3	23	19.5	18.9
6 – 9	22	13.8	34	16.7	9	9.9	16	13.6	16.4
10 – 19	23	14.4	31	15.3	18	19.8	14	11.9	14.2
20 – 29	4	2.5	3	1.5	2	2.2	6	5.1	3.0
30	1	0.6	1	0.5	0	0.0	1	0.8	0.6
F = 0.53, Not Significant									
TOBACCO:									
Frequency of Use in Last 30 Days		37.9		26.1		27.5		30.3	37.4
0	100	62.1	150	73.9	66	72.5	83	69.7	63.1
1 – 2	12	7.5	10	4.9	4	4.4	9	7.6	8.7
3 – 5	8	5.0	7	3.4	4	4.4	2	1.7	3.8
6 – 9	3	1.9	2	1.0	3	3.3	5	4.2	2.8
10 – 19	5	3.1	10	4.9	2	2.2	6	5.0	4.1
20 – 29	7	4.3	12	5.9	3	3.3	4	3.4	5.1
30	26	16.1	12	5.9	9	9.9	10	8.4	12.4
F = 2.28, p ≤ .08									
MARIJUANA:									
Frequency of Use in Last 30 Days		15.8		11.4		15.4		10.8	18.5
0	133	84.2	179	88.6	77	84.6	105	89.0	81.5
1 – 2	13	8.2	11	5.4	7	7.7	5	4.2	6.8
3 – 5	3	1.9	5	2.5	2	2.2	1	0.8	3.2
6 – 9	2	1.3	1	0.5	0	0.0	1	0.8	2.2
10 – 19	3	1.9	4	2.0	3	3.3	3	2.5	2.5
20 – 29	1	0.6	1	0.5	2	2.2	3	2.5	2.2
30	3	1.9	1	0.5	0	0.0	0	0.0	1.6
F = 0.53, Not Significant									

Table 3: Frequency of Usage: Alcohol, Other Substances (cont.)

Mode:	Telephone (n = 161)		Mail (n = 204)		Web (n = 115)		IVR (n = 128)		% Nat'l.
COCAINE:	n	%	n	%	n	%	n	%	
Frequency of Use in Last 30 Days		0.6		1.0		1.1		0.8	1.8
0	156	99.4	201	99.0	90	98.9	117	99.2	98.2
1 – 2	0	0.0	1	0.5	1	1.1	1	0.8	1.0
3 – 5	0	0.0	1	0.5	0	0.0	0	0.0	0.3
6 – 9	0	0.0	0	0.0	0	0.0	0	0.0	0.2
10 – 19	0	0.0	0	0.0	0	0.0	0	0.0	0.1
20 – 29	0	0.0	0	0.0	0	0.0	0	0.0	0.1
30	1	0.6	0	0.0	0	0.0	0	0.0	0.1
			F = 0.35, Not Significant						
HALLUCINOGENS:		1.9		0.5		1.1		1.9	2.4
0	153	98.1	202	99.5	89	98.9	117	99.2	97.7
1 – 2	2	1.3	1	0.5	1	1.1	1	0.8	1.6
3 – 5	1	0.6	0	0.0	0	0.0	0	0.0	0.4
6 – 9	0	0.0	0	0.0	0	0.0	4	1.1	0.1
10 – 19	0	0.0	0	0.0	0	0.0	0	0.0	0.1
20 – 29	0	0.0	0	0.0	0	0.0	0	0.0	0.0
30	0	0.0	0	0.0	0	0.0	0	0.0	0.1
			F = 0.82, Not Significant						
DESIGNER DRUGS:		2.6		1.5		1.1		2.5	1.4
0	152	97.4	200	98.5	86	98.5	115	97.5	98.6
1 – 2	4	2.6	2	1.0	1	1.1	3	2.5	0.9
3 – 5	0	0.0	1	0.5	0	0.0	0	0.0	0.2
6 – 9	0	0.0	0	0.0	0	0.0	0	0.0	0.1
10 – 1 9	0	0.0	0	0.0	0	0.0	0	0.0	0.1
20 – 29	0	0.0	0	0.0	0	0.0	0	0.0	0.0
30	0	0.0	0	0.0	0	0.0	0	0.0	0.1
			F = 0.19, Not Significant						