

# WEB-BASED DATA COLLECTION: AN EFFECTIVE STRATEGY FOR INCREASING AFRICAN AMERICANS' PARTICIPATION IN HEALTH-RELATED RESEARCH

Pamela E. Scott-Johnson, PhD; Susan M. Gross, PhD, MPH, RD;  
Dorothy C. Browne, MSW, MPH, DrPH

The study sought to explore response modes (via web-based vs paper surveys) and rates to a follow-up health questionnaire and to examine respondent characteristics by response modes. 192 young adult African Americans responded online or by paper. We found observable differences in follow-up responses, with more participants completing the online version first. No statistical differences were revealed in response modes based on academic discipline, sex, income or health status. The 60% follow-up response rate supports web-based data collection as a viable means of assessing health information from African Americans. This research provides evidence of the Internet as a viable alternative for increasing participation of young African American adults, a relatively understudied group, to obtain data on health status and behaviors, over time. (*Ethn Dis.* 2010;20[Suppl 1]:S1-201-S1-206)

**Key Words:** Web-based Survey, African American Research Participation, College Health and Wellness

## INTRODUCTION

The inclusion and participation of ethnic minorities and women in health-related research is imperative to eliminating disparities in health. Wendler, Kington, Madans<sup>1</sup> demonstrated that minorities were as willing to participate in studies as non-minorities, contradicting other research findings that demonstrated low willingness and participation.<sup>2-4</sup> Barriers, such as mistrust in researchers conducting studies on minority populations, access to study facilities, income and job status which decrease opportunities for participation in research studies, persist and must be addressed.<sup>5-7</sup> Corbie-Smith<sup>4</sup> reported that the socioeconomic status of minorities is a major impediment to participation. Low-income minority participants lack readily available access to transportation to get to research sites and may lose wages as a result of taking time from work to participate in studies. Mistrust and apprehension toward medical personnel and researchers, as well as suspicion of mistreatment and alternative motives, are acknowledged impediments that discourage minority participation.<sup>2</sup> Adams-Campbell and her colleagues<sup>3</sup> indicated that one major barrier to participation was the criteria used to determine participant eligibility for clinical trials. Thus, this aspect of the research design inhibited participation of African Americans and other ethnic minorities.

The Internet has been shown to be a powerful data collection tool.<sup>8,9</sup> Using the Internet as a tool for recruiting and assessing the health status of minorities offers new opportunities for behavioral researchers for both cross-sectional and longitudinal investigation.<sup>10</sup> When compared to other research response modes (eg, postal mail services, fax, or e-mail),

the Internet produced an increased response rate,<sup>9</sup> reduced costs in participant recruitment and retention,<sup>11,12</sup> and the ability to maintain higher data quality and analysis, since the data collection is automated and accurate.<sup>9,13,14</sup>

Of particular interest are the studies that compare response modes in assessing health and wellness of college students or young adults. For example, Pealer and her colleagues<sup>15</sup> examined the feasibility of a web-based method when collecting high-risk health-related information from college students. After comparing web-based and mailed responses, these researchers reported that participants in the web-based response mode answered more sensitive items than did their mail-based counterparts. Moore and her colleagues<sup>16</sup> compared web-based and postal mail methods in assessing the effectiveness of a college-based binge drinking intervention. The authors reported more web-based responses (83%) than print-based responses (75.4%) to the process evaluation. Morrell and her colleagues<sup>17</sup> used a web-based methodology to examine smoking behaviors among college students. This methodology allowed them to examine tobacco use among 21,000 university students. Sixty-eight percent of the males reported smoking at least once; 21% reported having used smokeless tobacco compared to 9% of females.

The purpose of this article is to explore response mode (via web-based vs paper survey) to the follow-up health survey, as well as examine respondent health indicators and related behaviors.

## METHODS

The College Health and Wellness Study was an investigation of health-

---

From Morgan State University, Department of Psychology, Baltimore, MD (PSJ); Johns Hopkins Bloomberg School of Public Health, Department of Population, Family and Reproductive Health, Baltimore, MD (SMG); and Norfolk State University, The Ethelyn R. Strong School of Social Work, Norfolk, VA (DCB).

Address correspondence and reprint requests to Pamela E. Scott-Johnson, PhD; Morgan State University; Department of Psychology; 1700 E. Cold Spring Lane; Baltimore, MD 21251; 443-885-3508; 443-885-8239 (fax); pamela.scottjohnson@morgan.edu

promoting practices and psychological determinants of health in a graduating class of seniors in spring 2003 from a historically Black university located in the mid-Atlantic region.<sup>18</sup> For the one-year follow-up study, only baseline participants who had graduated in 2003 and provided accurate contact information ( $N=357$ ) were recruited. 192 of those who participated in the 2003 baseline survey responded to a one-year follow-up.

### Recruitment and Data Collection

In 2003, participants completed a paper and pencil survey measuring their health status. In 2004, one year later, these same individuals were asked to complete a similar survey in order to determine potential changes in health status. However, for the follow-up, participants were given the choice between responding online or by paper. In March 2004, all participants from the baseline survey were sent a letter inviting them to participate in the follow-up survey. Each participant was informed of the survey's purpose and was offered a \$25 gift card to a national chain retail store or video rental store as an incentive for completing the survey. The first 100 respondents were also entered in a raffle for a concert and dinner worth \$500.

Respondents were given the choice of completing the survey via Internet or by mailed paper survey (participants were instructed to contact the study coordinator by telephone in order to receive a paper survey with a postage paid return-addressed envelope included). The contact letter provided the website for the follow-up survey and the respondent's unique identification code, which would serve as their password to access the survey.

Participants were given three months to respond using the online methods. Those who did not complete the online survey within one month were sent a reminder postcard, and three

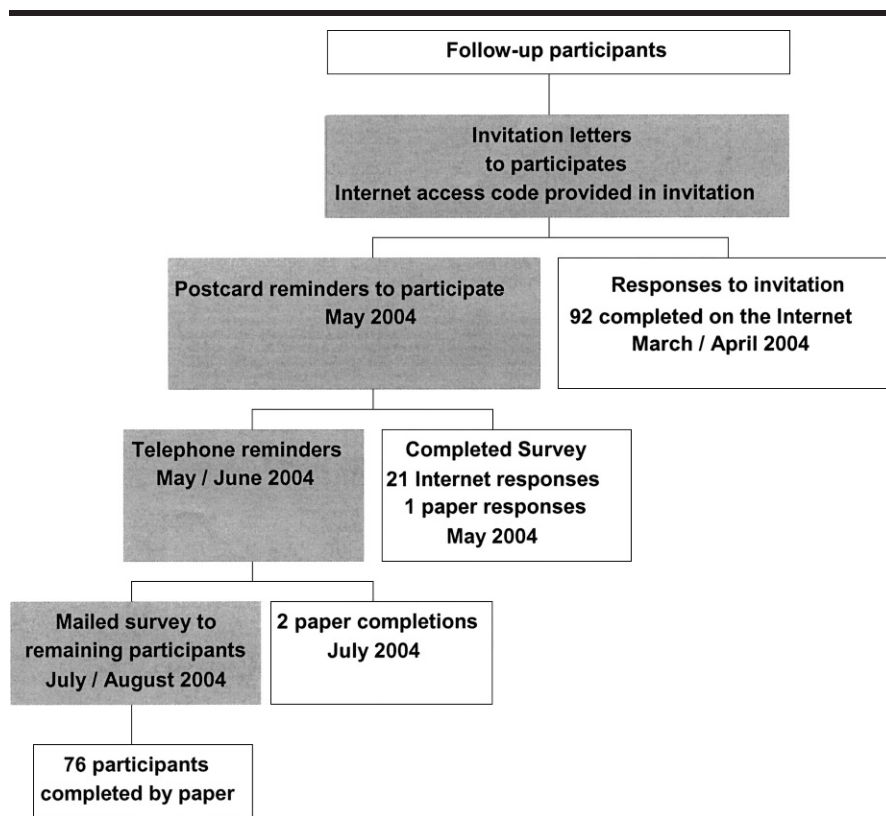


Fig 1. Flow diagram of the data collection illustrating process and the responses received at each time point

additional attempts were made to contact them by telephone (one attempt before noon, one attempt after 5:00 pm., and one attempt on a weekend day). Finally, after three months, all baseline participants who did not complete the survey were sent a paper survey with a self-addressed stamped envelope. Figure 1 presents the flow diagram of the data collection process and the responses received at each time point. All of the Internet responses ( $n=113$ ) were received during the first three months of data collection and the majority of the paper survey responses ( $n=79$ ) were received during the last three months of data collection. All data collected by paper questionnaires were entered into the website by study staff. A 10% random sample of paper questionnaires was used to check for accuracy of data entry.

### Measures

Responses from the follow-up surveys were received and categorized by response modes as either web-based survey or paper survey. Baseline data collection acquired personal demographic information, such as age, academic major study area, race and US citizenship. The follow-up data collection process included items on an individual's marital status, number of children, income and employment status.

In order to determine the general health status of participants, each respondent was asked to rate his/her general health as excellent, very good, good, fair, or poor in both baseline and follow-up surveys.

Body mass index (BMI) was calculated in  $\text{kg/m}^2$  using self-reported height and weight. Participants were

categorized according to NIH guidelines: underweight  $\leq 18$ , acceptable 19–24, overweight 25–29, obese 30–34, or extreme obesity (obesity II)  $\geq 35$ –39 at both baseline and follow-up surveys.

Respondents were asked about the frequency, duration, and intensity of leisure-time physical activities during the preceding 30 days. Level of physical activity was assessed using a continuous score expressed as metabolic equivalent (MET)-minutes per week. Activity level was assigned using the International Physical Activity Questionnaire guidelines for scoring: inactive is 0 to 599 MET-minutes per week, minimally active is 600 to 2999 MET-minutes per week, and Health Enhancing Physical Activities (HEPA) were activities of at least 3000 MET-minutes per week. Respondents' answers were dichotomized into "inactive" or "any activity" at both baseline and follow-up measures.

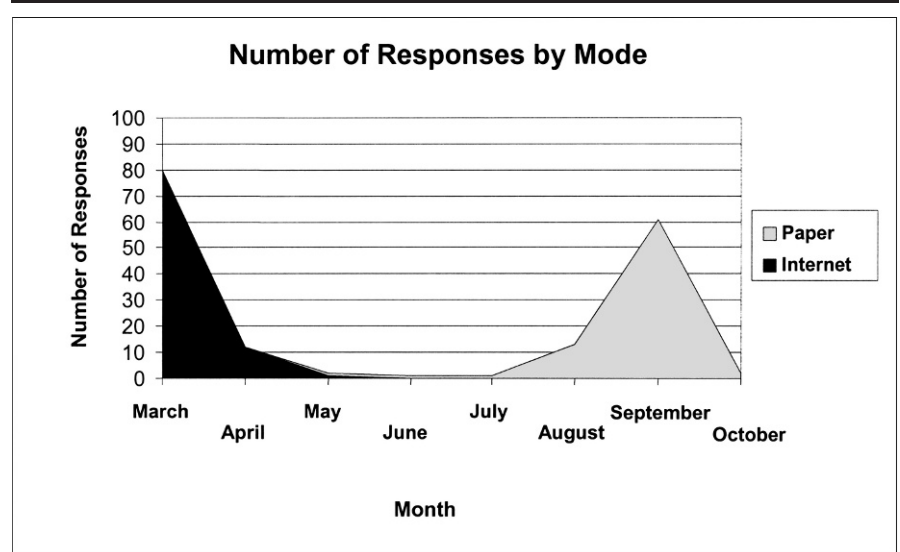
Dietary intake was assessed using the Eating Behaviors Patterns Questionnaire,<sup>19</sup> which is a 51-item scale that was designed to be a behavioral assessment of eating that is predictive of fat intake in African American women.

Symptoms of depression at baseline and one-year follow up were assessed using the Revised Center for Epidemiologic Studies Depression Scale (CESDR scale<sup>20</sup>). The CESDR includes 35 items that measure mood, somatic complaints, interactions with others and motor functions.

At baseline, respondents were asked to report on their exposure (age of opportunity to use) and experience (age of first use and use within past 30 days) with illicit drugs and alcohol. At follow-up respondents were also asked to indicate the appropriate exposure category for cigarette smoking and hard liquor use since graduation.

### Statistical Analysis

Response mode, sociodemographic, health status, weight-related, depression and substance use variables were sum-



**Fig 2.** Participants chose to respond by Internet more immediately and in greater numbers than those by paper. The Internet responses ( $n=113$ ) were received during the first three months of data collection and the majority of the paper survey responses ( $n=79$ ) were received during the last three months of data collection

marized using means and frequencies. Chi-square tests and  $t$  tests were performed to assess the association between response mode and sociodemographic, health status, weight-related behaviors and attitudes, symptoms of depression and substance use variables. A comparison of health status and related behaviors by response modes from those who completed the follow-up survey was conducted. We also compared baseline responses and follow-up responses according to the mode in which participants completed the follow-up. Data were analyzed using SPSS (Version 9.0, 1999 SPSS Inc., Chicago, IL) with statistical significance at the  $P < .05$ .

## RESULTS

Three hundred fifty-seven participants from the original study were invited to take part in the follow-up study. Of the 192 follow-up participants, and regardless of the response mode, survey participants included more females than males (52.3% vs. 35.0%,  $\chi^2=10.34$ ,  $df=1$ ). Most re-

spondents chose the web-based format (59%) and the rest completed the paper form. At baseline, the mean age was  $23.7 \pm 5.4$  years, 70% reported an annual income of less than \$20,000, and 91.5% reported their health status as good, very good or excellent. Approximately, 36% were liberal arts majors, 27% were business and management majors, 16% were computer science majors, 12% were education majors and 9% were engineering majors. At follow-up the majority of respondents (83%) reported being employed full-time or part-time.

More participants ( $n=113$ ) responded during the first three months using an Internet response mode than did those who responded using paper. The majority of the paper survey responses ( $n=79$ ) were received during the last three months of data collection in response to reminders either by phone or postcard (Figure 2).

There were more females who responded by paper mode (78.5%) than females who responded by Internet mode (66.4%;  $\chi^2=3.58$ ,  $df=1$ ). However, sex differences in response mode

**Table 1. Participant characteristics by response mode**

Characteristic	Internet % (n=113)	Paper % (n=79)	$\chi^2$	df	P value
<b>Age</b>	23.6 ± 4.2	23.6 ± 6.3	$F_{(1,190)} = 0.001$	1,190	.98
<b>Sex</b>					
• Male	33.6	21.5	3.58	1	.06
• Female	66.4	78.5			
<b>Major</b>					
• College of liberal arts	34.5	35.4	0.98	4	.91
• School of business and management	24.8	30.4			
• School of computer, mathematical and natural science	15.9	13.9			
• School of education and urban studies	14.2	11.4			
• School of engineering	10.6	8.9			
<b>Employment</b>					
• Other	15.5	17.9	0.42	2	.81
• Part-time	17.2	12.8			
• Full-time	67.2	69.2			
<b>Income</b>					
• <\$20,000 per year	25.2	29.5	1.32	2	.52
• \$20,000 to \$40,000 per year	56.1	48.7			
• >\$40,000 per year	18.7	21.8			

No statistically significant differences by age, sex, academic major, employment and income.

were not significant. In terms of academic disciplines, the highest percentage of responses in the study came from individuals majoring in the liberal arts. There were no statistically significant differences by academic discipline and response mode ( $\chi^2=.98$ ,  $df=4$ ). Most of the Internet participants were employed full-time (67%) and earned between \$20,000 and \$40,000 per year. No difference was found by income and response mode. (Table 1)

In addition to comparing demographic information by response mode, the study sought to compare the health status and related health behaviors from graduation (baseline) to one-year follow-up (Table 2). Weight status was the only health indicator that did not significantly differ from baseline; there were no differences in health status or related behaviors by response mode. From baseline to follow-up, we observed some trends in health status and related health behaviors: increased reporting of very good health status (39.5% vs 46.8%), higher BMI (25.7

± 6.3 kg/m<sup>2</sup> vs 26.2 ± 7.0 kg/m<sup>2</sup>), less physical activity (33.7% vs 21.0% with any activity) and more symptoms of depression (15.3 ± 10.5 vs 18.0 ± 10.5). The symptoms of depression were not significant. Dietary intake and cigarette smoking remained relatively stable from baseline to follow-up. These trends in health status and related health behaviors did not vary by mode of response at baseline or at follow-up. At baseline, 42% of follow-up Internet respondents and 33.3% of follow-up paper respondents reported having *very good* health; and at follow-up, 49.5% Internet respondents and 42.5% of paper respondents reported having *very good* health.

## DISCUSSION

The purpose of the study was to explore differences in response mode for those completing a one-year follow-up survey. Additionally, the study examined and compared the baseline and

follow-up health indicators of a cohort of college graduates by response mode at follow-up. Based on the literature,<sup>16,21,22</sup> we anticipated that differences in response modes by sex, income and employment status would be statistically significant. The data do not show these differences. These findings may suggest that given college students' widespread use of the Internet, web surveys are the method of choice.<sup>23</sup> Carini et al<sup>13</sup> report that college students have higher online usage for communication compared to the general population. Our research findings indicate that while there were observable differences, there were no statistically significant differences by response mode in any of the variables that were examined. Overall, the 60% follow-up response rate and the lack of differences in response mode support the Internet as a viable means of assessing health information in an African American population.

In this study, a web-based response mode was not a barrier to participation by African Americans and may have facilitated responding. This contradicts the idea that minorities are less likely to use the Internet or to respond to online health surveys.<sup>8,13,21</sup> Thus, the assertion reported earlier in this paper that African Americans do not respond favorably to web surveys is not applicable. Furthermore, the broader use and comfort with the use of web-based surveys with this population implies that Internet utilization for data collection on college campuses is beneficial. Moreover, most HBCUs have access to this technology and can utilize this method of data collection.

## Implications

As our research indicates, acceptance and comfort with the Internet, at least within this population, transcends the demographic and socioeconomic barriers associated with the mode of completion. Health educators and practitioners who are committed to increasing

**Table 2. Baseline responses and follow-up responses**

	Baseline %	Follow-up %
<b>Health status</b>		
• Excellent	15.8	10.5
• Very good	39.5	46.8
• Good	35.8	32.6
• Fair/poor	9.0	10.0
	$\chi^2_{(12)}=125.6, P<.01, df=3$	
<b>Weight status</b>		
• Underweight	4.7	6.4
• Optimal weight	48.3	47.7
• Overweight	27.9	32.6
• Obese	19.2	13.4
BMI mean (sd)	25.7 (6.3)	26.2 (7.0)
	$t_{(171)}=-1.91, P<.06, df=3$	
<b>Physical activity (baseline)</b>		
• Inactive	66.3	79.0
• Any activity	33.7	21.0
Average total MET minutes per week, Mean (sd)	4138.0 (12907)	1901.5 (6693)
	$t_{(177)}=-1.99, P<.05$	
<b>Dietary intake</b>		
• Low fat Diet	48.1	49.2
<b>Cigarette smoking (ever used)</b>		
• Ever used/since graduation	21.8	17.6
• No, never	78.2	82.4
<b>Hard liquor (ever used)</b>		
• Ever used/since graduation	36.0	56.5
• No, never	64.0	43.5
	$\chi^2_{(1)}=4.42, P<.05, df=1$	
<b>Depression (follow-up)</b>		
• Any symptoms	78.5	66.3
• No symptoms	21.5	33.7
CESD Score mean (sd)	15.3 (10.5)	18.0(10.5)
	$t_{(177)}=-2.48, P<.05$	

The chi-square and F-test relate to the test done comparing follow-up response categories (paper vs Internet) for each variable first at baseline and second at follow-up. All baseline data were collected using paper questionnaires; they do not test whether scores are different from baseline to follow-up.

the participation of ethnic minorities, and African Americans in particular, from diverse academic, income, and employment backgrounds should consider using an online means of data collection, which is viable, cost-effective, convenient, accessible, and acceptable.

**ACKNOWLEDGMENTS**

We would like to thank the following for their assistance: Dr. Natasha Williams, postdoctoral fellow in the joint Morgan-Hopkins Center for Health Disparities; Corey Bacchus, graduate research assistant

in the Morgan State University, School of Public Health and Policy; and Lynnett Gray, undergraduate research student in the Morgan State University, Department of Psychology.

This research was funded and supported from awards to Morgan State University by the NIH National Center on Minority Health and Health Disparities 1P60MD000214-01, 5P60MD00217-02, U24DA12390-04.

**REFERENCES**

1. Wendler D, Kington R, Madans J, et al. Are racial and ethnic minorities less willing to participate in health research? *PLoS Med.* 2006;3(2):e19.

2. Corbie-Smith G, Thomas SB, Williams MV, Moody-Ayers S. Attitudes and beliefs of African Americans toward participation in medical research. *J Gen Intern Med.* 1999; 14:537-546.

3. Adams-Campbell LL, Ahaghotu C, Gaskins M, et al. Enrollment of African Americans onto clinical treatment trials: study design barriers. *Am J Clin Oncol.* 2004;22(4): 730-734.

4. Corbie-Smith G. Minority recruitment and participation in health research. *N C Med J.* 2004;65(6):385-387.

5. Olin JT, Dagerman KS, Fox LS, Bowers B, Schneider LS. Increasing ethnic minority participation in Alzheimer disease research. *Alzheimer Dis Assoc Disord.* 2002;16(Suppl 2):S82-S85.

6. Miskimen T, Marin H, Escobar J. Psychopharmacological research ethics: Special issues affecting US ethnic minorities. *Psychopharmacology (Berl).* 2003;171(1):98-104.

7. Noah DA. The participation of underrepresented minorities in clinical research. *Am J Law Med.* 2003;29(2-3):221-245.

8. Fricker RD, Schonlau M. Advantages and disadvantages of Internet research surveys: Evidence from the literature. *Field Methods.* 2002;14(4):347-367.

9. Rhodes SD, Bowie DA, Hergenrather KC. Collecting behavioural data using the World Wide Web: Considerations for researchers. *J Epidemiol Community Health.* 2003;57(1): 68-73.

10. Pittenger DJ. Internet research: An opportunity to revisit classic ethical problems in behavioral research. *Ethics Behav.* 2003;13(1): 45-60.

11. Raziano DB, Jayadevappa R, Valenzula D, Weiner M, Lavizzo-Mourey R. E-mail versus conventional postal mail survey of geriatric chiefs. *The Gerontologist.* 2001;41:799-804.

12. Cobanoglu C, Warde B, Moreo P. A comparison of mail, fax, and web-based survey methods. *Int J Market Res.* 2001;43(4): 441-451.

13. Carini RM, Hayek JC, Kuh GD, Kennedy JM, Ouimet JA. College student responses to web and paper surveys: Does mode matter? *Res High Educ.* 2003;44(1):1-19.

14. Eysenbach G, Wyatt J. Using the Internet for surveys and health research. *J Med Internet Res.* 2002;4(2):E13.

15. Pealer LN, Weiler RM, Pigg RM, Miller D, Dorman SM. The feasibility of a web-based surveillance system to collect health risk behavior data from college students. *Health Educ Behav.* 2001;28(5):547-559.

16. Moore MJ, Soderquist J, Werch C. Feasibility and efficacy of a binge drinking prevention intervention for college students via internet

## WEB-BASED DATA COLLECTION - Scott-Johnson et al

- versus postal mail. *J Am Coll Health*. 2005;54(1):38-44.
17. Morrell HE, Cohen LM, Bacchi D, West J. Predictors of smoking and smokeless tobacco use in college students: A preliminary study using web-based survey methodology. *J Am Coll Health*. 2005;54(2):108-115.
  18. Gary TL, Gross SM, Browne DC, LaVeist TA. The college health and wellness study: Baseline correlates of overweight among African Americans. *J Urban Health*. 2006;83(2):253-65.
  19. Schlundt DG, Hargreaves MK, Buchowski MS. The eating behavior patterns questionnaire predicts dietary fat intake in African American women. *J Am Diet Assoc*. 2003;103(3):338-345.
  20. Eaton WW, Muntaner C, Smith C. *Revision of the Center for Epidemiologic Studies Depression (CES-D) Scale* Johns Hopkins University. Prevention Center: Baltimore, MD, 1998.
  21. McCabe SE, Boyd CJ, Couper MP, Crawford S, D'Arcy H. Mode effects for collecting alcohol and other drug use data: Web and U.S. mail. *J Stud Alcohol*. 2002;63(6):755-761.
  22. Hanauer D, Dibble E, Fortin J, Col N. Internet use among community college students: Implications in designing healthcare interventions. *J Am Coll Health*. 2004;52(5):197-202.
  23. McCabe SE. Comparison of web and mail surveys in collecting illicit drug use data: A randomized experiment. *J Drug Educ*. 2004;34(1):61-72.