

# Sense of Community Index, reliability of measurements and their differences among race

Chris Aga  
University of Minnesota, Morris  
600 East 4th St.  
Morris, Minnesota 56267  
agaxx010@morris.umn.edu

## Keywords

Sense of Community, Sense of Community Index, African Americans, Information response theory, Differential Item Functioning

## 1. INTRODUCTION

Different ethnic groups have differing cultural and historical experiences, because of this it can be difficult to determine if measurement techniques used to gather information between groups are biased towards one group or if they are equal among all groups. This phenomenon is known as differential item functioning (DIF). Psychological sense of community (SOC) is typically measured by sense of community index (SCI) and was analyzed by Coffman for determining if variances were indicative of true racial differences in SOC or if any of the measurements used were biased towards a particular group.

This paper gives a summary of Coffman's experiment and findings in [4] and gives deeper summaries on the approaches used along with reasoning behind mathematical and statistical constructs. The following section gives a background on sense of community, operational constructs and measurement techniques. Section 3 discusses the data-set used for this experiment. Validity of operational constructs for use in SCI is depicted in Section 4. Section 5 then explains calculations made for choosing a proper SCI model. Next, Section 6 shows how DIF is statistically tested against the chosen mode, and finally Section 7 shows the results of the entire experiment.

## 2. SENSE OF COMMUNITY BACKGROUND

Sense of Community (SOC) has various definitions and is extremely important in sociology and psychology. In particular, Sarson introduced SOC in [9] and described it as being one of the most important concepts for definition of self. Sense of Community as defined by McMillan in [5] as "a spirit of belonging together, a feeling that there is an authority structure that can be trusted, an awareness that

trade, and mutual benefit come from being together, and a spirit that comes from shared experiences that are preserved as art".

## 2.1 Operational Constructs

Varying definitions exist concerning this term, but most literature follows a 4-piece operational system of SOC. The first construct is membership, which describes the idea that when people invest themselves into a community, they have a certain sense of a right to belong [6]. Influence, as defined by Peterson as "a sense that one matters, or can make a difference, in a community and that the community matters to its members" [7]. The third construct, dealing with the idea that the things desired and values shared in a community are similar or are the same. This is termed 'reinforcement of needs' [4]. Finally, SOC is also measured in this four-part operational construct system by a 'shared emotional connection' between member of the community.

### *Validity of four factors.*

Many studies have been conducted on these four constructs to determine their relevancy and overall weight in the calculation of an individual's SOC, or Sense of Community Index (SCI). Some have found that these four factors should be measured individually and combined with differing weights to determine one's SCI, while others have found no validity in these measurements and insist that SCI be measured as a one-factor solution. SCI as both a one-factor and four-factor instrument was heavily studied by Chipuer in [3]. The one-factor approach was determined as valid, and the four-factor approach had validity but there was much less consistency of measurements between data-sets than with the one-factor approach.

## 3. EXPERIMENT

Coffman in [4] attempted to determine if SCI differences were attributable to true differences in the amount of SOC between blacks and whites or if any of the questions used to calculate the SCI values were nonequivalent measures between race. This concept is known as differential item functioning, and concerns the prevention of biases of a certain measurement when looking at exogenous variables such as gender sex or race.

### 3.1 Data

The data-set used for Coffman's experiment consisted of 1,463 interviews to determine SCI. The interviews were conducted "using computer assisted telephone interview (CATI)

technology” [4]. The entire data-set used consisted of 588 whites and 840 blacks and a total of 12 questions were asked to determine an individual’s SCI. The questions were yes/no and thus the overall score ranged from 0-12 with higher numbers indicating a stronger perceived SOC.

## 4. DIMENSIONALITY

### 4.1 Overview

Before undergoing the main experiment, Coffman determined the number of factors from the four SOC operational constructs to be used. The final number of factors was carefully determined by a mathematically rigorous process known as full information factor analysis.

### 4.2 Computation and explanation

Full information item factor analysis first involved calculating the probability of a *person<sub>i</sub>* with a SOC score of  $\theta$  answering yes to a particular *question<sub>j</sub>*. SOC score in the one-factor model is  $\theta$ , whereas in for example, a four-factor model (all four operational constructs for sense of community) is represented as scores in:

$\theta = (\text{membership}, \text{influence}, \text{similarity}, \text{reciprocity})$ .

Probability is a function of item facility (the number of people who answered yes to a question), standard deviation in answers to a question and the mean answers to a question.

A probability distribution is generated for the entire data-set and is then integrated over to determine the probability of answering *question<sub>j</sub>* with a certain response. A core factor in the score is using what is known as “factor loading”. Factor loading is a weight applied to each operational construct for SOC, and these factors are calculated by minimizing the error output of the model. The combination of factor loadings and a person’s ability are combined in the probability function as described by Equation 1.

$$\text{ability} = \sum_{k=1}^m \alpha_{jk} \theta_{ki} \quad (1)$$

Marginal probability is then computed for *question<sub>j</sub>* based on all the probabilities per person with each specific vector of SOC factor scores. Finally, this marginal probability is combined along with a null hypothesis (not explained at all by the author for this particular experiment) and then a goodness-of-fit test (g-test) is applied. Initially one factor is used, full information analysis is computed, then g-test is applied. If the output is not statistically significant, no further factors are added and the dimensionality is now known for the SCI model [2].

## 5. CHOICE OF PARAMETERS LOGISTIC MODEL

### 5.1 Models

After assessing proper dimensionality for the experiment, a proper item response theory model must be chosen to determine which parameters are applicable to each question for studying this data-set. The parameters include difficulty and discrimination. Difficulty in this experiment was the standard deviation from average SOC within a group, of a person having a 50% chance to answer yes to a question [4]. Discrimination describes

The second parameter is discrimination, which as defined by Baker in [1], “describes how well an item can differentiate between examinees having abilities below the item location and those having abilities above the item location.” Discrimination was calculated using factor loadings which were based off of standard deviation in responses.

Many times a third, ‘guessing’ parameter is used since IRT often is used for knowledge assessment to prevent persons with lesser knowledge from simply achieving a higher score through lucky guessing. This however was not included in Coffman’s experiment since most people will not have to resort to guessing to answer questions such as “I care about what my neighbors think of my actions”.

### 5.2 Statistical analysis for final choice

Three models were suggested and tested to determine the proper approach for this experiment. Two 1PL models, one with a discrimination of 1 for each question, the other with an estimated discrimination value (equivalent for each question), and a 2PL model with discrimination values determined by factors loadings. The model used was chosen by comparing “Akaike information criterion, and Bayesian information criterion and log-likelihood nested tests” for each approach [4].

## 6. DIFFERENTIAL ITEM FUNCTIONING

Once all of the setup is completed, the core component of this experiment can be evaluated. In particular, Coffman wanted to assure that all questions being used to evaluate an individual’s SOC were not biased towards a particular group. After parameter value(s) (depending on 1PL or 2PL model) are gathered for each question, differential item functioning (DIF) is applied to each parameter to see if they are variant from group to group. The null hypothesis used for this experiment will be discussed later seeing as they are dependent on the chosen IRT model.

### 6.1 Calculation of Differential Item Functioning

Swaminathan in [10] describes various approaches to calculating the DIF of items of single or multiple factors as well as working with models made up of single or multiple parameters for characteristics of questions. As will be depicted in Sections 4 and 5, a one-factor, two-parameter based model was chosen and thus the following procedure for calculating DIF was used.

First, a predictor for probability of answering yes to a question by modeling between different groups is described by Equation 2. The difficulty of a question is  $\beta_{0j}$ , variance of response on an item for people with the same SOC score is represented by  $\beta_{1j}$ , and  $\theta_{ij}$  describes the SOC of *person<sub>i</sub>* in *group<sub>j</sub>*.

$$P = \frac{e^{(\beta_{0j} + \beta_{1j} * \theta_{ij})}}{1 + e^{(\beta_{0j} + \beta_{1j} * \theta_{ij})}} \quad (2)$$

Swaminathan defines DIF as: “an item shows DIF if individuals with the same ability but from different groups do not have the same probability of success on the item” [10]. Coffman explains that these probabilities between groups are transformed into chi-squared distributions then statistically analyzed to determine if any of the parameters demonstrate DIF within the experiment [4].

<i>Item</i>	<i>Factor loading</i>	<i>Biserial correlations</i>	<i>Facility</i>
1) I think my neighborhood is a good place to live.	.84	0.656	0.918
2) People in this neighborhood believe the same things are important.	.84	0.694	0.842
3) My neighbors and I want the same thing.	.83	0.694	0.846
4) I can recognize many people who live in my neighborhood.	.55	0.494	0.841
5) I feel at home in my neighborhood.	.86	0.614	0.947
6) Very few of my neighbors know me.	.53	0.500	0.697
7) I care about what my neighbors think of my actions.	.54	0.541	0.764
8) I have almost no influence over what this neighborhood is like.	.49	0.505	0.638
9) If there is a problem in this neighborhood, the people here can solve it.	.75	0.664	0.826
10) It is very important to me to live in this neighborhood.	.78	0.703	0.726
11) I expect to live in this neighborhood for a long time.	.67	0.590	0.790
12) People around here are willing to help their neighbors.	.79	0.675	0.892

Figure 1: Table of questions used to determine SCI, includes factor loadings which describe the weight of each question when determining SCI and analyzing DIF [4]

## 7. RESULTS

### 7.1 Chosen dimensionality

Full information item factor analysis was conducted on the gathered data-set multiple times to determine the proper dimensionality to use for this experiment. Though statistical analysis in TESTFACT, a program for scoring tests, and calculating factor analysis. The results, due to loss of significance in g-tests showed that a one-factor model was to be used for the experiment instead of using multiple operational constructs for scoring of SOC. Factor loadings for each question were generated from the one-factor full-information-analysis and are shown in 4.1. Biserial correlations were a coefficient that described the relationship between a question and the average SOC of a group or groups. A large biserial correlation for a question would mean that people who got high SOC scores typically answered yes to this question. Facility was the percentage of people in the entire data-set who answered yes to a particular question [11].

### 7.2 Choice of information response theory model

The statistical analysis tool R was used, and a latent trait modeling (LTM) package was implemented to determine the proper IRT model for this experiment. The three statistics described briefly in Section 5.2 were calculated by LTM, compared for significance and the chosen model was 2PL. Thus both difficulty and discrimination parameters were included and examined for statistically significant DIF [4].

### 7.3 Differential item functioning analysis

Probability distributions were generated for each item, for both blacks, whites and a combination of both groups to illustrate the overall impact each question had based on scored SOC. These results are shown in Figure 7.1.

Information plots were also generated for the means of determining if particular items were more accurate descriptors of SOC for blacks or whites. Information in IRT, as described by Baker, is calculated using the precision of an

item's parameter. Baker in [1] uses Equation 3 to produce information  $I$ .  $\sigma$  represents the amount of standard deviation, this value is squared to get variance, then the reciprocal is computed to obtain  $I$ .

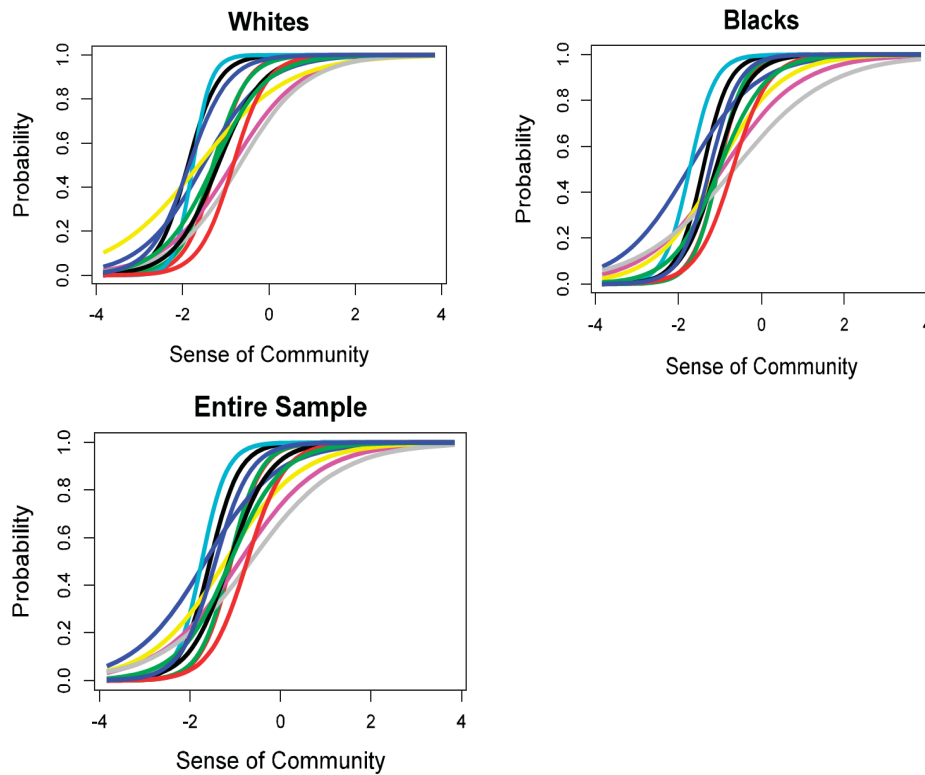
$$I = \frac{1}{\sigma^2} \quad (3)$$

Standard deviation in [4] is described by how much the SOC value varies from a probability of 0 to 1. Therefore, questions with low standard deviations will, based on Equation 3, depict how useful a question is at determining someone's SOC.

If information differences at the same value of SOC are statistically significant then DIF is computed to determine if a  $question_x$  is a good indicator at determining SOC differences between race or if  $question_x$  is not important for the actual determination of SOC [4]. Significant differences in a particular question, "I feel at home in my neighborhood" were found at areas of similar SOC between blacks and whites in the experiment, thus DIF was calculated on the results to determine how accurate of a determiner this question was for each group. After statistically analyzing parameters for DIF, none were found.

## 8. DISCUSSION

Full information factor analysis found that a one-factor approach (amount of SOC versus SOC calculated from amounts of membership etc) was a correct dimensionality to use for SCI in this experiment. Coffman described how other research found support for four-factor SCI models, but explained that differences in those studies versus this study demeaned assurance of accurate models. First, data was dichotomous since the questions asked were either yes or no, thus full information factor analysis for choosing proper dimensions was necessary [4]. Also, other studies such as [8] had used different scales for measuring SOC, thus their finding of four-factor support were not valid for justifying its automatic acceptance in this experiment.



**Figure 2: Probability distribution of all 12 questions for whites blacks and for all candidates**

After all the appropriate models were chosen, DIF was calculated for both difficulty and discrimination parameters on each question. Greater information was found for one question (5) when determining SOC for whites as opposed to blacks. But after statistical analysis on both question parameters, DIF was no existent thus differences in SCI between blacks and whites were likely indicative of true differences in SOC between race.

For question 5, “I feel at home in my neighborhood”, low standard deviation was found in particular or white, so it mean that this was a good indicator for measuring a white person’s SOC because of more consistent answers at particular SOC scores. But, white people with SOC score  $\theta$  did not tend to answer yes much more than black people with the same  $\theta$  score, they just had more consistent answers to this question. These results show that differences in SCI measurements between blacks and whites are not attributable to differential item functioning in a particular question, but are more likely measures of true race differences in SOC.

## 9. REFERENCES

- [1] F. B. Baker. *The basics of item response theory*. ERIC, 2001.
- [2] R. D. Bock, R. Gibbons, and E. Muraki. Full-information item factor analysis. *Applied Psychological Measurement*, 12(3):261–280, 1988.
- [3] H. M. Chipuer and G. M. Pretty. A review of the sense of community index: Current uses, factor structure, reliability, and further development. *Journal of Community psychology*, 27(6):643–658, 1999.
- [4] D. L. Coffman and R. BeLue. Disparities in sense of community: True race differences or differential item functioning? *Journal of community psychology*, 37(5):547–558, 2009.
- [5] D. W. McMillan. Sense of community. *Journal of community psychology*, 24(4):315–325, 1996.
- [6] D. W. McMillan and D. M. Chavis. Sense of community: A definition and theory. *Journal of community psychology*, 14(1):6–23, 1986.
- [7] N. A. Peterson, P. W. Speer, and D. W. McMillan. Validation of a brief sense of community scale: Confirmation of the principal theory of sense of community. *Journal of Community Psychology*, 36(1):61–73, 2008.
- [8] N. A. Peterson, P. W. Speer, and D. W. McMillan. Validation of a brief sense of community scale: Confirmation of the principal theory of sense of community. *Journal of Community Psychology*, 36(1):61–73, 2008.
- [9] S. B. Sarason. *The psychological sense of community: Prospects for a community psychology*. Jossey-Bass San Francisco, 1974.
- [10] H. Swaminathan and H. J. Rogers. Detecting differential item functioning using logistic regression procedures. *Journal of Educational measurement*, 27(4):361–370, 1990.
- [11] S. Varma. Preliminary item statistics using point-biserial correlation and p-values. *Educational Data Systems Inc.: Morgan Hill CA*. Retrieved, 16(07), 2006.