Computers in Human Behavior 58 (2016) 354–360

Contents lists available at ScienceDirect

Computers in Human Behavior

journal homepage: www.elsevier.com/locate/comphumbeh

Psychological research in the internet age: The quality of web-based data

Sarah R. Ramsey, Kristen L. Thompson, Melissa McKenzie, Alan Rosenbaum*

Department of Psychology, Northern Illinois University, DeKalb, IL, United States

ARTICLE INFO

Article history:
Received 16 July 2015
Received in revised form
15 December 2015
Accepted 18 December 2015
Available online xxx

Keywords:
Environmental effects
Human–computer interaction
Performance
Mass media
Test validity

ABSTRACT

The internet is increasingly used in psychological research to solicit participants and collect data. This paper includes two studies examining the quality of data obtained via web-based methods administered either inside or outside the lab. Both studies used item recognition accuracy as a proxy for attention to questions. Study 1 examined the extent to which undergraduate participants ($N = 504$) read and attended to questions either inside or outside the lab. Study 2 ($N = 744$) replicated Study 1, added a Mechanical Turk sample, and examined attention to non-intuitive survey instructions. Results indicated that participants demonstrated good item recognition, regardless of locale or sample; however, small sex effects on accuracy were found in both studies. Specifically, women were more accurate at identifying previously seen items than men in both Study 1 and Study 2. In Study 2, Mechanical Turk participants were more likely to read instructions than undergraduate participants, regardless of whether they participated inside or outside of the lab. The findings support the use of the internet for sampling purposes as well as survey administration, and suggest that researchers use care when studies include non-intuitive instructions.

© 2015 Published by Elsevier Ltd.

1. Introduction

Psychological research commonly utilizes undergraduate, subject pool participants, raising questions about both the internal and ecological validity of the research. Are subject pool participants, who are “voluntarily” compelled to participate, giving thoughtful responses? Or are they answering randomly and as quickly as possible? Are they even reading the questions? In traditional research, participants typically provide data in a supervised, laboratory setting (i.e., on-site). This allows investigators to monitor participant impairment (e.g., fatigue) or carelessness and observe whether participants complete surveys independently and without distraction (e.g., multi-tasking). The laboratory also confers a scientific aura that might result in more conscientious participation. Additionally, investigators are available to address questions or provide clarifications.

While much research is still collected in laboratory settings, web-based methodologies are proliferating (Prince, Litovsky, & Friedman-Wheeler, 2012). Prior to the internet, off-site study administration typically involved distributing and receiving materials via mail. However, the growth of internet usage has created new opportunities for researchers. Crowdsourcing services such as Amazon's Mechanical Turk (MTurk) permit recruitment of geographically and culturally diverse participants, enhancing external validity. Data is collected more quickly and prepared more accurately with web-administered surveys than via traditional methods (Gosling, Vazire, Srivastava, & John, 2004). Printing and mailing costs are reduced and participants are paid less, rendering these studies more eco-friendly and economical. Additionally, the flexibility and freedom regarding when and where one participates increases participant convenience and anonymity.

Despite the advantages, the quality of web-based data must be demonstrated. The lack of investigator oversight of off-site, web-based studies makes the amount of attention and care participants exercise in completing them unknown (Rosenbaum & Langhinrichsen-Rohling, 2006). However, some studies have begun to examine whether off-site, web-based methodologies produce comparable results to those administered on-site via traditional methods (e.g., Weigold, Weigold, & Russell, 2013; Zhang et al., 2012). Most, but not all, have demonstrated that on and off-site data are statistically comparable.

Chuah, Drasgow, and Roberts (2006) found no differences in the

* Corresponding author.
E-mail address: arosenbaum@niu.edu (A. Rosenbaum).

http://dx.doi.org/10.1016/j.chb.2015.12.049
0747-5632/© 2015 Published by Elsevier Ltd.

statistical equivalency of data collected via a proctored, on-site, web-administered survey and an un-proctored, off-site, web-administered survey. Templar and Lange (2008) also compared on-site to off-site web-based survey responses and found no differences between sites. However, some studies suggest important site differences beyond statistical equivalency. Paré and Cree (2009) identified differences in participants' item response latency when completing web-based tasks in on-site versus off-site conditions. When rating characteristics of images, off-site participants were significantly slower than on-site participants. Dandurand, Shultz, and Onishi (2008) compared participant accuracy on complex computer-administered problem-solving tasks and found that on-site participants were significantly more accurate than off-site participants. The authors of both studies speculated that decreased attention and/or multi-tasking in the off-site condition may explain these results.

While early studies assessing the comparability of on-site and off-site data collection utilized undergraduate samples, recent research has examined the comparability of research using undergraduate samples to those conducted entirely using web-based methods. The results of these studies have been mixed. Buhrmester, Kwang, and Gosling (2011) found consistency in questionnaire reliability scores across undergraduate and MTurk samples. Paolacci, Chandler, and Ipeirotis (2010) also found that undergraduate and MTurk participants responded similarly on judgment and decision-making tasks. Further, using one item to assess attention to survey items (i.e., "While watching the television, have you ever had a fatal heart attack?"), the authors found no significant differences in incorrect responses across groups. However, a study conducted by Goodman, Cryder, and Cheema (2013) incorporated complex instructions at the end of their survey and found undergraduate participants were significantly more likely to follow instructions than MTurk participants.

The internet has changed the way we live and it is starting to change the way we do research. Given that recent studies suggest there may be important differences in the extent to which on-site versus off-site participants attend to web-based surveys, we conducted two studies to further examine web-based research. Study 1 examined whether undergraduate participants (in both on-site, and off-site, web-based conditions) attended to survey questions and whether there were differences attributable to site. Study 2 replicated Study 1 and expanded upon it by adding an MTurk comparison group and also assessing the extent to which participants attended to survey instructions.

2. Study 1

Study 1 employed a novel strategy to examine the extent to which undergraduates attend to (i.e., read and process) survey items presented via the internet, either in the lab or at an off-site location (and time) of their choosing. Recognition accuracy of previously seen items served as a proxy for attending. Presumably, participants that read survey questions would be better at recognizing those questions (embedded in a second survey) compared with participants who responded without carefully reading the questions. Given the dearth of research regarding participant attending between on-site and off-site conditions, no formal hypotheses were proposed.

2.1. Method

2.1.1. Participants

The G*Power formula (Faul, Erdfelder, Lang, & Buchner, 2007) indicated that a total sample size of 158 was necessary to obtain a power of .95, assuming a small effect size (.25) and setting a significance level of $p < .05$ based on the statistical analyses conducted. Data was collected at a large, public Midwestern university for the duration of one semester, resulting in a sample of 504 undergraduate students ranging in age from 18 to 37 ($M = 19.35$; $SD = 2.02$). Additional descriptive statistics are presented in Tables 1 and 2. Participants received research credit applied to a course research requirement.

2.1.2. Measures

2.1.2.1. Demographic questionnaire. All participants completed a short demographic questionnaire assessing age, sex, race/ethnicity, and socioeconomic status. Off-site participants were also asked to specify the device they used (i.e., desktop computer, laptop, smart phone, other).

2.1.2.2. Questionnaires. Questionnaire 1 (Q1) included 50-items drawn from a variety of existing measures assessing both sensitive and non-sensitive topics. These included: the Balanced Time Perspectives Scale (Webster, 2011), the Attitudes Toward Emotions Scale (Harmon-Jones, Harmon-Jones, Amodio, & Gable, 2011), the Sexual Risk Survey (Turchik & Garske, 2009), the Short Sadistic Impulse Scale (O'Meara, Davies, & Hammond, 2011), the Behavioral Undercontrol Questionnaire (Stice, Myers, & Brown, 1998), the Caffeine Expectancy Questionnaire (Huntley & Juliano, 2012), the Alcohol Use Disorders Identification Test: Self-Report (Saunders, Aasland, Babor, De La Fuente, & Grant, 1993), and the Revised

Table 1
Participant descriptives: Study 1 and Study 2.

	On-site ($n = 250$)		Off-site ($n = 254$)		M-Turk –		Total ($N = 504$)	
	n	%	*N*	%	*n*	%	*N*	%
Study 1								
Gender								
Male	89	35.6	110	43.3	–	–	199	39.5
Female	161	64.4	144	56.7	–	–	305	60.5
Race/Ethnicity								
Caucasian	124	49.6	150	59.1	–	–	274	54.4
African American	62	24.8	50	19.7	–	–	112	22.2
Hispanic	24	9.6	26	10.2	–	–	50	9.9
Asian American	9	3.6	12	4.7	–	–	21	4.2
Native American	2	.8	0	0	–	–	2	.4
Indian	3	1.2	2	.8	–	–	5	1.0
Multiracial	25	10.0	13	5.1	–	–	38	7.5
Missing	1	.4	1	.4	–	–	2	.4
	On-site ($n = 251$)		Off-site ($n = 247$)		M-Turk ($n = 246$)		Total ($N = 744$)	
Study 2								
Gender								
Male	101	40.2	101	40.9	105	42.7	307	41.3
Female	150	59.8	144	58.3	141	57.3	435	58.5
Education								
Some high school	0	0	0	0	3	1.2	3	.4
High school diploma	0	0	0	0	24	9.8	24	3.23
Some college	242	96.4	228	92.3	73	29.7	543	72.98
Bachelor's/Assoc.	9	3.6	18	7.3	106	43	133	17.9
Graduate education	0	0	0	0	40	16.3	41	5.5
Race/Ethnicity								
Caucasian	145	57.8	143	57.9	188	76.4	476	64.0
African American	51	20.3	43	17.4	19	7.7	113	15.2
Hispanic	16	6.4	27	10.9	5	2.0	48	6.5
Asian American	16	6.4	13	5.3	14	5.7	43	5.8
Native American	1	.4	0	0	4	1.6	5	.7
Indian	1	.4	0	0	1	.4	2	.3
Pacific Islander	1	.4	1	.4	0	0	2	.3
Multiracial	16	6.4	17	6.9	13	5.3	46	6.2
Other	4	1.6	2	.8	1	.4	7	.9
Missing	0	0	1	.4	1	.4	2	.3

Download English Version:

https://daneshyari.com/en/article/6837578

Download Persian Version:

https://daneshyari.com/article/6837578

Daneshyari.com