

Sorting it out: pile sorting as a mixed methodology for exploring barriers to cancer screening

Hung-Wen Yeh · Byron J. Gajewski · David G. Perdue ·
Angel Cully · Lance Cully · K. Allen Greiner · Won S. Choi ·
Christine Makosy Daley

© Springer Science+Business Media Dordrecht 2013

Abstract We discuss a mixed methodology for analyzing pile sorting data. We created a list of 14 barriers to colon cancer screening and recruited 18, 13, and 14 participants from three American Indian (AI) communities to perform pile sorting. Quantitative data were analyzed by cluster analysis and multidimensional scaling. Differences across sites were compared using permutation bootstrapping. Qualitative data collected during sorting were compiled by AI staff members who determined names for the clusters found in quantitative analysis. Results showed five clusters of barriers in each site although barriers in the clusters varied slightly across sites. Simulation demonstrated type I error rates around the nominal 0.05 level whereas power depended on the numbers of clusters, and between and within cluster variability.

H.-W. Yeh (✉) · B. J. Gajewski
Department of Biostatistics, The University of Kansas Medical Center,
Kansas City, KS 66160, USA
e-mail: hyeh@kumc.edu

H.-W. Yeh · B. J. Gajewski · A. Cully · L. Cully ·
K. A. Greiner · W. S. Choi · C. M. Daley
Center for American Indian Community Health, The University
of Kansas Medical Center, Kansas City, KS 66160, USA

A. Cully
e-mail: acully@kumc.edu

L. Cully
e-mail: lcully@kumc.edu

K. A. Greiner
e-mail: agreiner@kumc.edu

W. S. Choi
e-mail: wchoi@kumc.edu

C. M. Daley
e-mail: cdaley@kumc.edu

Keywords Pile sorting · Multivariable · Cluster analysis · Multidimensional scaling · Permutation bootstrapping

1 Introduction

Cancer screening is a key component in cancer prevention. Screening is often not well implemented in practice due to barriers in many populations, particularly in underserved communities. Hence, it is crucial to identify screening barriers so effective and efficient interventions can be designed to overcome them. Researchers normally use one of two approaches to ascertain barriers in a given population. They either ask community members in interviews or focus groups to name barriers, or they identify barriers on their own (from the literature or elsewhere) and ask survey participants to select those barriers that are relevant to them. Both of these approaches have significant limitations. While qualitative data from interviews or focus groups are not generalizable when analyzed through text analysis or other qualitative techniques, they do allow participants to identify barriers on their own; however, participants may not identify all barriers without prompting. Surveys are generalizable, but do not allow participants to self-identify barriers. These approaches can identify and possibly rank order barriers, but do not show how people group barriers together. It may be helpful to learn how barriers are grouped or how they form taxonomies because this may help in determining which barriers work together to create a multiplicative effect on participants. Screening adherence might be dramatically improved through simultaneous intervention on the multiple barriers working together to impede screening.

With roots in anthropology, *pile sorting*, also known as *card sorting*, is a powerful methodology that can be used to explore and contextualize relationships between individual and group norms, values, feelings and fears, and complex constructs such as barriers to cancer screening or risk behaviors (Trotter and Potter 1993). This method combines qualitative and quantitative techniques, through qualitative data collection using structured or semi-structured interviews and primarily quantitative analysis. First, attributes to be listed on each card are determined. Then, participants are instructed to group cards based on their perceptions of the barriers. In successive pile sorting, as was done in this case, cards (barriers in this study) with the most similarity are grouped first, followed by cards with less similarity. The process creates unconstrained clusters and continues until all cards or clusters are grouped into one pile.

B. J. Gajewski
School of Nursing, The University of Kansas Medical Center, Kansas City,
KS 66160, USA
e-mail: bgajewski@kumc.edu

D. G. Perdue
The American Indian Cancer Foundation, and Minnesota Gastroenterology P.A.,
Minneapolis, MN 55441, USA
e-mail: dperdue@aicaf.org

K. A. Greiner · C. M. Daley
Department of Family Medicine, The University of Kansas Medical Center,
Kansas City, KS 66160, USA

K. A. Greiner · W. S. Choi · C. M. Daley
Department of Preventive Medicine & Public Health, The University of Kansas Medical Center,
Kansas City, KS 66160, USA

(Alternatively, the successive pile sorting may be performed in the opposite direction, i.e., all cards in a beginning pile separated into individual piles with one single card). Researchers then compute the “distance” between each pair of cards and convert the results of pile sorting into a distance or proximity matrix. Each participant thus has her/his own distance matrix. The distance matrix is then analyzed by cluster analysis to explore the hierarchical cluster relationships, or by multidimensional scaling (MDS) to create a “map” indicating visual relationships among barriers. In this way, participants actively demonstrate their thinking and researchers systematically learn participants’ thoughts.

Because researchers typically have multiple participants perform a pile sorting activity, which produces multiple distance matrices and thus MDS maps, it is crucial to find a consensus matrix or map that best represents these participants for inference purposes. Several methods have been proposed for Euclidean and non-Euclidean distances. For Euclidean distance measures, consensus matrices may be obtained by generalized procrustes analysis (GPA, [Gower 1975](#); [Dijksterhuis and Gower 1991](#)) and STATIS ([Lavit 1988](#); [Qannari et al. 1995](#)) or its descendent DISTATIS ([Abdi et al. 2007](#)). In pile sorting, however, the distance measures are ranks and hence not Euclidean. The individual difference scaling (INDSCAL, [Carroll and Chang 1971](#)) and alternating least square (ALSCAL, [Takane et al. 1977](#)) are two options for understanding distances in non-Euclidean scales.

Methods of comparing matrices between groups include the Mantel’s test for associations of elements between matrices ([Mantel 1967](#); [Dietz 1983](#); [Schneider and Borlund 2007](#)) and the GPA ([Sibson 1978](#); [Gower and Dijksterhuis 2004](#); [Schneider and Borlund 2007](#)). Although these methods are often applied to assess the similarity or resemblance between matrices, the latter may be used to test for differences. The significance of both methods is determined by the permutation bootstrapping test due to the fact that there is no known probabilistic distribution for this test statistic.

In this paper, we apply mixed methods pile sorting to investigate barriers to colon cancer screening in three American Indian (AI) communities. Specifically, we explain how items for pile sorting were determined using semi-structured interviews and free listing, how the process of pile sorting was done including both a structured procedure and a series of open-ended questions, and how analysis was completed using both quantitative and qualitative techniques in a community-based participatory research framework. In addition, we articulate how to prepare pile sorting data for analysis because literature either lacks clarity in this preparation or relies on specialized software. Because participants were recruited from three communities, we were also interested in differences in perception of colon cancer screening barriers among community sites. We use a statistic extended from the conventional analysis of variance and evaluate its significance by the permutation bootstrapping test ([Efron and Tibshirani 1994](#)). Further, because little is known about the statistical properties of distance matrices obtained in pile sorting, the actual data analysis motivates a simulation study to examine Type I error rate and power.

2 Methods

In this section, we first describe our method for determining pile sort items. We then use a hypothetical example to clearly articulate the procedure of pile sorting, data collection, and conversion of pile sorting data to a distance matrix. We then describe the data collected to investigate perceptions about colon cancer screening among AI from three separate sites. As the objective of this manuscript is to present a method for exploring barriers, the site names

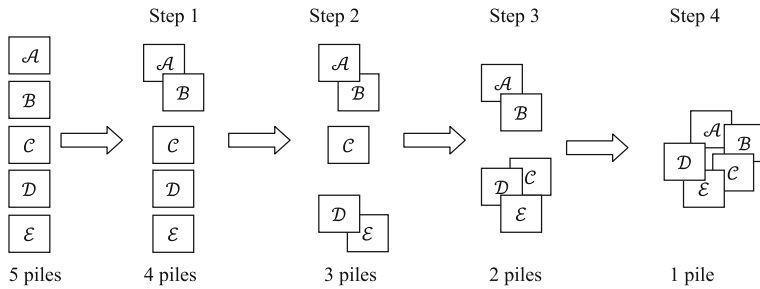


Fig. 1 Pile sorting procedure using a simple example of five cards

are kept confidential. Standard cluster analysis and MDS are summarized, followed by the methods for testing differences in perception among sites.

2.1 Free listing: identifying items for pile sorting

We interviewed 13 AI community leaders and 17 providers who work primarily with AI communities using semi-structured techniques that included a free list of barriers to colon cancer screening. Full details of methods and results of these interviews are published elsewhere (Daley et al. 2012). Free list items were analyzed using Smith's saliency index (Smith 1993), a weighted mean that takes into account the number of items on a participant's list, placement of the item on the participant's list, and the frequency with which an item was mentioned (Smith and Borgatti 1997). Items that were named by multiple participants, but were worded differently (e.g., money and finances) were grouped together prior to analysis. Community members on the research team used consensus to determine which items could be considered the same and grouped together. They also determined what word would be used for that grouping by consensus. Items chosen for the pile sort exercise included those with the highest saliency index. We included items until there was a natural break in the indices; the same point at which items began being named only by one individual.

To articulate the procedure of collecting and preparing pile sort data for analysis, we use a simple example of five items here. In successive pile sorting, as was done in this case, cards (barriers in this study) with the most similarity are grouped first, followed by cards with less similarity. The process creates unconstrained clusters and continues until all cards or clusters are grouped into one pile. (Alternatively, the successive pile sorting may be performed in the opposite direction, i.e., all cards in a beginning pile separated into individual piles ending with all cards in single card piles). Consider five cards (items in pile sorting are called cards in this article) labeled as A , B , C , D , and E , thus, we have five piles with one card in each pile at the beginning. Suppose a participant feels cards A and B are the most similar and places them into one pile in Step 1 (Fig. 1) and has four piles: $\{A, B\}$, $\{C\}$, $\{D\}$, and $\{E\}$ (We use a pair of brackets $\{\}$ to denote a pile, and elements inside the brackets to indicate the cards in the pile.) Among the four piles in Step 1, the participant regards cards D and E are the most similar and place them together in Step 2, so has three piles $\{A, B\}$, $\{C\}$, and $\{D, E\}$. The participant then thinks card C is more similar to the pile with cards D and E than the pile with cards A and B , and form two piles $\{A, B\}$ and $\{C, D, E\}$. In the last step, he/she put the two piles together and has $\{A, B, C, D, E\}$. The sequence of pile sorting is recorded in a spreadsheet as follows:

We then convert this spreadsheet to a $r \times c$ indicator matrix \mathbf{X} where the column dimension c is the number of cards and the row dimension $r = c - 1$ is the number of steps, which is

Step	Cards				
	\mathcal{A}	\mathcal{B}	\mathcal{C}	\mathcal{D}	\mathcal{E}
1	X	X			
2				X	X
3			X	X	X
4	X	X	X	X	X

one less than the number of cards. This matrix has an element $X_{ij} = 1$ if we see a mark “X” on i th row and j th column in the table shown above, and 0 otherwise. Here,

$$\mathbf{X} = \begin{matrix} & \text{Step} & \mathcal{A} & \mathcal{B} & \mathcal{C} & \mathcal{D} & \mathcal{E} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \end{matrix}$$

The indicator matrix is then used to calculate the distance matrix through a score vector $= (r, r - 1, r - 2, \dots, 1)'$ or $\mathbf{v} = (4, 3, 2, 1)'$ here. The distance matrix \mathbf{D} can be determined through a similarity matrix \mathbf{S} whose (i, j) element describes the similarity among barriers can be obtained by

$$S_{i,j} = \begin{cases} c & \text{if } i = j \\ \max(\mathbf{X}_i \# \mathbf{X}_j \# \mathbf{v}) & \text{if } i \neq j \end{cases}$$

where \mathbf{X}_i and \mathbf{X}_j denote the i th and j th columns, and “#” denote element-wise multiplication. Finally, the distance matrix \mathbf{D} is determined by subtracting each element in the similarity matrix from the number of cards, i.e., $\mathbf{D}_{ij} = c - \mathbf{S}_{ij}$, $i, j = 1, 2, \dots, 5$. In this 5-card example, the \mathbf{S} and \mathbf{D} matrices are hence

$$\mathbf{S} = \begin{matrix} & \mathcal{A} & \mathcal{B} & \mathcal{C} & \mathcal{D} & \mathcal{E} \\ \begin{matrix} \mathcal{A} \\ \mathcal{B} \\ \mathcal{C} \\ \mathcal{D} \\ \mathcal{E} \end{matrix} & \begin{bmatrix} 5 & 4 & 1 & 1 & 1 \\ 4 & 5 & 1 & 1 & 1 \\ 1 & 1 & 5 & 2 & 2 \\ 1 & 1 & 2 & 5 & 3 \\ 1 & 1 & 2 & 3 & 5 \end{bmatrix} \end{matrix}, \quad \text{and } \mathbf{D} = \begin{matrix} & \mathcal{A} & \mathcal{B} & \mathcal{C} & \mathcal{D} & \mathcal{E} \\ \begin{matrix} \mathcal{A} \\ \mathcal{B} \\ \mathcal{C} \\ \mathcal{D} \\ \mathcal{E} \end{matrix} & \begin{bmatrix} 0 & 1 & 4 & 4 & 4 \\ 1 & 0 & 4 & 4 & 4 \\ 4 & 4 & 0 & 3 & 3 \\ 4 & 4 & 3 & 0 & 2 \\ 4 & 4 & 3 & 2 & 0 \end{bmatrix} \end{matrix}$$

Because cards \mathcal{A} and \mathcal{B} are the most similar and put together in Step 1, their distance $D_{AB} = 1$; cards \mathcal{D} and \mathcal{E} are the second most similar and piled together in Step 2, so $D_{DE} = 2$; card \mathcal{C} joins the pile $\{\mathcal{D}, \mathcal{E}\}$ in Step 3, thus, $D_{CD} = D_{CE} = 3$. Piles $\{\mathcal{A}, \mathcal{B}\}$ and $\{\mathcal{C}, \mathcal{D}, \mathcal{E}\}$ are merged in Step 4 hence the distance between \mathcal{A} and \mathcal{C} , between \mathcal{A} and \mathcal{D} , between \mathcal{A} and \mathcal{E} , between \mathcal{B} and \mathcal{C} , between \mathcal{B} and \mathcal{D} , and between \mathcal{B} and \mathcal{E} is 4. The diagonal elements of matrix \mathbf{D} are all 0, reflecting each card is identical to itself (no distance). The distance matrix will be used in cluster and MDS analysis.

2.2 Data from study of barriers to colon cancer screening among AI

The barriers selected by our key informants and named by community members on our research team included access to care, cost of test, culture or tradition, do not trust Western

medicine, embarrassment, fear of results, fear of test, gender of the person doing the test, getting an appointment, knowledge or awareness, need for Native specific education, negative things about the test, no insurance, and transportation. A total of 45 participants performed pile sorting: 18 from Site 1, 13 from Site 2, and 14 from Site 3. We do not name the sites per agreements with tribes; all sites were in the Great Lakes region of the United States. Sites 1 and 3 were reservation communities; Site 2 was an urban community.

Items were then individually placed on 3-in. \times 5-in. laminated cards. Participants were instructed by community members trained as research assistants to group cards based on their perceptions of the barriers. They were instructed to first pile the two most similar cards, and repeat the process until all cards were placed into one pile as we illustrated in the example above. Community member researchers recorded the sequence of cards piled together, which was then used to produce distance matrices for analysis as described previously. In addition, participants were asked to name each pile as they were grouped based on why the cards were grouped together. These data were used to qualitatively understand what the piles meant to participants and provide a mixed methods approach to data interpretation. Because 14 cards were applied, the consequent distance and the similarity matrices would be 14 by 14 in dimension. The diagonal elements were zeros, and the off-diagonal elements ranged from 1 to 13. In the cases where participants thought any two cards were too different and never placed them together, the corresponding distance was given as 14.

2.3 MDS and cluster analysis

We first calculated the distance matrix for each participant. Because the measure of distances was virtually the rank of dissimilarities among cards that ranged from 1 to 14 (a distance of card to itself was 0 and not ranked), the magnitude of distances was not meaningful, thus it was natural to apply non-metric MDS. Kruskal's STRESS coefficient was used to examine the badness-of-fit. The results indicated excellent to good fit (inter-quartile range of STRESS 0.005–0.034) with 2D no-metric MDS for most individual matrices while far less satisfactory fit with metric MDS (IQR of STRESS 0.11–0.13). When matrices of all participants or of participants from the same sites were analyzed by the INDSCAL or ALSCAL, the fitness was poor (STRESS > 0.2) regardless of whether the scale was treated as metric or non-metric, unless the dimension was increased to 7 or more. Hence, we sought other approaches and obtained the averaged distance matrix for each site by the element-wise average of individual distance matrices. The averaged distance matrix was analyzed by MDS for each of the three sites and for the three sites combined. We also performed the agglomerative cluster analysis to evaluate the number of clusters in each MDS map using the Ward's minimum variance method (Johnson and Wichern 2001). The number of clusters was determined by the semi-partial R^2 (Timm 2002) as one uses the scree plot in principal component analysis to determine the number of principal components.

2.4 Hypothesis testing for comparing perceptions across groups

Comparison of participants' perceptions across groups was conducted by permutation bootstrapping. Consider $n = \sum_{i=1}^k n_i$ participants are recruited from k groups to perform pile sorting and the distance matrices of the k groups are to be compared. Let \mathbf{D}_{ij} denote the distance matrix from the j th participant in the i th group, ($j = 1, 2, \dots, n_i, i = 1, 2, \dots, k$), and \mathbf{D}_i and \mathbf{D} be the averaged distance matrices for participants in the i th group and for participants from all groups, respectively. Let $\mathbf{\Delta}_i$ denote the true mean distance matrix for group i . Assume the observed distance matrix for subject j in group i $\mathbf{D}_{ij} = \mathbf{\Delta}_i + \mathbf{E}_{ij}$ where

\mathbf{E}_{ij} is a residual matrix with mean $\mathbf{0}$ and variance Ψ that is homogeneous across subjects. To test the hypotheses $H_0: \Delta_1 = \Delta_2 = \dots = \Delta_k$ versus H_A : not all equal, we randomly assign participants to each of the groups to be compared, and then calculate the average distance matrix in each group. The test statistic is adopted from the concept of the between- versus within-group variance ratio in analysis of variance. We treat the individual distance matrix as “an observation value” and compute the sum of squares of off-diagonal elements in between- and within-difference matrices. Note diagonal elements are zeros and the sum of squares in matrix elements can be expressed as trace. In this way, the “variance ratio” becomes

$$\text{VR} = \frac{\sum_{i=1}^k n_i \times \text{trace}(\mathbf{D}_i - \mathbf{D}) / (k - 1)}{\sum_{i=1}^k \sum_{j=1}^{n_i} \text{trace}(\mathbf{D}_{ij} - \mathbf{D}_i) / (n - k)}$$

The distribution of the test statistics under the null hypothesis will be constructed in the following steps:

- (1) Randomly allocate all participants' distance matrices into the k groups.
- (2) Obtain the averaged distance matrix for each group and for all groups combined
- (3) Compute the test statistic VR
- (4) Repeat Steps (1) through (3) for B times.

As the F test in ANOVA, this test for homogeneity is one-sided with the rejection region in the right tail. The null hypothesis is rejected at the nominal 0.05 significance level.

2.5 Simulations

To examine the performance of this permutation bootstrapping test, we conducted a simulation study to examine its type I error rate and power. Without losing generality, we considered two groups of participants performing pile sorting using 14 cards, as in our AI example. The simulation study consisted of two parts. Part 1 was a simplistic scenario that participants' perception of the 14 cards formed only 2 clusters on a 2D map. We simulated card coordinates using two bivariate normal distributions, one for each cluster:

$$\mathbf{x}_1 \sim \text{BVN}\left(\begin{pmatrix} -1.5 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 & -0.62 \\ -0.62 & 1.2 \end{pmatrix}\right), \mathbf{x}_2 \sim \text{BVN}\left(\begin{pmatrix} 1.5 \\ 1 \end{pmatrix}, \begin{pmatrix} 0.8 & 0.42 \\ 0.42 & 2.4 \end{pmatrix}\right)$$

Figure 2a shows the ellipses that contain 95 % data generated from these parameters. The covariance (-0.62 and 0.42) suggests the direction (clockwise or counter-clockwise) and the angle of the clusters. In studying type I error rate, we simulated 7 cards from \mathbf{x}_1 and \mathbf{x}_2 for both groups. For investigating power, we kept one group unchanged whereas we simulated 6 and 8 cards from the two clusters in the other group. We considered three sample sizes: 5, 10, and 15 subjects per group. For each subject in each group, we then calculated the distance matrix. The upper (and/or lower)-triangular elements were then ranked based on the between-card distance to reflect the nature that the pile sorting data were actually ordinal. The distance matrices with ranked elements are then used to compute the variance ratio and to perform permutation bootstrapping test. We also considered that the within-cluster variation might also have impact on power, thus a scale parameter $c = 0.5, 1$, and 2 was multiplied to the cluster covariance matrices in separate simulations.

In Part 2 of the simulations, we considered a more realistic scenario where the 14 cards form 5 clusters. To examine the type I error rate, coordinates for 3, 2, 3, 4, and 2 cards were respectively generated by

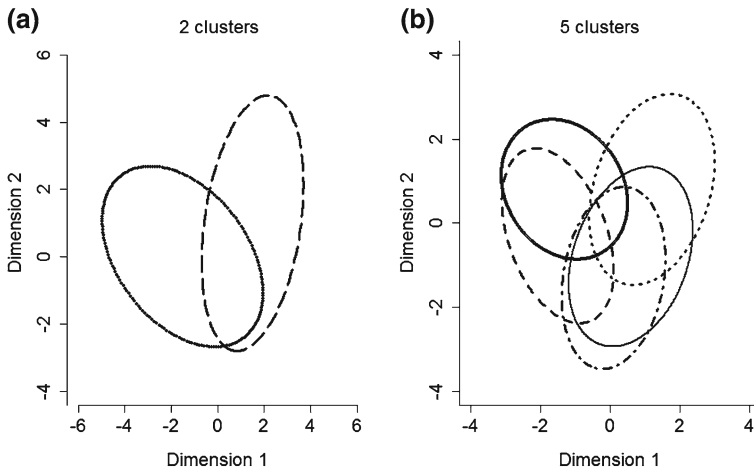


Fig. 2 The 95 % tolerance ellipses in simulation

$$\begin{aligned}
 \mathbf{x}_1 &\sim \text{BVN}\left(\begin{pmatrix} -1.3 \\ 0.8 \end{pmatrix}, \begin{pmatrix} 0.55 & -0.10 \\ -0.10 & 0.46 \end{pmatrix}\right), \mathbf{x}_2 \sim \text{BVN}\left(\begin{pmatrix} -1.5 \\ -0.3 \end{pmatrix}, \begin{pmatrix} 0.43 & -0.22 \\ -0.22 & 0.72 \end{pmatrix}\right) \\
 \mathbf{x}_3 &\sim \text{BVN}\left(\begin{pmatrix} 1.2 \\ 0.8 \end{pmatrix}, \begin{pmatrix} 0.54 & 0.20 \\ 0.20 & 0.86 \end{pmatrix}\right), \mathbf{x}_4 \sim \text{BVN}\left(\begin{pmatrix} 0.6 \\ -0.8 \end{pmatrix}, \begin{pmatrix} 0.53 & 0.19 \\ 0.19 & 0.76 \end{pmatrix}\right), \\
 \mathbf{x}_5 &\sim \text{BVN}\left(\begin{pmatrix} 0.1 \\ -1.3 \end{pmatrix}, \begin{pmatrix} 0.37 & 0.11 \\ 0.11 & 0.78 \end{pmatrix}\right)
 \end{aligned}$$

for both groups. The ellipses containing 95 % of the simulated data in these clusters are shown in Fig. 2b. To study power, we simulated 3, 2, 3, 4, and 2 cards from $\mathbf{x}_1, \mathbf{x}_2, \dots$, and \mathbf{x}_5 in one group, and switch one card from Cluster 5 to Cluster 3 for the other group, leaving 3, 2, 4, 4, and 1 cards in the 5 clusters. As the simplistic scenario of 2 clusters, we also did simulation for the three values of the scale parameter. In all scenarios, we permuted data 1,000 times, and replicated 1,000 simulation runs.

Quantitative analysis and simulation were performed on SAS version 9.2. The MDS and the cluster analyses were conducted using the procedures MDS and CLUSTER (Johnson 1998), respectively, and the data preparations as well as the rest of the analyses were conducted using the IML codes developed by the authors.

2.6 Analysis of qualitative data collected during pile sorting

Participants named each pile as it was created based on how they understood the barriers to be similar. Names of piles were compiled into a spreadsheet, based on the name given to the pile the first time each individual grouped a set of items together. Two community member researchers (one male and one female to ensure an emic or insider's perspective from each gender) went through the lists of pile names for each cluster and came to a consensus on the appropriate name. The two community member researchers (AC, LC) were two of the individuals who originally collected the data. This was done because not only could they look at the data from an emic perspective, but they could also remember additional qualitative comments that were made during collection. The use of community members as researchers

in both the data collection and data analysis phases of the research project helped us to maintain the community-based participatory nature of the research.

3 Results

3.1 The AI example

The respective averaged distance matrices for Sites 1, 2, and 3 are shown in Table 1. Figures 3, 4, 5, 6, 7, 8, and 9 show the results of cluster analysis and MDS, respectively. The dendrograms in Figs. 3, 4, and 5 demonstrate the hierarchical cluster relationships among the barriers to colon cancer screening at the three sites. Some barriers closely interconnect to each other, such as *cost of care* (barrier #2) and *no insurance* (barrier #13) among Site 1 participants, and *fear of results* (barrier #6) and *fear of test* (barrier #7) among participants from Sites 2 and 3. The horizontal axis of the dendrograms indicates the R^2 (defined as $R_k^2 = 1 - \sum_{i=1}^k SSE_i$ where SSE_i is the sum of squared errors of every item in cluster i from the cluster centroid) which implies the proportion of variation explained by the clusters. The plots suggest at least five clusters are required to achieve 60 % R^2 . Although the semi-partial R^2 in Fig. 6 does not show an obvious elbow, cluster numbers beyond five seem to have little contribution to improve fitness, suggesting five clusters may be a proper solution.

In data from Site 1, barriers *access to care* (#1), *getting an appointment* (#9), and *transportation* (#14) form a cluster that was named “Understanding the Process to Scheduling an Appointment”, barriers *cost of test* (#2) and *no insurance* (#13) form another named “Lack of Funds/Money/Resources”, barriers *embarrassment* (#5) and *gender of the person doing the test* (#8) form another named “Awkwardness”, barriers *fear of results* (#6), *fear of test* (#7), *knowledge or awareness* (#10), and *negative things about the test* (#12) form another named “Lack of Knowledge of Colonoscopy”. The final barriers, *culture or tradition* (#3), *do not trust Western medicine* (#4), and *need for Native specific education* (#11), form a cluster that was described by participants in a manner similar to the previous cluster and received the same name, “Lack of Knowledge of Colonoscopy”, by our community member researchers.

For Site 2 participants (the only urban site), the first two clusters consist of the same barriers as those in Site 1; however, only one of them received the same name based on how participants described them. The cluster containing the items *cost of test* (#2) and *no insurance* (#13) received the same name. The cluster containing the items *access to care* (#1), *getting an appointment* (#9), and *transportation* (#14) received the name “Lack of Knowledge of Colonoscopy”. The other clusters are slightly different from those of Site 1. The barriers *knowledge or awareness* (#10) joins barriers *culture or tradition* (#3), *do not trust Western medicine* (#4), and *need for Native specific education* (#11) to form one cluster that was also described as “Lack of Knowledge of Colonoscopy”. Barriers *embarrassment* (#5), *fear of results* (#6), *fear of test* (#7), and *gender of the person doing the test* (#8) form a cluster named “Issues of Trust”; the barrier *negative things about the test* (#12) itself represents a cluster.

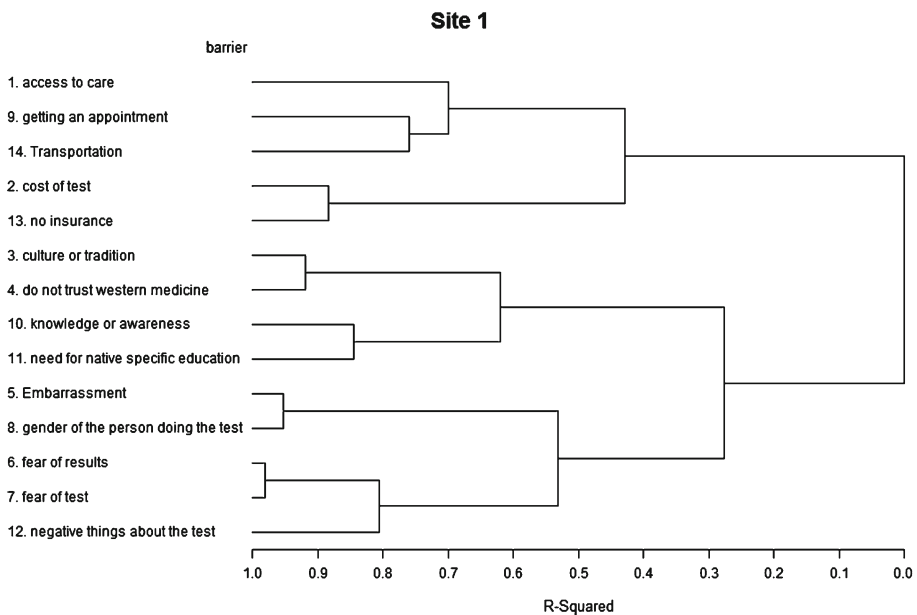
In data from Site 3, clusters yet again form differently. The two clusters seen in both Sites 1 and 2 now form one cluster that is described as “Lack of Funds/Money/Resources”. Barriers *culture or tradition* (#3), *do not trust Western medicine* (#4) form another cluster named “Trust Issues with Western Medicine”. Barriers *embarrassment* (#5) and *gender of the person doing the test* (#8) form another cluster as seen in Site 1, but in this case named “Awkwardness”. Barriers *fear of results* (#6), *fear of test* (#7), and *negative things about the test* (#12) form a cluster named “Fear”. Finally, barriers *knowledge or awareness* (#10) and *need for Native*

Table 1 The averaged distance matrix at each site

Card	Card													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
<i>(a) Site 1</i>														
1	0	10.5	12.5	13.2	12.4	10.6	11.8	12.4	9.1	12.3	12.5	12.4	10.3	7.8
2	10.5	0	12.5	12.4	11.2	11.7	11.1	12.1	11.4	12.8	12.6	12.2	4.8	9.3
3	12.5	12.5	0	8.5	11.4	11.4	11.2	11.8	12.4	10.8	7.8	11.1	12.5	13.2
4	13.2	12.4	8.5	0	11.9	11.9	11.5	12.0	13.3	10.9	10.2	11.7	12.1	12.6
5	12.4	11.2	11.4	11.9	0	9.9	8.9	7.6	12.1	11.5	12.0	10.5	11.5	12.7
6	10.6	11.7	11.4	11.9	9.9	0	6.3	9.3	11.8	9.7	11.3	8.3	11.6	12.1
7	11.8	11.1	11.2	11.5	8.9	6.3	0	10.5	11.7	9.6	11.1	6.8	11.1	12.0
8	12.4	12.1	11.8	12.0	7.6	9.3	10.5	0	12.1	11.1	12.3	11.1	12.2	12.6
9	9.1	11.4	12.4	13.3	12.1	11.8	11.7	12.1	0	11.2	11.7	12.2	11.3	9.2
10	12.3	12.8	10.8	10.9	11.5	9.7	9.6	11.1	11.2	0	8.9	9.4	12.8	12.9
11	12.5	12.6	7.8	10.2	12.0	11.3	11.1	12.3	11.7	8.9	0	10.7	12.6	13.2
12	12.4	12.2	11.1	11.7	10.5	8.3	6.8	11.1	12.2	9.4	10.7	0	12.3	13.1
13	10.3	4.8	12.5	12.1	11.5	11.6	11.1	12.2	11.3	12.8	12.6	12.3	0	9.9
14	7.8	9.3	13.2	12.6	12.7	12.1	12.0	12.6	9.2	12.9	13.2	13.1	9.9	0
<i>(b) Site 2</i>														
1	0	9.8	11.8	12.9	12.4	11.5	11.4	12.6	8.7	11.7	12.4	12.2	10.2	7.4
2	9.8	0	12.7	12.9	10.6	12.1	11.2	11.6	10.3	12.2	12.8	11.8	7.1	9.5
3	11.8	12.7	0	8.7	12.5	12.3	12.3	11.5	12.2	10.8	7.7	11.9	12.5	12.4
4	12.9	12.9	8.7	0	12.2	11.9	12.7	12.2	13.2	10.8	10.5	10.7	12.5	13.2
5	12.4	10.6	12.5	12.2	0	8.5	7.7	7.6	11.8	11.1	12.2	10.6	10.4	12.7
6	11.5	12.1	12.3	11.9	8.5	0	5.3	10.2	12.4	10.6	12.2	9.8	12.1	12.4
7	11.4	11.2	12.3	12.7	7.7	5.3	0	10.2	11.5	9.9	12.2	11.2	11.8	12.4
8	12.6	11.6	11.5	12.2	7.6	10.2	10.2	0	11.9	11.5	10.3	11.5	11.7	12.0
9	8.7	10.3	12.2	13.2	11.8	12.4	11.5	11.9	0	12.3	12.2	12.2	10.9	6.4
10	11.7	12.2	10.8	10.8	11.1	10.6	9.9	11.5	12.3	0	9.4	12.3	12.6	12.0
11	12.4	12.8	7.7	10.5	12.2	12.2	12.2	10.3	12.2	9.4	0	11.4	12.3	12.1
12	12.2	11.8	11.9	10.7	10.6	9.8	11.2	11.5	12.2	12.3	11.4	0	12.5	12.2
13	10.2	7.1	12.5	12.5	10.4	12.1	11.8	11.7	10.9	12.6	12.3	12.5	0	10.8
14	7.4	9.5	12.4	13.2	12.7	12.4	12.4	12.0	6.4	12.0	12.1	12.2	10.8	0
<i>(c) Site 3</i>														
1	0	9.2	11.9	12.4	11.6	10.7	11.1	12.4	9.9	10.2	10.3	11.2	9.2	9.0
2	9.2	0	11.7	11.9	11.9	11.5	12.1	12.2	10.1	11.7	11.9	11.6	6.4	8.5
3	11.9	11.7	0	10.0	11.1	11.7	12.2	11.0	12.7	10.6	10.1	11.7	11.7	12.4
4	12.4	11.9	10.0	0	12.0	11.1	10.9	11.1	10.8	11.7	10.6	10.6	11.9	11.1
5	11.6	11.9	11.1	12.0	0	10.6	10.0	8.9	11.9	10.7	11.9	10.4	12.1	10.9
6	10.7	11.5	11.7	11.1	10.6	0	5.3	12.6	11.5	8.9	10.0	9.6	11.2	11.6
7	11.1	12.1	12.2	10.9	10.0	5.3	0	12.1	11.2	8.6	9.9	8.7	12.1	11.6
8	12.4	12.2	11.0	11.1	8.9	12.6	12.1	0	11.7	10.1	11.3	11.2	12.2	12.3

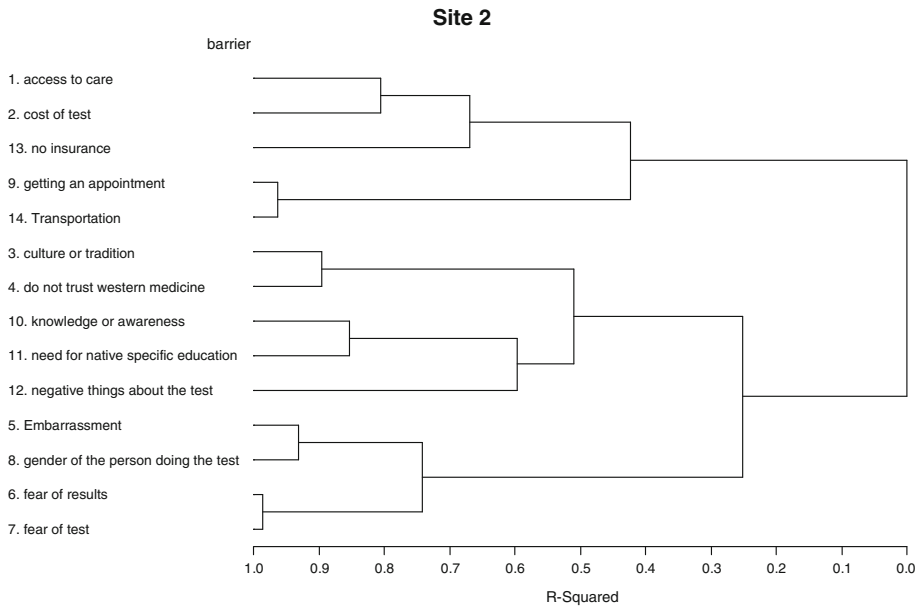
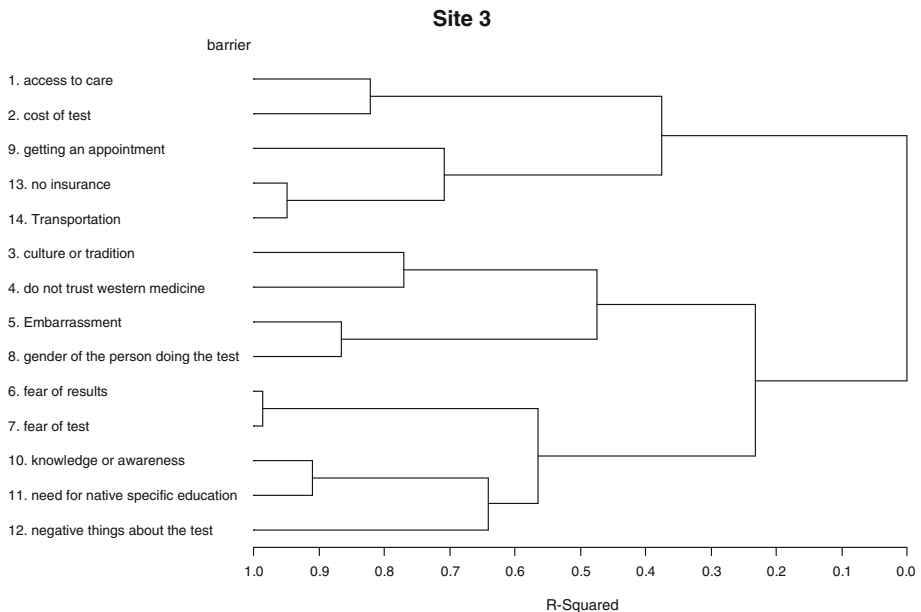
Table 1 continued

Card	Card													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
9	9.9	10.1	12.7	10.8	11.9	11.5	11.2	11.7	0	12.0	11.9	10.9	9.8	9.9
10	10.2	11.7	10.6	11.7	10.7	8.9	8.6	10.1	12.0	0	8.8	10.7	11.7	11.9
11	10.3	11.9	10.1	10.6	11.9	10.0	9.9	11.3	11.9	8.8	0	10.9	11.9	11.9
12	11.2	11.6	11.7	10.6	10.4	9.6	8.7	11.2	10.9	10.7	10.9	0	12.1	11.5
13	9.2	6.4	11.7	11.9	12.1	11.2	12.1	12.2	9.8	11.7	11.9	12.1	0	8.4
14	9.0	8.5	12.4	11.1	10.9	11.6	11.6	12.3	9.9	11.9	11.9	11.5	8.4	0

**Fig. 3** Dendrogram for Site 1

specific education (#11) cluster together and are named “Lack of Knowledge”. The cluster labels for the three sites are summarized in Table 2.

When performing MDS, we considered both non-metric and metric methods. The non-metric method generally did not fit the data well (Kruskal’s STRESS coefficients are 0.11, 0.16, and 0.19 for the three sites for 2D projection). The 1D metric MDS provides satisfactory badness-of-fit (the STRESS coefficients are 0.09 for all sites) though, we present the results of 2D MDS (Fig. 4, the STRESS coefficients are 0.06 for Site 1 and 0.07 for the other two sites) for better spatial presentation on relationships. Clearly, the barriers *access to care* (#1), *getting an appointment* (#9), *transportation* (#14), *cost of care* (#2), and *no insurance* (#13) are separate from other barriers in all sites, though their descriptions in each vary, with participants sometimes focusing on understanding of the process of colonoscopy or scheduling and other times focusing on lack of resources. In general, cluster results from participants from Site 1 are more similar to those from

**Fig. 4** Dendrogram for Site 2**Fig. 5** Dendrogram for Site 3

participants from Site 2 than either is similar to results from participants from Site 3, though Sites 1 and 3 are both reservation communities and Site 2 is an urban community. Specifically, in Site 3, the barriers *embarrassment* (#5) and *gender of the per-*

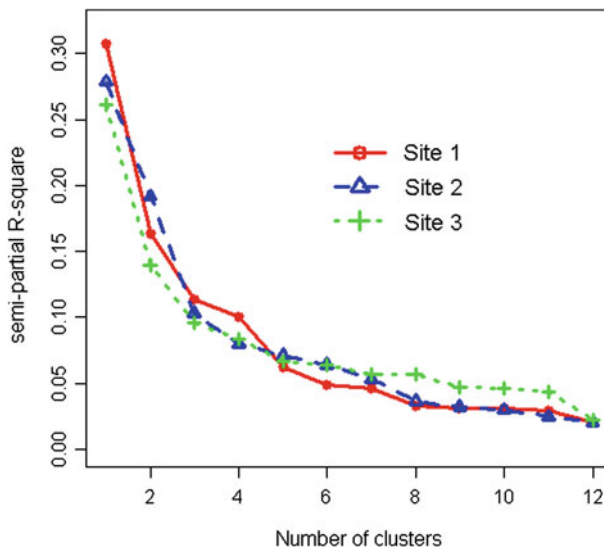


Fig. 6 Semi-partial R^2 versus cluster number

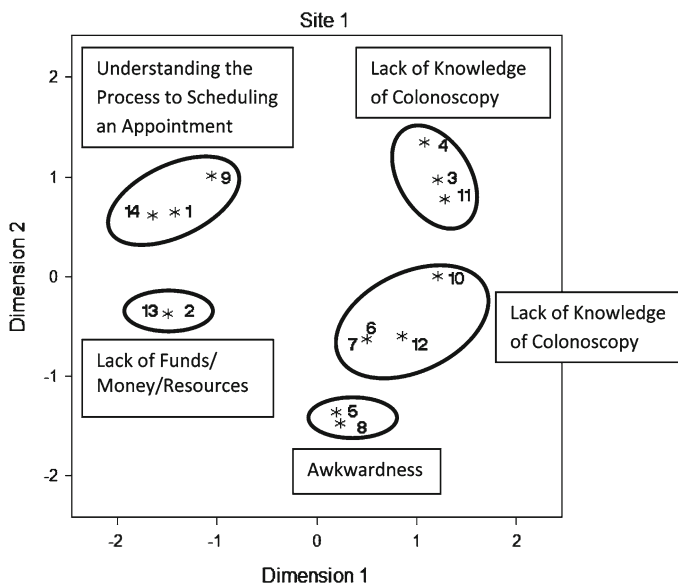


Fig. 7 2D MDS for Site 1. Card (barrier) labels: 1 access to care, 2 cost of test, 3 culture or tradition, 4 do not trust western medicine, 5 embarrassment, 6 fear of results, 7 fear of test, 8 gender of the person doing the test, 9 getting an appointment, 10 knowledge or awareness, 11 need for native specific education, 12 negative things about the test, 13 no insurance, and 14 transportation

son doing the test (#8) more similar to the barriers *culture or tradition* (#3) and *do not trust Western medicine* (#4) than the barriers *knowledge or awareness* (#10) and *need for Native specific education* (#11) are. Accordingly, the names of the clusters vary in Site 3. However, the differences do not achieve statistical significance according to the

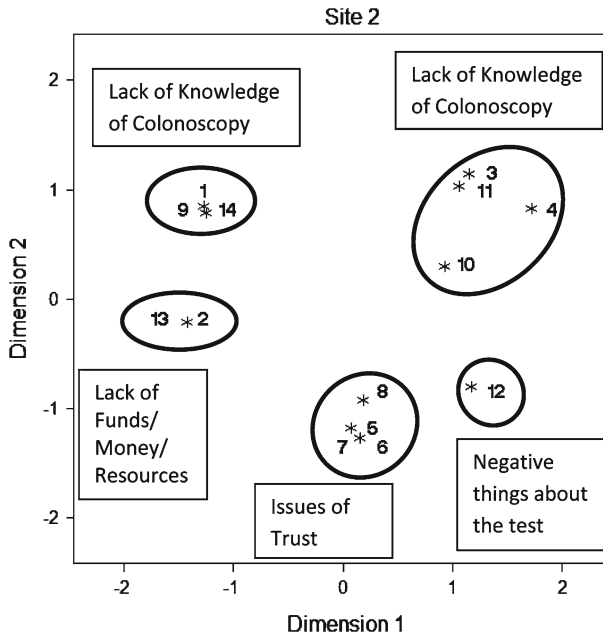


Fig. 8 2D MDS for Site 2. Card (barrier) labels: 1 access to care, 2 cost of test, 3 culture or tradition, 4 do not trust western medicine, 5 embarrassment, 6 fear of results, 7 fear of test, 8 gender of the person doing the test, 9 getting an appointment, 10 knowledge or awareness, 11 need for native specific education, 12 negative things about the test, 13 no insurance, and 14 transportation

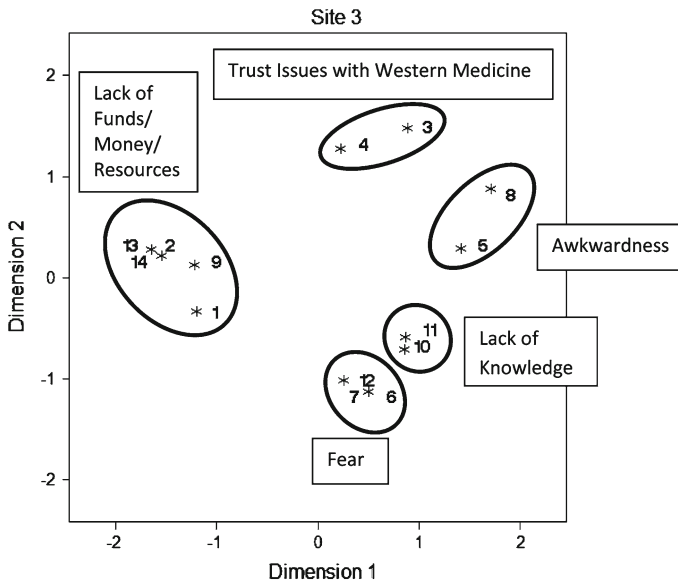


Fig. 9 2D MDS for Site 3. Card (barrier) labels: 1 access to care, 2 cost of test, 3 culture or tradition, 4 do not trust western medicine, 5 embarrassment, 6 fear of results, 7 fear of test, 8 gender of the person doing the test, 9 getting an appointment, 10 knowledge or awareness, 11 need for native specific education, 12 negative things about the test, 13 no insurance, and 14 transportation

permutation bootstrapping test using 1,000 samples (variance ratio = 1.12 across the three sites, $p = 0.219$).

3.2 Simulation results

Table 3 shows that using the variance ratio test statistic in permutation bootstrapping lead to type I error rates around the nominal 0.05 level. For the parameters specified in the distributions of \mathbf{x}_1 , \mathbf{x}_2 , ..., and \mathbf{x}_5 , and the numbers of cards specified above, the sample size of 15 participants per group is able to detect the difference with 59 % power. For greater within-cluster variation (larger values of the scale parameter), more participants will be required to achieve the same level of power.

4 Discussion

Pile sorting is a method that researchers use to understand complex relationships about a target community's perceptions as they may relate to health behaviors. In cancer screening research, pile sorting can be used to learn how participants perceive barriers, and how barriers relate to each other. Knowledge of how barriers cluster can guide researchers to design more effective and efficient interventions capable of tackling multiple barriers. In addition to this unique strength, there are several advantages over other qualitative and quantitative approaches. Compared to the pure qualitative methods of text analysis or similar techniques for analysis of interview or focus group data, pile sorting data, though collected during structured interviews, allow us to quantify the results into constructs that are more generalizable. In particular, the MDS maps provide direct visual distance among barriers and make presentation and communication an easier task. Survey research is often troubled by non-responses that create missing data and complicates analysis. On the other hand, pile sorting during structured interviewing is interactive and consequently, non-response is rarely an issue. Despite these advantages, pile sorting is not designed to scale barriers by importance.

In this work, we describe the procedure of pile sorting, and apply this technique to investigate perceptions about colon cancer screening barriers among AI participants. The proposed variance ratio statistic and permutation bootstrapping test provide researchers a tool to evaluate whether perceptions of barriers differ across populations. Evidence of differences in perceptions across populations may suggest the same intervention may not work equally well in a different population, indicating need for community- or culturally-tailored approaches. Simulation can be used to estimate the sample size in designing studies using pile sorting in multiple populations. Our simulation demonstrates and implies that several factors may influence the sample size: the number of clusters, the location or coordinates of clusters, the within-cluster variance (and covariance), and the number of items in the clusters. Although our simulation applies the same sample size per group, it needs not be equal in practice. For instance, if one population is small or the subjects are difficult to recruit, one may use an affordable sample size in this group and use larger sample sizes in other groups. Here we focus on comparing the perceptions across groups without using individual information. In order to include and control participants' characteristics, further investigation on regression models is necessary.

In our AI example, we see piles developing around particular themes that are then named, using word from community members. Across the sites there are differences in how items are grouped together, showing a difference in how people from each site perceive barriers to relate to each other. This shows us possibilities for intervention on multiple barriers targeting a par-

Table 2 Cluster labeling in the three sites

Site	Cluster of barriers	Cluster label
1	(#1) Access to care	Understanding the Process to Scheduling an Appointment
	(#9) Getting an appointment	
	(#14) Transportation	
	(#2) Cost of test	
	(#13) No insurance	Lack of Funds/Money/Resources
	(#5) Embarrassment	
	(#8) Gender of the person doing the test	Awkwardness
	(#6) Fear of results	
	(#7) Fear of test	Lack of Knowledge of Colonoscopy
	(#10) Knowledge or awareness	
	(#12) Negative things about the test	
	(#3) Culture or tradition	
	(#4) Do not trust Western medicine	Lack of Knowledge of Colonoscopy
	(#11) Need for Native specific education	
2	(#2) Cost of test	Lack of Funds/Money/Resources
	(#13) No insurance	
	(#1) Access to care	Lack of Knowledge of Colonoscopy
	(#9) Getting an appointment	
	(#14) Transportation	
	(#3) Culture or tradition	
	(#4) Do not trust Western medicine	Lack of Knowledge of Colonoscopy
	(#10) Knowledge or awareness	
	(#11) Need for Native specific education	
	(#5) Embarrassment	
	(#6) Fear of results	Issues of Trust
	(#7) Fear of test	
	(#8) Gender of the person doing the test	
	(#12) Negative things about the test	
3	(#1) Access to care	Lack of Funds/Money/Resources
	(#2) Cost of test	
	(#9) Getting an appointment	
	(#13) No insurance	
	(#14) Transportation	Trust Issues with Western Medicine
	(#3) Culture or tradition	
	(#4) Do not trust Western medicine	
	(#5) Embarrassment	
	(#8) Gender of the person doing the test	Awkwardness
	(#6) Fear of results	
	(#7) Fear of test	Fear
	(#12) Negative things about the test	
	(#10) Knowledge or awareness	Lack of Knowledge
	(#11) Need for Native specific education	

Table 3 Simulation of the permutation bootstrapping test

(a) Type I error rate

Scale factor ^a	2 clusters ^b			5 clusters ^c				
	No. subjects per group			No. subjects per group				
	5	10	15	5	10	15	20	30
0.5	0.062	0.049	0.050	0.049	0.056	0.052	0.051	0.052
1	0.049	0.045	0.056	0.050	0.058	0.044	0.048	0.051
2	0.053	0.046	0.051	0.056	0.049	0.047	0.045	0.047

(b) Power of detecting differences at 0.05 significance level

Scale factor ^a	2 clusters ^d			5 clusters ^e				
	No. subjects per group			No. subjects per group				
	5	10	15	5	10	15	20	30
0.5	0.958	1.000	1.000	0.451	0.868	0.982	1.000	1.000
1	0.505	0.909	0.991	0.185	0.346	0.592	0.774	0.964
2	0.164	0.391	0.574	0.087	0.118	0.187	0.266	0.416

^a A constant multiplied to the covariance matrix influencing the within cluster variability^b 7 cards per cluster^c No. cards in the five clusters: 3, 2, 3, 4, 2^d Group 1: 7 cards per cluster versus Group 2: 6 and 8 cards in the two clusters^e Group 1: 3, 2, 3, 4, 2 cards in Clusters 1–5 versus Group 2: 3, 2, 4, 4, 1 cards in the corresponding clusters

ticular community. When barriers are grouped similarly across sites, it shows us possibilities for intervention in multiple AI communities. We are currently analyzing data from the Midwest region to determine if there are similarities here as well. If we can determine that there are similar barriers across regions, it may be possible to develop a multiple-barrier intervention that can be used in multiple locations, possibly one large multi-site trial. With the small size of the Native population in the United States, it is often necessary to intervene in multiple locations in order to gain a large enough sample size to determine statistical significance. Pile sort data from multiple locations can help us to determine how best to intervene in these multi-site trials.

The qualitative data collected during the pile sort exercise allows us to better see why participants are grouping particular barriers together. By seeing how people name their combinations and if it is similar across sites with similar clusters, we can determine more specific interventions and accompanying educational materials that are appropriate across sites. We can also determine if accompanying educational materials in particular need to be modified across sites to better target a community and how individuals within that community conceptualize barriers. For example, in both Sites 1 and 2, participants grouped together barriers *access to care* (#1), *getting an appointment* (#9), and *transportation* (#14); however, Site 1 participants focused their naming of the cluster on understanding scheduling whereas participants in Site 2 focused on understanding colonoscopy itself. While this may not seem to be a large difference, it could change the type of education that needs to be done in site. It could also change an intervention, with the intervention for Site 1 focused specifically on modifying how appointments are scheduled and the intervention for Site 2 focusing on better explanation of the test itself and no intervention on the scheduling process, even though the

same barriers are being clustered together. Without the additional qualitative information, this difference would not be known. The combination of identifying clusters quantitatively and naming the clusters qualitatively gives us a fuller picture of how to develop an intervention that is appropriate to a community and has a better chance at success.

5 Conclusions

Pile sorting provides a unique value to understand community members' perceptions of colon cancer screening barriers. Results from cluster analysis and MDS may offer guidance in designing future interventions that more effectively overcome barriers and improve screening rates. Individuals' perceptions about screening barriers may be compared across populations using the proposed variance ratio and the permutation bootstrapping test. The addition of qualitative analysis of naming of piles adds further information to understand how participants create the clusters and can help in intervention development.

Acknowledgements This work was supported in part by the National Institute on Minority Health and Health Disparities Center of Excellence Grant, Center for the American Indian Community Health (CAICH) P20MD004805 and by the National Cancer Institute (R03121828). HY and BJG were also supported by the NIH Grant 1UL1RR033179. DP was supported by NIH Grant U01 CA114642. The contents are solely the responsibility of authors and do not necessarily represent the official view of the NIH.

Conflict of interest The authors declare that they have no conflict of interests.

References

- Abdi, H., Valentin, D., Chollet, S., Chrea, C.: Analyzing assessors and products in sorting tasks: DISTATIS, theory and applications. *Food Qual. Prefer.* **18**, 1–16 (2007)
- Carroll, J.D., Chang, J.J.: Analysis of individual differences in multidimensional scaling via an n -way generalization of "Eckert–Young" decomposition. *Psychometrika* **35**, 283–319 (1971)
- Daley, C.M., James, A.S., Filippi, M., Weir, M., Braiuca, S., Kaur, B., Choi, W.S., Greiner, K.A.: American Indian community leader and provider views of need and barriers to colorectal cancer screening. *J. Health Dispar. Res. Pract.* **5**(2), 10–23 (2012)
- Dietz, E.J.: Permutation tests for association between two distance matrices. *Syst. Zool.* **32**, 21–26 (1983)
- Dijksterhuis, G.B., Gower, J.C.: The interpretation of generalized procrustes analysis and allied methods. *Food Qual. Prefer.* **3**, 67–87 (1991)
- Efron, B., Tibshirani, R.J.: *An Introduction to the Bootstrap*. Chapman & Hall/CRC, New York (1994)
- Gower, J.C.: Generalized procrustes analysis. *Psychometrika* **40**, 33–51 (1975)
- Gower, J.C., Dijksterhuis, G.B.: *Procrustes Problems*. Oxford University Press, Oxford (2004)
- Johnson, D.E.: *Applied Multivariate Methods for Data Analysts*, 1st edn. Duxbury Press, Pacific Grove (1998)
- Johnson, R.A., Wichern, D.W.: *Applied Multivariate Statistical Analysis*, 5th edn, pp. 690–692. Prentice Hall, Upper Saddle River (2001)
- Lavit, C.: *Analyse conjointe de tableaux quantitatifs*. Masson, Paris (1988)
- Mantel, N.: Detection of disease clustering and generalized regression approach. *Cancer Res.* **27**, 209–220 (1967)
- Qannari, E.M., Wakeling, I., MacFie, H.J.H.: A hierarchy of models for analyzing sensory data. *Food Qual. Prefer.* **6**, 309–314 (1995)
- Schneider, J.W., Borlund, P.: Matrix comparison, part 2: measuring the resemblance between proximity measures or ordination results by use of the mantel and procrustes statistics. *J. Am. Soc. Inf. Sci. Technol.* **58**, 1596–1609 (2007)
- Sibson, R.: Studies in the robustness of multidimensional scaling: procrustes statistics. *J. R. Stat. Soc. B* **40**, 234–238 (1978)
- Smith, J.J.: Using ANTHROPAC 3.5 and a spreadsheet to compute a free-list salience index. *Cult. Anthropol. Methods* **5**, 1–3 (1993)

- Smith, J.J., Borgatti, S.P.: Saliency counts—and so does accuracy: correcting and updating a measure for free-list-item saliency. *J. Linguist. Anthr.* **7**(2), 208–209 (1997)
- Takane, Y., Young, F.W., de Leeuw, J.: Nonmetric individual differences multidimensional scaling: an alternating least squares method with optimal scaling features. *Psychometrika* **42**, 8–67 (1977)
- Timm, N.H.: *Applied Multivariate Analysis*, pp. 522–533. Springer, New York (2002)
- Trotter, R.T., Potter, J.M.: Pile sorts, a cognitive anthropological model of drug and AIDS risks for Navajo teenagers: assessment of a new evaluation tool. *Drug Soc.* **7**, 23–39 (1993)