Journal of Gerontology: SOCIAL SCIENCES
2008, Vol. 63B, No. 6, S385–S393

Using Freelisting to Identify, Assess, and Characterize Age Differences in Shared Cultural Domains

Robert W. Schrauf[1] and Julia Sanchez[2]

[1]Department of Applied Linguistics, Pennsylvania State University, University Park.
[2]The Chicago School of Professional Psychology, Illinois.

Objectives. Freelisting is a brief, paper-and-pencil technique in which participants make lists of items that they believe belong in a particular domain. Where cultural domains are shared, as for young and old in the same society, subtle intracultural differences may be difficult to detect. This article presents a series of techniques for revealing and describing this intracultural variation in freelisted data among young versus old age groups.

Methods. Older ($N = 30$) and younger ($N = 31$) Mexicans in Mexico City made freelists in four quotidian domains: animals, emotions, illnesses, and gendered occupations.

Results. We used minimum residual factor analysis (consensus analysis) to establish domain *coherence* and assess overall *consensus* concerning contents of the domains. We established *subvariation* within the overall consensus by comparing levels of observed versus predicted inter-informant agreement. Results showed divergent patterns of inter-informant agreement between young and old participants across domains. Qualitative examination of items with higher salience for young versus old revealed age differences consistent with prior findings in each domain.

Discussion. The concatenation of these techniques renders freelisting an accessible, easily administered tool for probing age and group differences in cultural domains.

***Key Words*:** Free listing—Cultural variation—Ethnography—Mexican—Anthropology and aging.

FREELISTING is a simple paper-and-pencil technique, requiring little more than 5 min, that makes it possible to identify, measure, and describe variation in cultural knowledge between groups. It is a standard anthropological method for establishing the coherence or boundedness of particular cultural domains, for identifying what items belong in those domains, and for discerning which items are more salient or representative of the domain (Borgatti, 1999; Weller, 1998; Weller & Romney, 1988). The technique has the distinct advantage of providing *emic* data (from the point of view of the "cultural insider") and generally requires small samples of about 20 to 30 individuals (Schrauf and Sanchez *in press*; Weller & Romney, 1988).

A *cultural domain* is a symbolic category described by a cover term (e.g., "emotion words," "barriers to managing chronic illness," "things older people do"), a semantic relation (X is a kind of, X is a result of, X is found in, X is a step in the process of), and items included within the category (Spradley, 1979). One can approach cultural domains variously from semantic (lexical–referential), cognitive (mental models, scripts, prototypes), or discourse (topical) perspectives (Schrauf, 2002). Approaching culture in terms of domains of knowledge and practice has a long history in cognitive anthropology, and freelisting is one of the key methods used to investigate them (D'Andrade, 1995).

In gerontology, researchers have used freelisting to investigate a variety of cultural domains. Barg and colleagues (2006) investigated the cultural domain of *depressive symptoms* and asked older (M age = 75.8, $SD = 6.4$) African American and White participants ($n = 60$) to list words that described "a person who is depressed" and "you when you are depressed, down in the dumps, or blue" (p. S331). Both Whites and African Americans mapped to a common understanding of depression, with both groups viewing "loneliness" as a key component of the experience. Bayliss, Steiner, Fernald, Crane, and Main (2003) investigated the domain of *barriers to self-care* among adults with multiple comorbid chronic illnesses ranging in age from 31 to older than 70 ($N = 16$, with 12 older than age 51) and asked participants the following: "List everything you can think of that affects your ability to manage your medical conditions" (p. 17). Though no formal approach addressed domain coherence, the study identified 15 common barriers to self-management of chronic conditions. Harman (2001) looked at perceptions of *activities of the elderly* among the Maya (Guatemala) and asked participants ($N = 31$, age range = 17–74), "What are the activities of the elder Maya?" Subsequent analyses of the 18 most frequent items (out of 276) showed three thematic groupings: productive work, socioreligious obligations, and frivolous/self-indulgent activities.

These studies focus on topics of particular relevance to older samples (geriatric depression, managing chronic illnesses, activities of the elderly) but do not explicitly compare older and younger adults on the same topic. This latter is a new and potentially fruitful application of the method that focuses attention on subtle age or generational differences in shared domains. In a first study of this kind, Schrauf and Sanchez (2004) made an explicit comparison between young and old samples of Mexicans in Mexico and Anglos in the United States on emotion words. They asked 30 participants in each of the four age-by-culture groups "to list as many emotions as you can think of" and to rate each emotion label as positive, neutral, or negative (p. 273). Results showed that older groups in both cultural

groups generated a greater diversity of emotion labels than young groups and that all individuals produced more negative (50%) than positive (20%) and neutral (30%) emotion terms.

In anthropology, differences between younger and older samples in knowledge and/or structuring of cultural domains may be conceived as a form of *intracultural variation* (Dressler, Borges, Balieiero, & Dos Santos, 2005; Romney, Boyd, Moore, Batchelder, & Brazill, 1996; Romney & Moore, 1998; Romney, Weller, & Batchelder, 1986; Weller, 1987; Weller & Baer, 2002). In the main, anthropologists are interested in measuring intracultural variation as the extent to which individual members of a culture know about (or are competent in) the cultural models and schemata of the cultures to which they belong (D'Andrade, 1987; Romney et al., 1996; Weller, 1987). These individual differences can give rise to subgroupings within cultures (Boster, 1986; Dressler et al., 2005; Ross, 2004). However, within this tradition, freelisting has generally been used as a preparatory step to identify items commonly thought to belong to a domain, which would then be incorporated into other data collection techniques, including paired comparisons, triad sorts, and multiple choice items (Weller & Romney, 1988). Here we propose the refinement of freelisting itself as a means for exploring intracultural variation between age groups within the same society. Thus, given a shared domain (e.g., emotions), we propose a quick, "user-friendly" means of characterizing age differences in emphases, valuation, and perception of items in the domain.

To display the analytic techniques involved, we present freelisted data that we collected from older and younger Mexicans in Mexico City on four domains: animals, emotions, illnesses, and "women's work." We selected these domains because both young people and old people share extensive knowledge about them, and age-driven differences in emphasis, valuation, familiarity, and knowledge should be rather subtle. In addition, these domains have been extensively studied in the anthropological and psychological literature, and we know a good deal about their content and structure. This makes it possible to test the results from our freelisting analyses against expectations generated by the literature. The following paragraphs summarize this past literature.

Animals

First, long lists are common in this domain. Henley (1969) found that American undergraduates produced lists of animals ranging between 21 and 110 items. Second, young people produce longer lists than old people. Among Mexican participants, Ostrosky-Solis, Gutierrez, Flores, and Ardila (2007) found that younger participants generated more items than older participants in the animals category of the semantic fluency test—a neuropsychological version of freelisting. Finally, the domain is usually structured along two dimensions of contrast: (a) wild versus domestic and (b) big versus small (Chan, Butters, Salmon, & McGuire, 1993; Henley, 1969; Rubin & Olson, 1980).

Emotions

As noted above, in a previous analysis of the emotion data presented in this article, Schrauf and Sanchez (2004) found that both young and old participants' lists consisted of 50% negative, 30% positive, and 20% neutral emotion words. Nevertheless, studies of emotional development over the life span have shown that maturation brings with it an increased ability to coordinate positive and negative emotions (affect complexity) and an increased emphasis on expressing positive versus negative emotions (affect optimization; Labouvie-Vief & Medler, 2002; see also Schrauf & Hoffman, 2007). Thus, we might expect younger people to find contradictory emotions to be more salient than older people, and older people to find positive versus negative emotions to be more salient than younger people.

Illnesses

In research on the domain of illnesses among Mexicans, age has emerged as *the* most salient dimension of contrast (i.e., diseases affecting children and youth vs diseases affecting adults). Other dimensions of contrast include life-threatening versus non-life-threatening, hot versus cold, and contagious versus non-contagious (Weller, 1984a). Thus, we would predict that old people would emphasize different ("age-likely") illness conditions than would younger people.

Women's Work

To our knowledge, no empirical studies of occupational gender typing are available for the Mexican context. However, successive studies done with college students in the United States have shown that from 1975 (Shinar, 1975) to 1993 (Beggs & Doolittle, 1993), students' occupation stereotyping decreased toward "gender neutral." A replication of these studies with college students in France and Spain showed less stereotyping in those countries than in the United States (Muñoz-Sastre, Fouquereau, Igier, Salvatore, & Mullet, 2000). Thus, we might predict that the younger cohort in this study will show less occupational gender typing than the older cohort.

For each domain we pose the three questions that characterize our analytic approach:

1. Is there a bounded, unitary cultural domain (domain coherence)?
2. Are there significant subvariations or submodels within the overall consensus (subvariations within domains)?
3. What characterizes these submodels or subvariations (qualitative differences)?

Given our focus on intracultural variation, we predict two possible patterns of age group differences. A strong form of the prediction is that both young and old will show a higher within-group agreement than between-group agreement, reflecting the fact that both groups hold different subvariations within the larger consensus. A weaker version would be that one group will show higher within-group agreement versus the other group, suggesting that the latter group shares the consensus model to a larger degree. Either version is consistent with age differences in the domains. Finally, we expect that our methods of characterizing the differences from the freelisting results will corroborate previous work on these domains in the social science literature.

Methods

Study Sample

We recruited participants in this study as part of a larger project in the anthropology of aging that investigated shifts in

Table 1. Participant Characteristics and Scores on the Spanish Woodcock–Muñoz Language Survey

Characteristic	Young ($N = 31$)		Old ($N = 30$)		t (df)		$p<$
Age	22.55	(2.08)	64.83	(6.19)	35.99	(59)	.001
Gender (M/F)	20/10		18/13				
Education	14.16	(1.85)	15.07	(3.88)	1.17	(59)	*ns*
MMSE	29.81	(0.40)	28.53	(1.85)	3.74	(59)	.001
Woodcock: Picture Vocabulary[a]	526.43	(7.91)	528.43	(9.67)	0.88	(58)	*ns*
Woodcock: Analogies[a]	517.00	(10.41)	516.27	(10.40)	0.26	(54)	*ns*
Woodcock: Letter–Word Identification[a]	555.43	(9.12)	560.30	(8.09)	2.19	(58)	.05
Woodcock: Dictation[a]	518.96	(10.21)	514.63	(9.70)	1.61	(53)	*ns*

Notes: Data are M (SD) unless otherwise specified. M = male; F = female; MMSE = Mini-Mental State Examination.

[a]Scores are given in W scores, a transformation of the Rasch logits ability scale used to rescale raw scores in the Woodcock–Muñoz Language Survey (Woodcock & Muñoz-Sandoval, 1993b). A W score of 500 reflects the average performance of beginning fifth-grade students.

the semantic structure of common cultural domains due to immigration. This larger study compares the sending community (monolingual, Spanish-speaking, young and old Mexicans in Mexico City) with age-matched first- and second-generation immigrants (bilingual, Spanish/English speaking, young and old Mexican Americans in Chicago) and the receiving community (monolingual, English-speaking, young and old Anglos in Chicago). The data presented here came from the first group only. This sample included 61 Mexicans in Mexico City, comprising (a) a younger group ($N = 31$; 13 women, 18 men) with a mean age of 22.55 ($SD = 2.08$; range = 18–26) and (b) an older group ($N = 30$; 10 women, 20 men) with a mean age of 64.83 ($SD = 6.19$; range = 59–86). Occupations of older participants were as follows: teachers from elementary school to college ($n = 21$), business owners or employees ($n = 8$), and a government employee, a veterinarian, and a nurse. The younger group included students ($n = 17$), sales personnel ($n = 3$), office workers ($n = 6$), business owners or employees ($n = 3$), and a government employee and a reporter. Table 1 shows the demographic data for the samples. Groups were equated on years of education, cognitive health, and linguistic ability.

We assessed the cognitive status of all participants with the Spanish version of the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975; Karno, Burnam, Escobar, Hough, & Eaton, 1983), and all participants scored above the usual cutoff (24) for possible cognitive impairment (Ostrosky-Solis, Lopez-Arango, & Ardila, 2000). Note that small variance in young and old MMSE scores yielded a statistically significant but substantively meaningless difference between the groups. We administered the Spanish Woodcock–Muñoz Language Survey (Woodcock & Muñoz-Sandoval, 1993b) to control for possible effects of verbal ability on the freelisting exercise. The Woodcock–Muñoz is a test of cognitive academic language proficiency, equated with an English standardization sample of 6,000 and further standardized on Spanish-speaking populations ($N = 2{,}000$) from six countries (including Mexico) and five U.S. states (Woodcock & Muñoz-Sandoval, 1993a). The test classifies speakers according to age and grade equivalents on four subtests that assess oral language and abilities in reading and writing. No significant differences emerged between the groups on the Picture Vocabulary subtest, the Analogies subtest, or the Dictation subtest. However, on the Letter-Word Identification subtest, older participants outperformed young participants. In general, older participants did not differ from younger participants in linguistic abilities.

Data Collection

A native Spanish-speaking Mexican American bilingual (Julia Sanchez) collected lists in paper-and-pencil format in Spanish. Participants were asked to list as many items as they could in each domain. Domains were queried in counterbalanced order, and time constraints were not imposed on respondents. When respondents indicated that they had finished a list, the investigator asked if any additional items came to mind for that category, and, if none did, then moved on to the next domain (Brewer, 2002). We coded and standardized responses (Weller & Romney, 1988). For example, we coded morphological derivatives (e.g., sad, sadness), synonyms (e.g., flu, influenza), or phrases with similar meaning with the same term.

Data Analysis

Analysis involved three steps addressing domain coherence and consensus, subvariation within domains, and qualitative interpretation of subvariation.

Domain coherence: Is there a bounded, unitary cultural domain with broad consensus about its contents?—We used factor analytic methods to determine whether a unitary, bounded domain existed in the data. We organized the data into a standard matrix with rows representing participants and columns representing items. Cells contained a 1 if a respondent mentioned the item and a 0 if he or she did not. Including zeros in cells in which individuals did *not* mention an item potentially inflates respondent-by-respondent agreement patterns. However, because freelisting is not an exhaustive test of the contents of the mind, but rather a sampling of salient items in response to a cue, it seems reasonable to assume that two individuals who do not mention an item agree that the item is not salient. We then transformed the respondent-by-item matrix into a respondent-by-respondent matrix with values in the cells representing correlations between each pair of respondents.

We then submitted each of the four respondent-by-item matrices (animals, emotions, illnesses, and women's work) separately to consensus analysis, which is a minimum residuals factor analysis of the data (Batchelder & Romney, 1988; Romney et al., 1986; Weller, 1987) using UCINET 6 software (Borgatti, Everett, & Freeman, 2002). In essence, the respondent-by-item matrix is transposed, participants are treated as variables, and items are treated as cases. Each individual's response pattern is weighted in relation to the most frequent response patterns in

the group, and the factors represent the proportion of agreement between individuals and the group. When a large, first eigenvalue emerges relative to the second (a 3:1 ratio is the standard), and when all loadings on the first factor are positive, then a single, bounded cultural domain is said to exist (the "one-culture assumption"). Loadings on the first factor represent each individual's agreement with the group consensus. These competency scores are measures of intracultural variation among respondents concerning their knowledge of the domain in question.

Subvariations within domains: Are there significant subvariations or submodels within the overall consensus?—Some investigators have assessed group differences by comparing mean competency scores (Barg et al., 2006; Bennet, Switzer, Aguirre, Evans, & Barg, 2006), but, as is evident below, these scores can be equivalent while masking real differences. That is, two groups can show substantial consensus on what belongs in a domain and yet hold different patterns of inter-informant agreement within that overall consensus. As Ross (2004) pointed out, it is possible that 70% of all informants agree on each item in a data set. However, if the remaining 30% are not randomly distributed but "constitute a subgroup of 30% who usually disagree with the majority" (p. 132), then subvariation may well exist within the overall consensus. Residuals analysis addresses this issue by comparing the observed inter-informant agreement with the predicted agreement generated by the model (Ross, 2004; Ross & Medin, 2005). First, a correlation matrix of observed agreement is constructed by correlating each informant's actual pattern of responses with those of every other informant. Second, a correlation matrix of predicted agreement is constructed by correlating each individual's loading on the first factor of the minimum residuals factor analysis with every other individual's loading on the first factor. Third, a table of residual agreement is computed by subtracting the predicted from the observed agreement. Finally, patterns of residual agreement are compared across groups and tested via analyses of variance on group means. Group differences are indicated when either or both groups show significant variation from the general consensus (Ross, 2004, pp. 145–152).

A second method of probing for group differences is to investigate the distribution of items, or degree of sharing of items both between groups and within groups. At a between-group level, items produced by each group may be divided into those shared and those not shared with the other group. With two groups, this gives three sets of items for each domain. For instance, in comparisons of young and old groups there are items listed by both old and young, items mentioned by old but not young, and items mentioned by young but not old. Comparing numbers of unshared items across groups can give an informal measure of "knowledge" or familiarity with a domain. That is, if one group has a higher number of items unshared with the other group, that group may be more "knowledgeable" about the domain or more experienced with it.

At the within-group level, we investigated the degree of sharing by attending to within-group frequencies of items. Frequency is a simple but potent indicator of item salience. Obviously an item mentioned by 90% of the respondents is more salient at the level of the group than an item mentioned by 10%. In this analysis we assigned to each item in a participant's list the within-group frequency score of that item and then computed an average frequency for each participant. For instance, let us assume that participants make lists of common illnesses, and the word *cancer* appears in 19 participants' lists, *AIDS* appears in 13 lists, *smallpox* appears in 8 lists, and *diarrhea* appears in 7 lists. If one participant's list contained only these four items, that list would have a mean frequency score of 11.75. Lists with higher mean frequency scores are lists containing more items mentioned by others in the group. Thus, higher mean frequencies are indicators of higher within-group sharing. Independent samples *t* tests on frequency scores are then conducted between groups.

Qualitative differences: What characterizes these submodels or subvariations?—In addition to frequency across lists, rank within lists is also an indicator of how important or prominent an item is. Arguably, the items that people think of first would seem to have more salience than items listed later. However, rank is conditioned by list length, and therefore a more formal measure of salience, called *Smith's* S (Smith, 1993; Smith & Borgatti, 1998), takes into account both frequency and rank. The formula (corrected by Sutrop, 2001, p. 269) is

$$S = \left\{\sum[(L_i - R_j + 1)/L_i]\right\}/N$$

where S is the salience of an individual item, L_i is the length of an individual list, and R_j is the rank of the item in that list. Scores range from 1.0 (maximal salience: first item on every list) to 0.0 (where zero is "the extreme case being a phantom item that was never mentioned in any list at all"; Smith, 1993, p. 1).

Qualitative differences in lists produced by young versus old groups are examined as follows. Taking into account the overall consensus, we ask which shared items show the maximal difference in salience in young versus old lists. For each domain, the salience scores of items produced by the older group are subtracted from salience scores of items produced by the younger group. This gives a list of difference scores ranging from positive values (items with higher salience for the young vs the old) to negative values (items with higher salience for the old vs the young). By selecting subsets of items at either end of this continuum, we derive lists of items that can be used to characterize the emphases of one group relative to the other. The number of items chosen for this extreme groups comparison should be kept to between 5 and 10 so that the mathematical differences in salience scores reflect meaningful differences between groups.

Results

List Length and Content

The mean list length in the four domains for young and old *combined* was animals ($M = 26.34$, $SD = 5.94$), emotions ($M = 9.10$, $SD = 3.74$), illnesses ($M = 15.82$, $SD = 5.62$), and women's work ($M = 13.20$, $SD = 6.12$). The total numbers of labels (after coding for similarity) was animals (182), emotions (100), illnesses (230), and women's work (173). Table 2 lists the 10 most frequently cited items in each domain.

Domain Coherence: Cultural Consensus Analysis

A first step involves testing whether freelists represent one coherent, bounded cultural domain, and one assesses this via

Table 2. Ten Most Frequently Mentioned Items in Each Domain (Young and Old Combined)

Rank	Domain			
	Animals	Emotions	Illnesses	Women's Work
1	dog (*perro*)	happiness (*alegría*)	cancer (*cáncer*)	nurse (*enfermera*)
2	cat (*gato*)	sadness (*tristeza*)	AIDS (*SIDA*)	secretary (*secretaria*)
3	lion (*león*)	anger (*coraje*)	flu (*gripe*)	doctor (*médica*)
4	elephant (*elefante*)	love (*amor*)	diabetes (*diabetes*)	teacher (*maestro*)
5	horse (*caballo*)	hate (*odio*)	measles (*sarampión*)	cook (*cocinera*)
6	bird (*pájaro*)	affection (*cariño*)	smallpox (*viruela*)	seamstress (*costurera*)
7	tiger (*tigre*)	depression (*depresión*)	hypertension (*hipertensión*)	housewife (*ama de casa*)
8	giraffe (*jirafa*)	nostalgia (*nostalgia*)	chicken pox (*varicela*)	accountant (*contadora*)
9	fish (*pez*)	fear (*miedo*)	cough (*tos*)	stylist (*estilista*)
10	cow (*vaca*)	resentment (*rancor*)	syphilis (*sífilis*)	architect (*arquitecta*)

a minimum residuals factor analysis on respondents. Again, a 3:1 loading on the first versus the second eigenvalue supports domain coherence. First and second eigenvalues, respectively, in each domain were as follows: animals (39.41, 1.17), emotions (45.95, 1.25), illnesses (47.63, 0.67), and women's work (46.08, 0.78). In each case, the ratio of the first eigenvalue to the second was well beyond the required 3:1. Furthermore, the mean competency scores were high for both young and old in each domain as well: animals ($M_{old} = .80$, $SD = .05$; $M_{young} = .81$, $SD = .04$), emotions ($M_{old} = .87$, $SD = .07$; $M_{young} = .87$, $SD = .06$), illnesses ($M_{old} = .87$, $SD = .05$; $M_{young} = .91$, $SD = .03$), and women's work ($M_{old} = .85$, $SD = .08$; $M_{young} = .90$, $SD = .04$).

Subvariation Within the Domain

Residuals analysis.—Given this expected general agreement between young and old, we turn now to whether there were subpatterns of inter-informant agreement that might have discriminated old from young. We made comparisons between the mean within-group residual agreement among the young and the mean between-group residual agreement of the young with the old (and vice versa) and tested them via one-way analyses of variance (Ross, 2004; Ross & Medin, 2005). Omnibus *F*s for each domain showed significant differences in each domain: animals, $F(2, 88) = 10.37$, $MSE = .001$, $p < .0003$; emotions, $F(2, 88) = 7.50$, $MSE = .0004$, $p < .001$; illnesses, $F(2, 86) = 10.03$, $MSE = .0002$, $p < .001$; and women's work, $F(2, 86) = 5.55$, $MSE = .0002$, $p < .01$. Table 3 shows the results of the planned comparisons. In two domains, both groups showed subvariation from the overall consensus. That is, within-group agreement of young and within-group agreement of old were significantly different from their between-group agreement. In animals and illnesses, the old agreed more among themselves than they did with the young (animals: $F = 20.00$, $p < .001$; illnesses: 19.97, $p < .001$), and the young agreed more with themselves than with the old (animals: $F = 8.38$, $p < .01$; illnesses: 6.39, $p < .05$). In two domains, only one group showed a pattern significantly different from the overall group. In emotions, the young differed significantly from the larger group ($F = 11.81$, $p < .001$), but the old did not ($F = 0.03$, *ns*). In women's work, the opposite pattern held. The old differed significantly from the group consensus ($F = 11.11$, $p < .01$), but the young did not ($F = 2.78$, *ns*).

Between-group distribution of items.—At the level of the group, the older participants produced more unique items in

Table 3. Mean Residual Correlations Within and Between Age Groups in Four Cultural Domains

Domain	Old ($N = 30$) (Within)		Young ($N = 31$) (Within)		Old–Young (Between)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Animals	$.036_a$	.006	$.034_b$	.006	$.029_{a,b}$	.005
Emotions	.029	.006	$.035_a$	.008	$.029_a$	.006
Illnesses	$.022_a$	.005	$.020_c$	.004	$.017_{a,c}$	.004
Women's work	$.026_b$	.006	.024	.005	$.021_b$	.005

Note: Means in a row sharing subscripts are significantly different. Subscripts also indicate significance levels: a ($p > .001$), b ($p > .01$), and c ($p > .05$).

Table 4. Total Items Produced by Younger ($N = 31$) Versus Older ($N = 30$) Age Groups in Each Domain and Mean List Lengths of Individual Participants

Domain	Age Group	Total Items	Proportion Not Shared With Other Age Group	List Length *M* (*SD*)		*t* (*df*)	*p*<
Animals	Young	128	.18	28.06	(5.82)	2.39 (59)	.05
	Old	155	.32	24.57	(5.62)		
Emotions	Young	63	.33	11.19	(4.96)	2.38 (59)	.05
	Old	79	.47	8.53	(3.66)		
Illnesses	Young	119	.29	14.42	(4.72)	2.05 (58)	.05
	Old	178	.53	17.31	(6.18)		
Women's work	Young	97	.27	11.19	(4.96)	2.50 (58)	.05
	Old	143	.50	15.34	(7.72)		

absolute terms than the younger participants in every domain, and, by implication, the older group produced a higher proportion of items not shared with the younger group (see Table 4). At the level of individuals, Table 4 shows significant differences in average list length in each domain, but neither age group dominated: Young produced more animal and emotion names, old produced more illness and women's work names.

Within-group distribution of items.—The young also showed more within-group sharing than the old for three domains: animals, emotions, and illnesses. For three of the four domains, frequency scores of old versus young were significantly different. For animals, young ($M = 14.85$, $SD = 1.55$) showed significantly higher frequencies than old ($M = 11.40$, $SD = 1.95$), $t(59) = 7.68$, $p < .001$; for emotions, young ($M = 12.66$, $SD = 2.43$) were significantly higher than old ($M = 10.03$, $SD = 3.61$), $t(58) = 3.34$, $p < .01$; and for illnesses, young ($M = 11.86$, $SD = 3.24$) were significantly higher than old ($M = 7.55$, $SD = 2.09$), $t(58) = 6.08$; $p < .001$. On women's work, mean frequencies of young ($M = 8.87$, $SD = 1.64$) were not significantly different from those of old ($M = 8.39$, $SD = 2.35$), $t(58) = 0.93$, *ns*.

These distributional analyses show the following. First, older participants produced a greater number and variety of labels than did the young in every domain. Second, the young showed higher levels of within-group sharing than did the old (except in women's work).

Table 5. Top Five Shared Items With Greatest Difference in Salience Between Young and Old

Higher Salience for Young Group ($N = 31$)		Higher Salience for Old Group ($N = 30$)	
Item	Difference in Salience	Item	Difference in Salience
	Animals		
Bird (*pajaro*)	.424	Hen (*gallina*)	.205
Bear (*oso*)	.263	Rooster (*gallo*)	.174
Fish (*pez*)	.209	Horse (*caballo*)	.155
Hippopotamous (*hipopotamo*)	.195	Rabbit (*conejo*)	.148
Frog (*rana*)	.185	Burro (*burro*)	.142
	Emotions		
Love (*amor*)	.333	Happiness (*alegría*)	.133
Hate (*odio*)	.306	Nostalgia (*nostalgia*)	.061
Passion (*passion*)	.117	Excitement (*emoción*)	.057
Affection (*cariño*)	.095	Desire (*deseo*)	.044
Resentment, anger (*rancor*)	.078	Apathy (*apatía*)	.029
	Illnesses		
Flu (*gripe*)	.482	Diabetes (*diabetes*)	.248
AIDS (*SIDA*)	.460	Hypertension (*hypertensión*)	.193
Gonorrhoea (*gonorrea*)	.268	Rheumatism (*reumatismo*)	.136
Cough (*tos*)	.249	Chickenpox (*varicela*)	.109
Hepatitis (*Hepatitis*)	.196	Arthritis (*arthritis*)	.106
	Women's work		
Stylist (*estilista*)	.206	Nurse (*enfermera*)	.267
Psychologist (*psicologa*)	.095	Doctor (*médica*)	.263
Designer (*diseñadora*)	.089	Professor (*profesora*)	.164
Secretary (*secretaria*)	.068	Laundress (*lavandera*)	.143
Cashier (*cajera*)	.065	Architect (*arquitecta*)	.132

Qualitative Differences: Saliencies

For each domain, Table 5 shows the first five items with higher salience for the young (left) and for the old (right).

Animals.—The differential salience list of the old group contained almost exclusively domesticated animals (hen, rooster, horse, burro), whereas the young list seemed to include largely undomesticated animals (bird, bear, fish, hippos, and frogs).

Emotions.—In the older group, positive emotions (happiness, nostalgia, excitement, and desire) were more salient than for the young group. Alternatively, in the young but not the old group, contradictory emotion words were more salient.

Illnesses.—Among older individuals, the names of chronic illnesses (diabetes, hypertension, rheumatism, and arthritis) had higher salience than they did for young people. Among young people, sexually transmitted diseases (AIDS and gonorrhea) and common ailments (flu and cold) were more salient than they were for older individuals.

Women's work.—Because data on occupational gender typing for Mexico are not available, we rely on the Muñoz-Sastre and colleagues (2000) data for Spain and France to interpret our results. In the Muñoz-Sastre study, college students rated occupations on a continuous scale as male (1.0), neutral (4.0), and female (7.0). According to the Muñoz-Sastre ratings (with values in parentheses representing France and Spain), our data show that older Mexicans listed a mix of occupations falling on the male (doctor 3.01, 3.14; professor

4.00, 3.97; architect 3.54, 3.21) and female (nurse 4.59, 4.75; laundress, rating not available) end of the Muñoz-Sastre continuum, whereas younger Mexicans listed a majority of female occupations (cashier 5.00, 4.32; psychologist 4.55, 4.35; secretary 5.72, 3.97; stylist, rating not available; laundress, rating not available).

Discussion

This study demonstrates the usefulness of a suite of methods for analyzing freelisted data to explore age differences in cultural domains. By focusing on very common, quotidian cultural domains and by testing for differences between younger and older adults who are members of the same cultural group, we have sought to highlight the sensitivity of these methods for detecting and describing subtle differences in both domain content and patterns of agreement between age groups. Across four domains we addressed group consensus, age group subvariation within that consensus, and qualitative interpretation of that subvariation. To wit, we asked the following questions.

First, do the freelisted items in each of the domains represent bounded coherent domains with broad consensus about the content? Consensus analysis of the four domains, with young and old data pooled, showed both thematic coherence and high consensus. This is consistent with our focus on a technique that addresses cultural domains that are clearly shared across generations within the same society and/or cultural group.

Second, do young versus old respondents represent significant subvariations within this larger consensus? Analysis of the inter-informant agreement residuals showed significant differences between young and old in each domain. This is the specific *intracultural variation* within the larger consensus that we set out to examine. Inspection of group production showed that the older group generated more items than the younger group in every domain. This in itself is an interesting effect because it demonstrates that, at least in domains shared across young and old, older adults will produce a greater variety of items than will young adults (Schrauf & Sanchez, 2004, in press).

Third, what are the qualitative differences between these subvariations? Within each domain, inspection of items with highest salience for old versus young, and young versus old, provided a means of describing group differences at the level of meaning. Again, given that these are common, quotidian domains, we expected that differences would be subtle, and this step in the analysis probed those differences. Furthermore, for methodological reasons we chose domains that have been examined via other methods to verify that our results from freelisting would be consistent with prior findings.

In the animals domain, domestic animals were more salient for the old, whereas wild animals were more salient for the young. As indicated above, previous research has shown that domesticity is a common dimension of contrast in the domain of animals (Henley, 1969; Rubin & Olson, 1980), but it is not clear from these data alone why older versus younger Mexicans would show higher salience for domestic animals.

In emotions, the old participants emphasized positive emotions, whereas the young listed contradictory emotions. This pattern is consistent with the literature on affect optimization among older adults (Labouvie-Vief & Medler, 2002; Labouvie-Vief, Manfred, Elizabeth, & Zhang, 2007). Our own previous work on these emotion data showed that, in absolute terms, both young and old listed proportionally more negative than positive affect words (Schrauf & Sanchez, 2004), whereas the finding here suggests an interaction between these. That is, although both groups may list more negative versus positive labels, positive labels are more salient for the old versus the young. This is consistent with the literature finding a positivity bias in reports of emotion in later life (Chipperfield, Perry, & Weiner, 2003) and in memory for emotion (Kennedy, Mather, & Carstensen, 2004; Schrauf & Hoffman, 2007).

In illnesses, older individuals found chronic diseases to be more salient, whereas the young highlighted common ailments and sexually transmitted diseases. This accords with common experience as well as prior research. On the one hand, older individuals are more likely to experience chronic illness than young people. On the other hand, young people usually enjoy better health than older people and are often more sexually active. Hence, sexually transmitted diseases may be more salient for them. In research on the structure of the illness domain among Mexicans (Weller, 1984a) and among Americans (Weller, 1984b), the classification of illnesses by age emerged as the most salient dimension.

Finally, in the category of gendered work, the older participants seemed less prone to gender stereotyping than the young. This is an interesting effect, because we might have expected older individuals to retain more traditional occupational gender typing. One possible explanation for the difference is that two thirds of the older sample were teachers and may not have shared societal prejudices concerning occupation.

Several comments are in order regarding the nature and extension of these techniques to other domains. First, as noted in the introductory paragraphs, our purpose was to examine differences in age groups concerning cultural domains that are largely shared. It is interesting to note that the effect sizes (differences in group mean residuals) were low, though statistically significant (see Table 3). This implies very little difference between young and old. This is not surprising, given that both age groups have high exposure to and familiarity with these domains. Nevertheless, what qualitative differences we did find made sense in light of what we know or might expect about young and old participants.

Second, these techniques are not limited to freelists consisting of single words. Although in the research reported here participants were asked to give exemplars within categories (essentially a lexical-semantic exercise), in the studies employing freelisting noted at the beginning of this article participants were asked to make lists of more complex domains (e.g., symptoms of depression, topics of discussion with one's doctor, activities of elders). For that reason, throughout this article we have spoken of *items* in domains rather than *words*.

Third, there is the issue of appropriate sample size. The literature suggests that 30 participants are usually enough to serve as a representative sample for testing for domain coherence and consensus (Borgatti, 1999; Weller & Romney, 1988). In additional analyses of the data in this article, we confirmed that the top 10 to 15 items in aggregate begin to

become stable at around 15 participants and that 25 participants will provide 95% of the structure of 30 participants (Schrauf & Sanchez, in press).

Conclusion

In summary, the methods discussed in this article provide sophisticated ways of analyzing simple freelisted data for the purpose of detecting and describing age or generational differences in the perception and understanding of commonly shared cultural domains. In general, the article shows that the concatenation of analytic methods, moving from *consensus analysis* to address domain coherence and group consensus, *residuals and frequencies analysis* to address subvariation, and *salience analysis* to address qualitative differences, produces results consistent with what we might expect in each domain, both from prior literature and our own intuition. In short, freelisting provides a simple method of testing for differences in cultural domains and characterizing those differences.

Acknowledgments

This research was supported by National Institute on Aging Grant AG16340.

R. W. Schrauf planned the study, conducted the data analysis, and wrote the paper. J. Sanchez collected the data and assisted in data analysis.

Correspondence

Address correspondence to Robert W. Schrauf, PhD, Department of Applied Linguistics, Pennsylvania State University, 305 Sparks Building, University Park, PA 16802. E-mail: rws23@psu.edu

References

Barg, F. K., Huss-Ashmore, R., Wittink, M. N., Murray, G. F., Bogner, H. R., & Gallo, J. J. (2006). A mixed-methods approach to understanding loneliness and depression in older adults. *Journal of Gerontology: Social Sciences, 61B*, S329–S339.

Batchelder, W. H., & Romney, A. K. (1988). Test theory without an answer key. *Psychometrika, 53*, 71–92.

Bayliss, E. A., Steiner, J. F., Fernald, D. H., Crane, L. A., & Main, D. S. (2003). Descriptions of barriers to self-care by persons with comorbid chronic diseases. *Annals of Family Medicine, 1*(1), 15–21.

Beggs, J. M., & Doolittle, D. C. (1993). Perceptions now and then of occupational sex typing: A replication of Shinar's 1975 study. *Journal of Applied Social Psychology, 23*, 1435–1453.

Bennett, I., Switzer, J., Aguirre, A., Evans, K., & Barg, F. (2006). 'Breaking it down': Patient-clinician communication and prenatal care among African American women of low and higher literacy. *Annals of Family Medicine, 4*, 334–340.

Borgatti, S. P. (1999). Elicitation techniques for cultural domain analysis. In J. J. Schensul, M. D. LeCompte, B. K. Nastasi, & S. P. Borgatti (Eds.), *Enhanced ethnographic methods: Audiovisual techniques, focused group interviews, and elicitation techniques* (pp. 115–151). Walnut Creek, CA: Altamira Press.

Borgatti, S. P., Everett, M. G., & Freeman, L. C. (2002). *Ucinet for Windows: Software for social network analysis*. Harvard, MA: Analytic Technologies.

Boster, J. S. (1986). Exchange of varieties and information between aguarana manioc cultivators. *American Anthropologist, 88*, 428–436.

Brewer, D. D. (2002). Supplementary interviewing techniques to maximize output in free listing tasks. *Field Methods, 14*(1), 108–118.

Chan, A. S., Butters, N., Salmon, D. P., & McGuire, K. A. (1993). Dimensionality and clustering in the semantic network of patients with Alzheimer's disease. *Psychology and Aging, 8*, 411–419.

Chipperfield, J. G., Perry, R. P., & Weiner, B. (2003). Discrete emotions in later life. *Journal of Gerontology: Psychological Sciences, 58B*, P23–P34.

D'Andrade, R. (1987). Modal responses and cultural expertise. *American Behavioral Scientist, 31*(2), 194–202.

D'Andrade, R. (1995). *The development of cognitive anthropology*. Cambridge, England: Cambridge University Press.

Dressler, W. W., Borges, C. D., Balieiro, M. C., & Dos Santos, J. E. (2005). Measuring cultural consonance: Examples with special reference to measurement theory in anthropology. *Field Methods, 17*(4), 331–355.

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research, 12*, 189–198.

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

Henley, N. M. (1969). A psychological study of the semantics of animal terms. *Journal of Verbal Learning and Verbal Behavior, 8*, 1765–1184.

Karno, M., Burnam, A., Escobar, J. I., Hough, R. L., & Eaton, W. W. (1983). Development of a Spanish language version of the National Institute of Mental Health Diagnostic Interview Schedule. *Archives of General Psychiatry, 40*, 1183–1188.

Kennedy, Q., Mather, M., & Carstensen, L. L. (2004). The role of motivation in the age-related positivity effect in autobiographical memory. *Psychological Science, 15*, 208–214.

Labouvie-Vief, G., Manfred, D., Elizabeth, J., & Zhang, F. (2007). Six-year change in affect optimization and affect complexity across the adult life span: A further examination. *Psychology and Aging, 22*, 738–751.

Labouvie-Vief, G., & Medler, M. (2002). Affect optimization and affect-complexity: Modes and styles of regulation in adulthood. *Psychology and Aging, 17*, 571–588.

Muñoz-Sastre, M. T., Fouquereau, E., Igier, V., Salvatore, N., & Mullet, E. (2000). Perception of occupational gender typing: A replication on European samples of Shinar's (1975) and Beggs and Doolittle's (1993) studies. *Journal of Applied Social Psychology, 30*, 430–441.

Ostrosky-Solis, F., Gutierrez, A. L., Flores, M. R., & Ardila, A. (2007). Same or different? Semantic verbal fluency across Spanish-speakers from different countries. *Archives of Clinical Neuropsychology, 22*, 367–377.

Ostrosky-Solis, F., Lopez-Arango, G., & Ardila, A. (2000). Sensitivity and specificity of the Mini-Mental State Examination in a Spanish-speaking population. *Applied Neuropsychology, 7*(1), 25–31.

Romney, A. K., Boyd, J. P., Moore, C. C., Batchelder, W. H., & Brazill, T. (1996). Culture as shared cognitive representations. *Proceedings of the National Academy of Sciences, USA, 93*, 4699–4705.

Romney, A. K., & Moore, C. C. (1998). Toward a theory of culture as shared cognitive structures. *Ethos, 26*(3), 314–337.

Romney, A. K., Weller, S. C., & Batchelder, W. H. (1986). Culture as consensus: A theory of culture and informant accuracy. *American Anthropologist, 88*, 313–338.

Ross, N. O. (2004). *Culture and cognition: Implications for theory and method*. Thousand Oaks, CA: Sage.

Ross, N., & Medin, D. (2005). Ethnography and experiments: Cultural models and expertise effects elicited with experimental research techniques. *Field Methods, 17*(2), 131–149.

Rubin, D. C., & Olson, M. J. (1980). Recall of semantic domains. *Memory and Cognition, 8*, 354–366.

Schrauf, R. W. (2002). Comparing cultures within subjects: A cognitive account of acculturation as a framework for cross-cultural study. *Anthropological Theory, 2*(1), 101–118.

Schrauf, R. W., & Hoffman, L. (2007). The effects of revisionism on remembered emotion: The valence of older, voluntary immigrants' pre-migration autobiographical memories. *Applied Cognitive Psychology, 21*, 895–913.

Schrauf, R. W., & Sanchez, J. (2004). The preponderance of negative emotion words in the emotion lexicon: A cross-generational and cross-linguistic study. *Journal of Multilingual and Multicultural Development, 25*(2/3), 266–284.

Schrauf, R. W., & Sanchez, J. (in press). Age-effects and sample size in free listing. *Field Methods*.

Shinar, E. H. (1975). Sexual stereotypes of occupations. *Journal of Vocational Behavior, 7*, 99–111.
Smith, J. J. (1993). Using Anthropac 3.5 and a spreadsheet to compute a free-list salience index. *Cultural Anthropology Methods, 5*(3), 1–3.
Smith, J. J., & Borgatti, S. P. (1998). Salience counts—and so does accuracy: Correcting and updating a measure for free-list-item salience. *Journal of Linguistic Anthropology, 7*(2), 208–209.
Spradley, J. P. (1979). *The ethnographic interview*. New York: Holt, Rinehart & Winston.
Sutrop, U. (2001). List task and a cognitive salience index. *Field Methods, 13*(3), 263–276.
Weller, S. C. (1984a). Consistency and consensus among informants: Disease concepts in a rural Mexican village. *American Anthropologist, 86*, 966–975.
Weller, S. C. (1984b). Cross-cultural concepts of illness: Variation and validation. *American Anthropologist, 86*, 341–351.
Weller, S. (1987). Shared knowledge, intracultural variation, and knowledge aggregation. *American Behavioral Scientist, 31*(2), 178–193.
Weller, S. (1998). Structured interviewing and questionnaire construction. In R. Bernard (Ed.), *Handbook of methods in cultural anthropology* (pp. 365–409). Walnut Creek, CA: Altamira Press.
Weller, S., & Baer, R. D. (2002). Measuring within- and between-group agreement: Identifying the proportion of shared and unique beliefs across samples. *Field Methods, 14*(1), 6–25.
Weller, S. C., & Romney, A. K. (1988). *Systematic data collection*. Newbury Park, CA: Sage.
Woodcock, R. W., & Muñoz-Sandoval, A. F. (1993a). An IRT approach to cross-language equating and interpretation. *European Journal of Psychological Assessment, 9*, 233–241.
Woodcock, R. W., & Muñoz-Sandoval, A. F. (1993b). *Woodcock-Munoz Language Survey*. Itasca, IL: Riverside.

Received August 15, 2007
Accepted August 4, 2008
Decision Editor: Kenneth F. Ferraro, PhD