Social Science Computer Review

Mode Effects in Free-list Elicitation : Comparing Oral, Written, and Web-based Data Collection

Clarence C. Gravlee, H. Russell Bernard, Chad R. Maxwell and Aryeh Jacobsohn

The online version of this article can be found at:

Additional services and information for *Social Science Computer Review* can be found at:

Email Alerts: http://ssc.sagepub.com/cgi/alerts

Subscriptions: http://ssc.sagepub.com/subscriptions

Reprints: http://www.sagepub.com/journalsReprints.nav

Permissions: http://www.sagepub.com/journalsPermissions.nav

Citations: http://ssc.sagepub.com/content/31/1/119.refs.html

>> Version of Record - Jan 18, 2013

OnlineFirst Version of Record - Sep 10, 2012

What is This?

Article

Mode Effects in Free-list Elicitation: Comparing Oral, Written, and Web-based Data Collection

Clarence C. Gravlee[1], H. Russell Bernard[2], Chad R. Maxwell[3], and Aryeh Jacobsohn[4]

Abstract
The growth of the Internet opens new possibilities for web-based data collection in cognitive anthropology. This study examines whether free-list data collected online are comparable to those collected with face to face interviews or with self-administered paper questionnaires. We collected free lists for two cultural domains in the United States: one diffuse (*things that mothers do*) and one relatively well defined (*racial and ethnic groups*). We selected a purposive sample of 318 university students and randomly assigned participants to provide free lists for one of these domains using a web-based survey, a face to face interview, or a self-administered paper questionnaire. All three modes identified the same set of salient concepts in each domain. Median list length per respondent varied across modes in response to a standard free-list question and to supplementary probes. For the well-defined domain of "racial and ethnic groups," supplementary probes widened differences among modes; for the more diffuse domain of "things that mothers do," probes erased evidence of mode effects. Collecting free lists online is viable but may yield different results depending on the study population and attributes of the cultural domains.

Web-based data collection has become a staple of survey research (Couper, 2008) and there is a vast literature comparing the results of web-based surveys to mail, face to face, and telephone questionnaires (Dillman et al., 2009; Eaton et al., 2010; Jäckle, Roberts, & Lynn, 2010). Here,

---

[1] University of Florida, Gainesville, FL, USA
[2] University of Florida, Gainesville, FL, USA
[3] Starcom Worldwide, Chicago, IL, USA
[4] Razorfish c/o Aryeh Jacobsohn, Chicago, IL, USA

Corresponding Author:
Clarence C. Gravlee, Department of Anthropology, University of Florida, PO Box 117305, 1112 Turlington Hall, Gainesville, FL 32611, USA.
Email: cgravlee@ufl.edu

we extend this literature on mode effects to free listing, a technique commonly used by cognitive anthropologists, and examine whether data collected on the Internet are comparable to those elicited with face to face interviews or with self-administered, paper questionnaires.

Free listing is a simple, open-ended technique in which researchers ask respondents to "list as many Xs" as they can, where X refers to some cultural domain (e.g., illnesses, plants). This technique was developed in psychology in the 1950s as a way to study associations among, and the salience of, concepts (Bousfield & Barclay, 1950). It has been used in anthropology since the 1960s to elicit the contents of cultural domains, including kinship terms (Romney & D'Andrade, 1964), life events (Hurwicz, Durham, Boyd-Davis, Gatz, & Bengtson, 1992), everyday activities (Harman, 2001), characteristics of innovative teaching (Jaskyte, Taylor, & Smariga, 2009), ethnic categories (Gravlee, 2005), sexual health problems faced by men (Verma, Rangaiyan, Singh, Sharma, & Pelto, 2001), and many others.

In the last 20 years, free listing has become widely used in cognitive anthropology as well as in marketing, psychology, public health, and other disciplines. This trend is due in part to the method's intrinsic appeal: It is a simple yet effective technique with many applications. We also suspect that the availability of software, such as ANTHROPAC (Borgatti, 1996), has propelled the growth of free listing because it makes the analysis of free-list data relatively easy. Now, as mobile and web-based data collection proliferate (Buskirk & Andrus, 2012; Couper, 2005; Gravlee, 2002; Peytchev & Hill, 2010; Vicente, Reis, & Santos, 2009) software has the potential to change the way free-list data are collected, too.

Traditionally, free lists are collected verbally or in writing, though many researchers recommend collecting written free lists if informants are literate (Borgatti, 1998; Schrauf & Sanchez, 2008). Quinlan (2005), for example, collected oral free lists in rural Dominica but found it difficult to obtain independent responses because people nearby often contributed items or prompted recall of items that respondents otherwise might not have listed. To avoid this bias, Quinlan recommends collecting free lists in writing whenever possible. She also suggests that written free lists may be more comprehensive because the process is more formal and encourages respondents to go at their own pace. The goal of this article is to test the common preference for written over face to face free lists and to determine whether the free lists collected over the web are comparable to those collected with conventional methods.

Method

Study design

We tested for mode effects in a purposive sample of university students in the United States who were randomly assigned to one of two cultural domains—"racial and ethnic groups" and "things that mothers do"—and to one of three data collection modes—face to face interview, paper questionnaire, and web survey. The rationale for studying two domains was that we hypothesized that the presence and magnitude of mode effects may depend on how "listable" a domain is. For example, respondents in self-administered modes might give up sooner in hard-to-list domains (like *things that mothers do*, which is expressed in phrases) than they would in more listable domains (like *racial and ethnic groups*, which can be listed as single items).

Sampling

Our sample consisted of university students between the ages of 17 and 22. We used a purposive sampling design with nested quotas for gender (men, women) and self-identified race or ethnicity (Black or African American, Hispanic or Latino, and White). We set quotas for these attributes because they are germane to the two cultural domains we studied. The design called for eight

respondents for each of the six race-gender categories (e.g., Black men), for a total of 48 respondents per mode per cultural domain ($N = 288$). Interviewers recruited participants from a variety of places on campus where they were likely to find different types of students (e.g., business school, student union, select student organizations, classrooms). We recruited all participants face to face and randomly assigned them to a mode and cultural domain in which the relevant quotas had not been met; to fill some quotas we used chain referral. Respondents assigned to face to face or paper questionnaires participated on the spot; respondents assigned to the web survey were given a card with the URL of the survey. We offered a \$5 Starbucks® gift card to all respondents who completed the interview, questionnaire, or web survey.

Development of Instruments

For each cultural domain, we developed a structured interview guide, a paper questionnaire, and an Internet survey using Educara Survey software. Educara Survey (http://sourceforge.net/projects/educarasurvey) is open-source software. In addition to standard survey question formats, Educara Survey supports structured ethnographic methods commonly used in cognitive anthropology, including free lists, pile sorts, triad tests with lambda designs, frame substitution, and paired comparisons with lambda designs (Weller & Romney, 1988). Data collected with Educara Survey are exported as Microsoft Excel® files or as XML files. English and Spanish localizations of the software are available, and the built-in templates for the user interface can be customized.

Weller and Romney (1988) offer suggestions about how to phrase questions and follow-up probes in face to face interviews. Brewer (2002) and colleagues (Brewer & Garrett, 2001; Brewer, Garrett, & Kulasingam, 1999) showed that three kinds of probing increased the mean number of items per respondent in free lists by between 4% and 49% in face to face interviews. To the extent possible, we incorporated these probes into the study design. The first probe, nonspecific prompting, involves asking generic questions such as "What other things can you think of?" when respondents stop listing items. This technique was only possible in face to face interviews. The second probe involves reading respondents' initial list back to them and asking if the list is complete. We used this technique in face to face interviews and adapted it to the pen-and-paper questionnaire by including a follow-up question on a separate page asking respondents to add any additional terms that occurred to them after rereading their original lists.

The final probe, semantic cueing, presents respondents with each item in their list, one at a time, and asks what other items that they have not already mentioned are like that item. We used this technique in face to face interviews and on the web for both domains, taking advantage of the software's ability to automate probes based on previous responses. For example, if a respondent in the "racial and ethnic groups" domain initially listed "African American," "Hispanic," "White," and so on in the web survey, the next screen would present a new open-ended text box with the following prompt: "You mentioned African American. Please list all the other racial and ethnic groups like African American that you haven't already mentioned." This probe would be repeated for each term in the initial free list, as would be the case in face to face interviews.

Analysis

We imported free-list data to ANTHROPAC software (Borgatti, 1996) and examined the frequency distribution of the raw free-list items. Next, we cleaned and recoded the data to correct misspellings (e.g., *Caucasion* → *Caucasian*) and to combine idiosyncratic expressions of the same concept (e.g., *care for kids* → *take care of children*). Some coding decisions were less straightforward. For example, do "African American" and "Black" mean the same thing? Are "cook" and "make dinner" equivalent? We answered such questions based on our own interpretations and, in some cases, by asking respondents in face to face interviews to clarify whether closely related categories were distinct.

Our analysis focused on three outcomes: coherence, completeness, and content. To test our expectation that "racial and ethnic groups" would be more coherent than "things that mothers do," we examined the relative frequency of items as a proxy for agreement among respondents about what constitutes the core of each domain. To test for mode effects in the completeness of lists, we compared the average list length and total number of items produced by each mode within domains. Last, we tested for mode effects in the content of each domain by conducting qualitative and quantitative comparisons of the most salient concepts identified across modes.

Results

The final, pooled sample ($N = 318$) exceeded the 288 specified by the quota sampling design but maintained the expected distribution of gender and self-identified race or ethnicity within each mode and cultural domain. For "racial and ethnic groups," we collected 48 free lists in face to face interviews, 53 with paper questionnaires, and 59 over the Internet. For "things that mothers do," we collected 50 free lists in face to face interviews, 53 with paper questionnaires, and 55 over the Internet. In general, the distribution of attributes potentially related to cultural knowledge about the domains, such as gender, self-identified race, mother's education (a proxy for socioeconomic status), or political affiliation does not vary by mode. The one exception is that respondents in face to face interviews for "things that mothers do" reported lower levels of mothers' education, on average, than did respondents in other modes ($\chi^2 = 10.98$, $df = 4$, $p = .027$). For example, only 8% of face to face respondents reported that their mothers had a graduate degree, as compared to 28% among respondents assigned to paper questionnaires.

Coherence

To compare the relative coherence of domains, Figure 1 shows scree plots for the frequency distribution of items across all modes. Both domains yielded many items. Respondents identified 523 "racial and ethnic groups" and 853 "things that mothers do." However, most items were idiosyncratic, illustrated by the long tails in panel (a) of Figure 1. Beyond this general pattern, Figure 1 confirms our expectation that there would be more idiosyncratic terms and fewer shared items in the motherhood domain. Only three terms—cook, love, and household cleaning—were mentioned by more than half of respondents, and no term had a frequency of greater than 68% across all three modes. By contrast, nine terms surpassed 50% agreement among respondents for the "racial and ethnic groups" domain, with the most common term appearing in 82% of respondents' lists. The difference is more apparent in panel (b), which focuses on only the 100 most frequently listed items in each domain. This comparison highlights the larger core of shared terms in the domain of "racial and ethnic groups," with almost twice as many items listed by at least 20% of respondents in the race and ethnicity domain, as compared to the motherhood domain (35 vs. 16, respectively).

Completeness

Figure 2 tests for mode effects in the completeness of free lists, as measured by the median number of items elicited per respondent, before and after probes. We highlight three findings.

1. There is some indication that self-administered free lists are slightly more productive than are face to face interviews. Despite using three supplementary probes in the face to face interviews and only one in the other two modes, median list length in oral free lists is never higher than it is in the other two modes. However, neither paper- nor web-based free lists are consistently longer than the other.
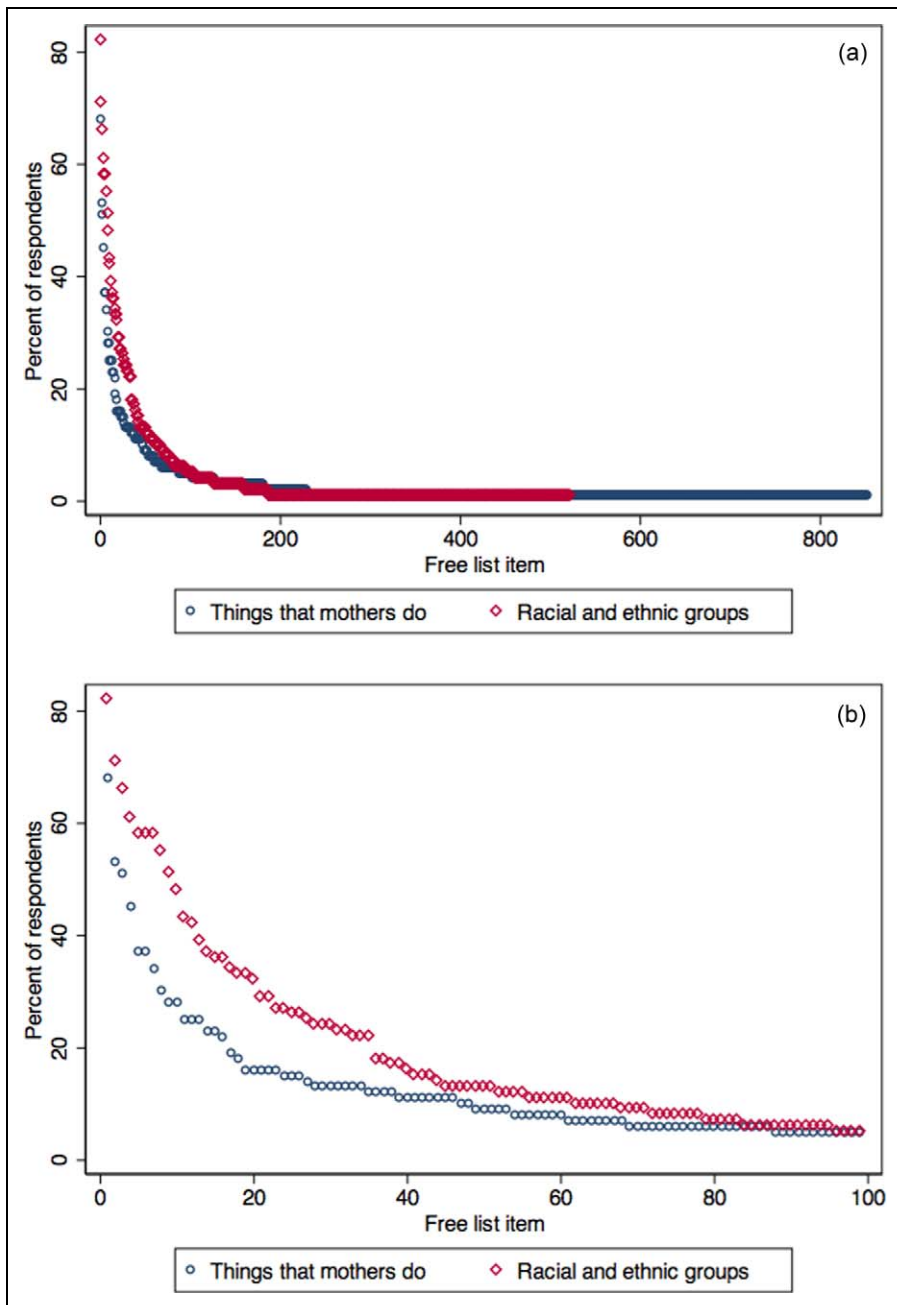
Figure 1. Frequency distribution of free-list items in two cultural domains. *Note.* Figure is available in full color in the online version at ssc.sagepub.com

2. In both domains, supplementary probes were most productive on the web. Semantic cueing on the web increased the median number of items per respondent by 111% for "racial and ethnic groups" and 86% for "things that mothers do."

3. The differential effectiveness of probes had opposite impacts on the two cultural domains. For "racial and ethnic groups," probes widened the differences among modes, but for "things that
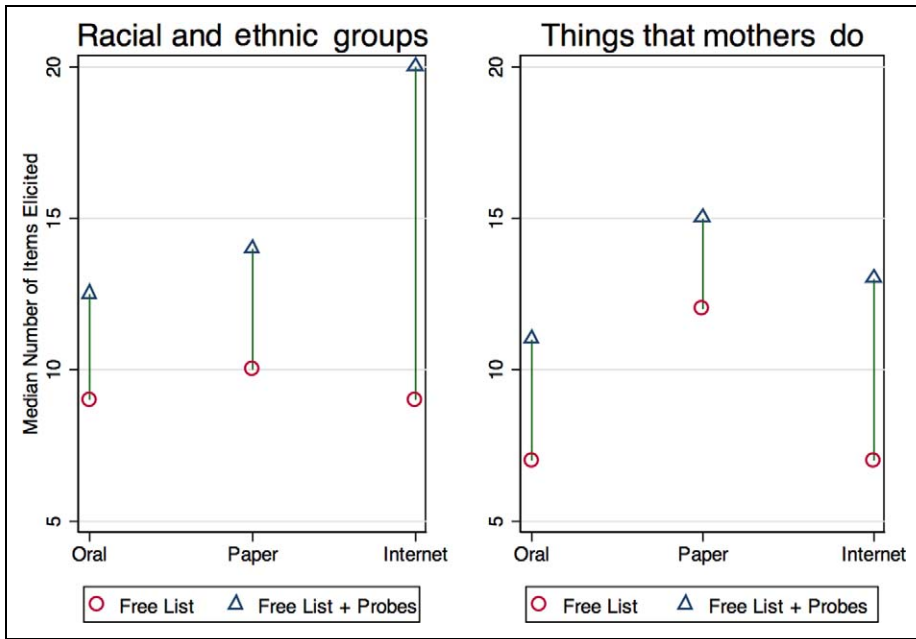
Figure 2. Median number of items listed per respondent, before and after probes, by cultural domain. *Note.* Figure is available in full color in the online version at ssc.sagepub.com

mothers do,'' they helped close the gap. This trend is confirmed by a Kruskal-Wallis test of differences in list length across modes. Before probes, differences are statistically significant for ''things that mothers do'' ($\chi^2 = 19.15$, $df = 2$, $p = .001$) but not for ''racial and ethnic groups'' ($\chi^2 = 1.30$, $df = 2$, $p = .519$). After probes, the pattern reverses: Differences are significant for ''racial and ethnic groups'' ($\chi^2 = 11.59$, $df = 2$, $p = .003$) but not for ''things that mothers do'' ($\chi^2 = 3.72$, $df = 2$, $p = .155$).

Content

Last, we examine possible mode effects in the content of cultural domains, as determined by the most salient concepts across modes. To evaluate salience, we use frequency (percentage of respondents who listed an item) and Smith's S, which incorporates both frequency and the average rank of an item within respondents' lists (Smith, 1993; Sutrop, 2001). These measures help identify core-periphery structures in a domain, and they are often used to select items for inclusion in subsequent types of data collection, like pile sorts or triad tests (Weller & Romney, 1988).

Tables 1 and 2 show the 25 most salient items in each domain, by mode. Overall, we find substantial overlap in the most salient terms across modes. For ''racial and ethnic groups'' (Table 1), the five most salient terms are identical across modes, although the rank order varies. Subsequent terms are generally at a lower level of contrast but occur in a recognizable pattern across modes (e.g., ''German'' holds ranks of 21, 18, and 20). There is also substantial overlap for ''things that mothers do'' (Table 2). ''Cook,'' ''household cleaning,'' and ''love'' are among the most salient terms in each mode; others that rank near the top include ''take care of kids'' and ''work.'' As salience decreases, there is generally more variability across modes, but some less salient terms occur at similar places on all three lists (e.g., ''breastfeeding'' holds ranks of 20, 23, and 21).

Table 1. Most Salient "Racial and Ethnic Groups," by Mode of Free-list Elicitation.

Rank	Face to Face interview			Paper Qestionnaire			Web Survey		
	Item	Frequency (%)	Salience	Item	Frequency (%)	Salience	Item	Frequency (%)	Salience
1	White	83	.759	White	83	.635	White	79	.718
2	Asian	88	.718	Hispanic	68	.523	Hispanic	68	.553
3	African American	69	.583	Asian	66	.446	Asian	63	.501
4	Hispanic	63	.537	African American	51	.412	Black	54	.479
5	Black	54	.475	Black	45	.380	African American	55	.410
6	Chinese	88	.469	Native American	60	.335	Indian	64	.388
7	Japanese	85	.459	Jewish	47	.244	Native American	59	.384
8	European	69	.434	Indian	47	.223	Chinese	63	.342
9	Native American	54	.380	Mexican	32	.207	Japanese	59	.310
10	Indian	63	.370	Chinese	34	.190	Mexican	48	.244
11	African	65	.368	African	38	.181	African	45	.244
12	Cuban	56	.298	Spanish	28	.170	Cuban	39	.220
13	Mexican	50	.298	European	32	.167	Spanish	34	.199
14	Jamaican	48	.271	Pacific Islander	34	.166	Puerto Rican	32	.186
15	Haitian	50	.265	Puerto Rican	23	.166	Haitian	34	.184
16	Korean	58	.260	Cuban	23	.139	Latino	27	.179
17	Puerto Rican	50	.260	Haitian	26	.136	Jamaican	30	.178
18	English	46	.226	German	25	.126	Vietnamese	39	.166
19	Middle Eastern	35	.214	Arab	23	.122	European	29	.160
20	Spanish	50	.210	Dominican	19	.122	German	25	.157
21	German	40	.209	Japanese	23	.117	Italian	27	.154
22	Latino	31	.206	Latino	21	.103	English	32	.153
23	French	40	.205	English	23	.101	Jewish	27	.153
24	Caribbean	29	.196	Italian	21	.099	French	27	.145
25	South American	40	.188	French	17	.095	Pacific Islander	25	.139

125

Table 2. Most Salient "Things That Mothers Do," by Mode of Free-list Elicitation.

	Face to Face Interview			Paper Questionnaire			Web Survey		
Rank	Item	Frequency (%)	Salience	Item	Frequency (%)	Salience	Item	Frequency(%)	Salience
1	Cook	74	.643	Cook	74	.552	Cook	56	.387
2	Household cleaning	58	.461	Household cleaning	53	.376	Love	53	.413
3	Take care of kids	54	.459	Love	60	.373	Household cleaning	44	.291
4	Work	58	.417	Work	49	.309	Feed	40	.283
5	Love	46	.361	Take care of kids	28	.247	Discipline	35	.225
6	Nurture	42	.321	Nurture	25	.205	Take care of kids	29	.224
7	Give birth	34	.293	Teach	26	.181	Teach	29	.205
8	Laundry	36	.251	Give advice	34	.180	Work	29	.178
9	Discipline	42	.247	Feed	26	.179	Protect	27	.178
10	Feed	34	.238	Care	21	.178	Care	25	.196
11	Protect	32	.223	Discipline	36	.163	Nurture	25	.182
12	Transport kids	36	.216	Give birth	23	.144	Care for sick	24	.132
13	Teach	28	.188	Laundry	21	.136	Give advice	22	.117
14	Teach values	38	.175	Take kids to school	23	.133	Change diapers	20	.143
15	Take kids to school	24	.175	Hug	17	.106	Listen	20	.118
16	Support emotionally	30	.168	Listen	17	.099	Give birth	18	.153
17	Care	22	.163	Support	19	.097	Worry	18	.138
18	Give advice	28	.160	Breastfeed	13	.095	Hug	18	.108
19	Provide safe and stable homes	24	.155	Support emotionally	13	.086	Laundry	18	.104
20	Breastfeed	20	.155	Protect	17	.082	Sing	16	.096
21	Help with homework	26	.152	Comfort	15	.082	Kiss	16	.093
22	Household chores	18	.144	Teach values	15	.074	Comfort	15	.110
23	Support financially	18	.126	Guide	13	.074	Cry	13	.089
24	Provide	20	.119	Give kids money	15	.067	Shout/Yell/Scream	13	.082
25	Shop for groceries	26	.116	Help with homework	13	.063	Breastfeed	11	.098

Note. Items are sorted in descending order of salience (Smith's S).
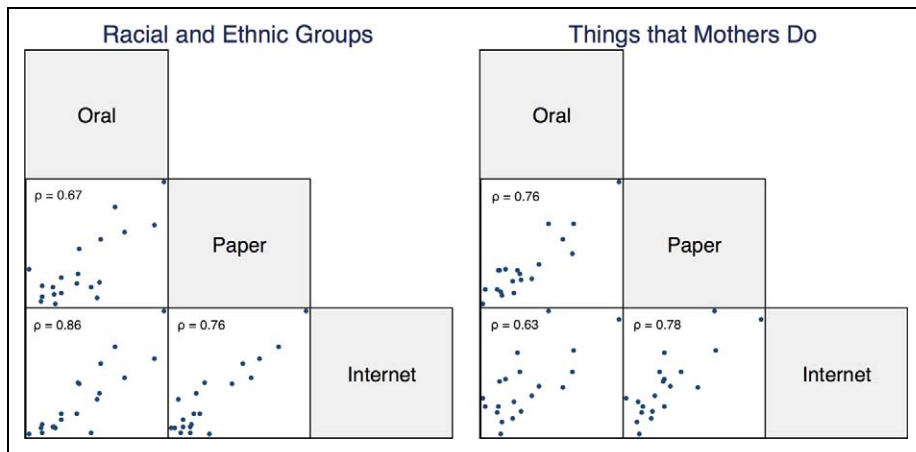
Figure 3. Scatterplot matrix of item salience for three modes of free-list elicitation, by cultural domain.

An unexpected finding in Tables 1 and 2 is that the level of agreement among respondents is higher in face to face interviews than it is in either self-administered questionnaires or web-based surveys. This pattern holds for both domains. For example, in the domain of "racial and ethnic groups" (Table 1), 17 terms were listed by at least half of respondents in face to face interviews, but only five terms met that threshold for paper questionnaires and nine for web surveys. Likewise, for "things that mothers do" (Table 2), 18 terms were mentioned by at least one-quarter of respondents in face to face interviews, but only 10 were in paper questionnaires and 11 in web surveys. Thus, for both domains, face to face interviews yielded greater consensus about the core items, while paper questionnaires and web surveys resulted in less agreement among respondents. This finding did not negate the overall similarity in content, however.

Figure 3 confirms the overlap in content across modes. In this display, the unit of analysis is the item, and the axes represent items' salience, as estimated for each mode. In general, the associations are strong, with Spearman's rank-order correlation (ρ) ranging from .63 to .86. There is no clear pattern regarding which modes are most alike. For "racial and ethnic groups," the strongest correlation is between face to face interviews and Internet surveys ($\rho = .86$); for "things that mothers do," the strongest correlation is between Internet and paper questionnaires ($\rho = .78$), followed closely by oral and paper administration ($\rho = .76$).

Discussion

The goal of this study was to determine whether web-based free listing produced results that are comparable to traditional methods, including oral and written elicitation. We found that all three modes identified the same set of salient concepts in two cultural domains, but there were differences among modes in the number of items elicited per respondent, the effectiveness of probes, and the level of agreement among respondents. These findings suggest that collecting free lists online is viable but may produce different results, depending on the study population and attributes of the cultural domains.

Free lists are most often used to identify the consensual elements of a domain, items identified by a large share of respondents. For the two domains we studied, all three modes identified similar core elements and yielded similar estimates of items' salience. In particular, web-based free lists and conventional methods produced comparable estimates. In both domains, the strongest correlations in salience between modes involved web-based free lists ($\rho = .86$ between web and oral for "racial

and ethnic groups''; ρ = .78 between web and paper for "things that mothers do"). In terms of content, then, our study indicates that collecting free lists over the Internet provides the same answers that traditional methods do.

The similarity in content persisted despite modest but significant differences in the number of items elicited per respondent under certain circumstances. Before probes, there was no evidence of mode effects in list length for "racial and ethnic groups," but there was for "things that mothers do." It is possible that the difference in the motherhood domain was a result of sampling error rather than a mode effect: List length was greatest among respondents to paper questionnaires, who also reported higher levels of education among their own mothers. At any rate, the difference among modes was erased by probing, largely because supplementary probes were more productive on the web than in face to face interviews or paper questionnaires. For "racial and ethnic groups," the effectiveness of online probes produced a difference among modes where none had been evident before.

It appears that the effectiveness of probing on the web is at least partially domain dependent. Probing was more effective for "racial and ethnic groups" than for "things that mothers do" because participants in the race and ethnicity domain generally responded to semantic cues by listing items at a lower level of contrast. For example, when asked what other terms are like "Hispanic," a typical response might include specific Hispanic groups defined by nationality (e.g., Puerto Rican, Cuban, Mexican, Guatemalan). In other words, the items generated by semantic cueing reflect the taxonomic structure of the cultural domain; in domains with a clear hierarchical taxonomy, semantic cueing is likely to produce many terms. This interpretation helps explain the relative productivity of supplemental probes on the web as compared to paper questionnaires, in which we were not able to implement semantic cueing. But it does not explain why semantic cueing on the web was more productive than semantic cueing face to face—even in the motherhood domain, which does not have the same hierarchical taxonomic structure that the domain of racial and ethnic groups does.

We see four possible reasons that semantic cueing was more effective online. First, web-based free-list elicitation may be less susceptible to social desirability effects than are traditional face to face interviews. It is possible that face to face respondents censored themselves more often than web respondents did to maintain a socially desirable self-presentation. This possibility is consistent with the mode effects literature in survey research. Respondents who complete self-administered questionnaires generally report higher levels of stigmatized behaviors and beliefs (Baker et al., 2010; Link & Mokdad, 2005), and there is some evidence that computer-assisted questionnaires are even less susceptible to social desirability bias than are paper-and-pencil questionnaires (Eaton et al., 2010; Vereecken & Maes, 2006; Wright, Aquilino, & Supple, 1998). Future research should be designed specifically to test for social desirability effects across modes of free-list elicitation.

Second, some fraction of the mode effect in semantic cueing may be due to satisficing (Krosnick, 1991). Satisficing refers to behaviors that survey respondents engage in to manage the cognitive demands of responding to a survey. The mode effects literature on satisficing is mixed. Some studies suggest that computer modes are associated with less satisficing (Chang & Krosnick, 2010), but others show the reverse (Heerwegh & Loosveldt, 2009). It is likely that satisficing depends on multiple design features, not just mode. In our study, it is possible that interviewers did not motivate face to face respondents to answer thoroughly, resulting in fewer additional items than in web-based free lists. Another possibility is that the combination of three types of probes in face to face interviews, as compared to one in web surveys, imposed an additional cognitive burden that respondents managed by listing fewer additional items in response to semantic cueing.

Third, inherent technological advantages of web surveys may help explain the productivity of online semantic cueing. Some survey researchers suggest that the visual layout and ease of typing

on web surveys encourages respondents to provide rich responses to open-ended questions (Denscombe, 2007). Indeed, Smyth and colleagues (2009) anticipate the proliferation of open-ended questions in survey research because web surveys make it possible to collect more detailed responses than are feasible in other modes. Another technological advantage of collecting free lists online is that software automates probing in a way that is not possible in other modes. Paper questionnaires cannot change dynamically to incorporate previous answers into subsequent questions, as semantic cueing requires. In face to face mode, semantic cueing places a large burden on interviewers who may be tempted to cut back. Web-based software eliminates these reasons for item nonresponse.

Fourth, our recruitment strategy may have contributed to differences among modes. Respondents assigned to web-based free listing could participate at their leisure, but face to face and questionnaire respondents were asked to participate on the spot. If we intercepted respondents at an inconvenient time (e.g., between classes, while they were talking to friends), respondents assigned to conventional methods may have curtailed their answers. Web respondents, by contrast, could respond at their leisure and may have been able to spend more time participating in the study. Future research should alter recruitment strategies, including web-based sampling strategies, and measure how much time people spend on free lists collected via different modes.

The effectiveness of web-based probes must be weighed against our finding that the level of agreement among respondents was highest among face to face respondents in both domains. Agreement among respondents is an important metric in free-list data because it is regarded as evidence of shared cultural knowledge. Differential agreement across modes, then, could alter the inferences researchers make about a cultural domain.

We wonder whether satisficing and social desirability might also help explain this finding. Respondents in self-administered modes (web and paper) may exert less effort to list the items they know because there is no interviewer to encourage them to answer as fully as possible. This effect would be consistent with some survey research indicating that satisficing is more common in self-administered modes (Heerwegh & Loosveldt, 2009), but it is difficult to reconcile with our other findings. If respondents exerted less effort in self-administered modes, we would expect them to generate fewer items, on average, than do respondents in face to face interviews. Yet we find the reverse. In both domains, face to face interviews produced the *fewest* items per respondent. Satisficing, therefore, is an unlikely explanation for the lower level of agreement among respondents in self-administered modes.

Another possibility is that social desirability might actually enhance measurement validity in face to face free listing. Typically, the purpose of collecting free lists is to elicit cultural knowledge. From the perspective of cognitive anthropology, cultural knowledge refers to shared and socially transmitted understandings that orient action and guide interpretation of everyday life (D'Andrade, 1981). It could be that the social interaction involved in a face to face interview primes respondents to focus on shared items at the core of a domain because presenting oneself as culturally knowledgeable is socially desirable. Thus, respondents in face to face interviews may be more likely than respondents in other modes to concentrate on the items that everybody is presumed to know. This possibility is an intriguing direction for further research.

Like many studies on mode effects, ours is limited by the fact that the sample is drawn from university students in the United States. This population is technologically savvy and accustomed to ubiquitous Internet access. Thus, their performance on web-based free lists may be a best-case scenario. A critical direction for future research is to examine mode effects in other populations that have less fluency with computers and the Internet. This step is especially important to assess the viability of web-based free listing in the contexts where many anthropologists and other field researchers collect free lists.

Web-based collection of free lists is likely to grow because it promises the same benefits that have made web surveys pervasive: reduced cost per respondent, potential for automation (e.g., skip

patterns, text piping), ability to incorporate multimedia, and the opportunity to obtain larger numbers of geographically dispersed respondents than ever before. The Internet also opens new opportunities for local collaborators or informants to continue collecting data even if the primary researcher leaves the field. Our study suggests that researchers could reap these benefits while producing results that are comparable to conventional methods, but this possibility remains to be tested in other domains and cultural contexts.

Authors' Note

Declaration of Conflicting Interests

Funding

References

Baker, R., Blumberg, S. J., Brick, J. M., Couper, M. P., Courtright, M., Dennis, J. M., . . . & Zahs, D. (2010). Research synthesis: AAPOR report on online panels. *Public Opinion Quarterly*, *74*, 711–781.

Borgatti, S. (1996). *ANTHROPAC*. Lexington, KY: Analytic Technologies. Retrieved June 22, 2012, from http://www.analytictech.com/anthropac/anthropac.htm

Borgatti, S. P. (1998). Elicitation techniques for cultural domain analysis. In J. J. Schensul & M. D. LeCompte (Eds.), *The ethnographer's toolkit* (Vol. 3, pp. 1–26). Walnut Creek, CA: AltaMira Press.

Bousfield, W., & Barclay, W. (1950). The relationship between order and frequency of occurrence of restricted associative responses. *Journal of Experimental Psychology*, *40*, 643–647.

Brewer, D. (2002). Supplementary interviewing techniques to maximize output in free listing tasks. *Field Methods*, *14*, 108–118.

Brewer, D., & Garrett, S. B. (2001). Evaluation of interviewing techniques to enhance recall of sexual and drug injection partners. *Sexually Transmitted Diseases*, *28*, 666–677.

Brewer, D., Garrett, S. B., & Kulasingam, S. (1999). Forgetting as a cause of incomplete reporting of sexual and drug injection partners. *Sexually Transmitted Diseases*, *26*, 166–176.

Buskirk, T. D., & Andrus, C. (2012). Smart surveys for smart phones: Exploring various approaches for conducting online mobile surveys via smartphones. *Survey Practice*. Retrieved June 22, 2012, from http://surveypractice.wordpress.com/2012/02/21/smart-surveys-for-smart-phones/

Chang, L., & Krosnick, J. A. (2010). Comparing oral interviewing with self-administered computerized questionnaires: An experiment. *Public Opinion Quarterly*, *74*, 154–167.

Couper, M. P. (2005). Technology trends in survey data collection. *Social Science Computer Review*, *23*, 486–501.

Couper, M. P. (2008). *Designing effective web surveys*. New York: Cambridge University Press.

D'Andrade, R. (1981). The cultural part of cognition. *Cognitive Science*, *5*, 179–195.

Denscombe, M. (2007). The length of responses to open-ended questions: A comparison of online and paper questionnaires in terms of a mode effect. *Social Science Computer Review*, *26*, 359–368.

Dillman, D. A., Phelps, G., Tortora, R., Swift, K., Kohrell, J., Berck, J., & Messer, B. L. (2009). Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the Internet. *Social Science Research*, *38*, 1–18.

Eaton, D. K., Brener, N. D., Kann, L., Denniston, M. M., McManus, T., Kyle, T. M., . . . & Ross, J. G. (2010). Comparison of paper-and-pencil versus web administration of the Youth Risk Behavior Survey (YRBS): Risk behavior prevalence estimates. *Evaluation Review*, *34*, 137–153.

Gravlee, C. C. (2002). Mobile computer-assisted personal interviewing (MCAPI) with handheld computers: The Entryware System v3.0. *Field Methods*, *14*, 322–336.

Gravlee, C. C. (2005). Ethnic classification in southeastern Puerto Rico: The cultural model of "color." *Social Forces*, *83*, 949–970.

Harman, R. C. (2001). Activities of contemporary Mayan elders. *Journal of Cross-Cultural Gerontology*, *16*, 57–77.

Heerwegh, D., & Loosveldt, G. (2009). Face-to-face versus web surveying in a high-Internet-coverage population: Differences in response quality. *Public Opinion Quarterly*, *72*, 836–846.

Hurwicz, M. L., Durham, C. C., Boyd-Davis, S. L., Gatz, M., & Bengtson, V. L. (1992). Salient life events in three-generation families. *Journal of Gerontology*, *47*, 11–13.

Jäckle, A., Roberts, C., & Lynn, P. (2010). Assessing the effect of data collection mode on measurement. *International Statistical Review*, *78*, 3–20.

Jaskyte, K., Taylor, H., & Smariga, R. (2009). Student and faculty perceptions of innovative teaching. *Creativity Research Journal*, *21*, 111–116.

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, *5*, 213–236.

Link, M. W., & Mokdad, A. H. (2005). Effects of survey mode on self-reports of adult alcohol consumption: A comparison of mail, web and telephone approaches. *Journal of Studies on Alcohol*, *66*, 239–245.

Peytchev, A., & Hill, C. A. (2010). Experiments in mobile web survey design: Similarities to other modes and unique considerations. *Social Science Computer Review*, *28*, 319–335.

Quinlan, M. (2005). Considerations for collecting freelists in the field: Examples from ethnobotany. *Field Methods*, *17*, 219–234.

Romney, A., & D'Andrade, R. (1964). Cognitive aspects of English kin terms. *American Anthropologist*, *66*, 146–170.

Schrauf, R. W., & Sanchez, J. (2008). Using freelisting to identify, assess, and characterize age differences in shared cultural domains. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *63*, S385–S393.

Smith, J. (1993). Using ANTHROPAC 3.5 and a spreadsheet to compute a free list salience index. *Cultural Anthropology Methods*, *5*, 1–3.

Smyth, J. D., Dillman, D. A., Christian, L. M., & Mcbride, M. (2009). Open-ended questions in web surveys. *Public Opinion Quarterly*, *73*, 325–337.

Sutrop, U. (2001). List task and a cognitive salience index. *Field Methods*, *13*, 263–276.

Vereecken, C. A., & Maes, L. (2006). Comparison of a computer-administered and paper-and-pencil-administered questionnaire on health and lifestyle behaviors. *Journal of Adolescent Health*, *38*, 426–432.

Verma, R. K., Rangaiyan, G., Singh, R., Sharma, S., & Pelto, P. J. (2001). A study of male sexual health problems in a Mumbai slum population. *Culture, Health & Sexuality*, *3*, 339–352.

Vicente, P., Reis, E., & Santos, M. (2009). Using mobile phones for survey research. *International Journal of Market Research*, *51*, 613–633.

Weller, S. C., & Romney, A. K. (1988). *Systematic data collection*. Newbury Park, CA: Sage.

Wright, D. L., Aquilino, W. S., & Supple, A. J. (1998). A comparison of computer-assisted and paper-and-pencil self-administered questionnaires in a survey on smoking, alcohol, and drug use. *Public Opinion Quarterly*, *62*, 331–353.

Author Biographies

Clarence C. Gravlee is an associate professor in the Department of Anthropology at the University of Florida. Gravlee is a medical anthropologist whose research focuses on social inequities in health, with an emphasis on the health effects of racism in the African Diaspora. He can be reached at cgravlee@ufl.edu.

H. Russell Bernard is a professor emeritus of anthropology at the University of Florida. His areas of research include technology and social change, language death, and social network analysis. Bernard is the author of *Research Methods in Anthropology: Qualitative and Quantitative Approaches* (5th edition, AltaMira Press, 2011). He can be reached at ufruss@ufl.edu.

Chad R. Maxwell works in design anthropology as well as consumer research for marketing, media, and advertising. Chad's professional experience has primarily been with digital marketing and advertising consultancies where he works as an integrated measurement and people-centered research consultant. He can be reached at chad@chadrmaxwell.com.

Aryeh Jacobsohn is a design, user-experience, and business researcher at Razorfish, where he applies anthropology to defining and creating human–computer interactions. He can be reached at aryehj@gmail.com.