Assessing the Validity of Diversity Indices

Assessing the Validity of Diversity Indices

By

Salomón Alcocer Guajardo

Cambridge
Scholars
Publishing

Assessing the Validity of Diversity Indices

By Salomón Alcocer Guajardo

This book first published 2024

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-5275-7378-8
ISBN (13): 978-1-5275-7378-9

This book is dedicated to Professor Gábor J. Székely
and dedicated posthumously to Professor Lee Joseph Cronbach
(1916 – 2001) and to Professor Charles Edward Spearman (1863 – 1945)
for their statistical methods that are used extensively in this book and
for their invaluable contributions to the quantitative assessment
of measurement reliability and validity.

CONTENTS

EXHIBITS

NYC DEPARTMENTS

NYC Department	Acronym
Administration for Children's Services	ACS
Board of Corrections	BOC
Board of Election	BOE
Borough President-Bronx	BP-BX
Borough President-Brooklyn	BP-BK
Borough President-Manhattan	BP-MAN
Borough President-Queens	BP-QNS
Borough President-Staten Island	BP-SI
Business Integrity Commission	BIC
Campaign Finance Board	CFB
City Commission on Human Rights	CCHR
Civilian Complaint Review Board	CCRB
Conflicts of Interest Board	COIB
Department for the Aging	DFTA
Department of Buildings	DOB
Department of City Planning	DCP
Department of Citywide Administrative Services	DCAS
Department of Consumer Affairs	DCA
Department of Correction	DOC
Department of Cultural Affairs	DCLA
Department of Design & Construction	DDC
Department of Education	DOE
Department of Environment Protection	DEP
Department of Finance	DOF
Department of Health/Mental Hygiene	DOHMH

Department of Homeless Services	DHS
Department of Info Tech & Telecomm	DOITT
Department of Investigation	DOI
Department of Parks & Recreation	PARKS
Department of Probation	DOP
Department of Records & Information Service	DORIS
Department of Sanitation	DSNY
Department of Small Business Services	SBS
Department of Transportation	DOT
Department of Youth & Community Development	DYCD
District Attorney - Bronx County	DA-BX
District Attorney - Kings County	DA-BK
District Attorney - Manhattan	DA-MAN
District Attorney - Queens County	DA-QNS
District Attorney - Richmond County	DA-SI
District Attorney – Special Narcotics	DA-NARC
Equal Employment Practices Commission	EEPC
Financial Information Services Agency	FISA
Fire Department	FDNY
Housing Preservation & Development	HPD
Human Resources Administration / Social Services	HRA
Independent Budget Office	IBO
Landmarks Preservation Committee	LPC
Law Department	LAW
MAYORALTY	MAYORALTY
Municipal Water Finance Authority	MWFA
New York City Council	COUNCIL
New York City Fire Pension Fund	FDNYPF
New York City Police Pension Fund	NYCPPF
New York City Tax Commission	NYCTAX

NYC Civil Service Commission	NYCCSC
NYC Employees Retirement System	NYCERS
NYC Health + Hospitals	NYCHH
NYC Housing Authority	NYCHA
Office of Administrative Trials & Hearings	OATH
Office of Collective Bargaining	OCB
Office of Emergency Management	NYCEM (OEM)
Office of Payroll Administration	OPA
Office of the Actuary	ACTUARY
Office of the City Clerk	CLERK
Office of the Comptroller	COMPTROLLER
Office of the Public Advocate (PA)	PA
Offices of the Public Administrators	PUBADMIN
Police Department	NYPD
School Construction Authority	SCA
Taxi & Limousine Commission	TLC
Teachers Retirement System	TRS

ACKNOWLEDGEMENTS

I would like to express my appreciation and gratitude to Courtney Dixon, Sophie Edminson, Amanda Millar, Adam Rummens, and the production staff of Cambridge Scholars Publishing for making this book a reality. I also would like to thank Andre Hall, Tiffany Nichole Peña, and Steve Santiago for their immeasurable and unending encouragement and support. This book would not be possible without them.

About the Author

Salomón Alcocer Guajardo, Ph.D., is a management consultant and researcher. He is a former associate professor of nonprofit and public management. Dr. Guajardo has held executive budgeting and finance positions in municipal government. He graduated from the University of California, Los Angeles (UCLA) with a Bachelor of Arts in political science. Dr. Guajardo received graduate degrees in public management and educational psychology from the University of Pittsburgh. He received a doctorate degree in public policy research and analysis from the Graduate School of Public and International Affairs, University of Pittsburgh. His research on compensation, diversity, and organizational efficiency has been published in leading peer-reviewed journals.

PREFACE

Despite the voluminous number of studies which have assessed and measured demographic (or social) diversity in nonprofit, private, and public organizations, there is a shortage of studies which have assessed the measurement validity and reliability of diversity indices that have been developed and used to measure demographic heterogeneity in the workplace. As discussed in previous companion books, the measurement reliability of some of the diversity indices used to measure age, ethnic, and gender diversity is questionable (Guajardo, 2023a, 2023b, 2023c, 2023d, 2023e, and 2024a).

Several studies have conducted cursory assessments of measurement validity and reliability of several indices of diversity that have been discussed in previous companion books (e.g., Guajardo, 1996, 2013, and 2015; Mercil and Williams, 1984). For instance, Guajardo (1996) conducted a narrow measurement reliability assessment of the index of qualitative variation (IQV or measure of variation (MV); Mueller, Schuessler, and Costner, 1970; Wilcox, 1967). In another study, Guajardo (2013) assessed how ethnic diversity scores obtained by the IQV, the McIntosh (1967; M_E) evenness index, and the Simpson's (1949) index of diversity changed when the number of categories were consolidated. However, the study did not perform statistical tests for construct and measurement validity. Guajardo (2015) also conducted a limited measurement reliability assessment of ethnic diversity scores obtained with the IQV, McIntosh, and Simpson indices when the indices were applied to the police and civilian workforces of the New York City Police Department (NYPD). The findings indicated the indices produced consistent diversity scores.

Mercil and Williams (1984), on the other hand, correlated and factor analyzed 14 indices of diversity and segregation that were applied to ethnic-based demographic data collected from 605 private and public schools in the US in 1980. The factor analysis extracted 3 factors and indicated the ethnic diversity and segregation scores were multidimensional. However, Mercil and Williams did not assess the construct and measurement validity of the indices specifically. They also did not assess measurement reliability

in terms of the indices yielding consistent ethnic diversity or segregation scores.

In addition to a lack of empirical studies focused specifically on the measurement validity and reliability of indices of diversity, the discriminatory power and measurement efficiency of diversity indices has not been assessed thoroughly. With respect to the discriminatory power of the indices, it is unclear how well indices of diversity detect subtle differences in the level of age, ethnic, and gender heterogeneity when they are applied to nonprofit, private, or public organizations. Consequently, it is unclear whether indices with complex mathematical formulas (e.g., $D_W = 1 - \sqrt{\frac{\sum p^2 - \frac{1}{n}}{\left(\frac{n-1}{n}\right)}}$) produce distributions with greater levels of variation (or variability) in comparison to indices with simplistic formulas (e.g., $D_{SW3} = \frac{-\sum p^2}{ln(n)}$).

Due to a lack of sufficient empirical research on the validity and reliability of indices of diversity, this book focuses exclusively on assessing the measurement reliability and the construct and measurement validity of the diversity indices that were applied to employment data reported by NYC departments for the fiscal year (FY) 2019. In so doing, this book applies correlation and factor analysis exclusively on the evenness (or standardized) diversity scores obtained for NYC departments for FY 2019. The rationale for doing so is that the standardized (or evenness) diversity scores are compatible amongst NYC departments. For consistency with previous companion books, Pearson correlation analysis is used to assess how well the heterogeneity scores for age, ethnicity, and gender that were obtained by the diversity indices are aligned with each other (Guajardo, 2023a, 2023b, 2023c, 2023d, 2023e, and 2024a). Factor analysis with unrotated and rotated principal component factors is also performed to assess whether the diversity indices obtain measures of the same construct. In addition, the assessment of construct validity is supplemented with distance correlation analysis because most of the indices produce nonnormal (or asymmetrical) distributions of diversity scores that are negatively skewed and heavy tailed. By undertaking distance correlation analysis, it is possible to assess the extent to which the diversity scores are alike. Estimates of measurement reliability are assessed by obtaining Cronbach alpha (α) coefficients.

This book also compares each diversity index discussed in previous companion books (Guajardo, 2023a, 2023b, 2023c, 2023d, 2023e, and

2024a). Specifically, each diversity index is assessed with respect to its discriminatory power and measurement efficiency. The coefficient of variation (CV) used in previous companion books is used to examine the variation within the distributions of diversity scores obtained by each index. Measurement efficiency is assessed by the volume of mathematical operations that are required by an index to obtain an evenness (or standardized) diversity score.

In conducting the measurement validity and reliability assessments of the heterogeneity scores obtained by the indices of diversity, this book addresses fundamental measurement issues and questions that arise when indices of diversity are applied to qualitative demographic or employment data to obtain measures of diversity. The issues and questions addressed in this book include the following:

- Do standardized indices of diversity produce consistent scores of age, ethnic, and gender heterogeneity based on the demographic employment data reported by organizations?
- Do standardized indices of diversity produce unidimensional scores of age, ethnic, and gender heterogeneity when they are applied to qualitative employment data reported by organizations?
- Do standardized indices of diversity with complex mathematical formulas produce distributions of heterogeneity scores of age, ethnic, or gender with greater levels of variation in comparison to indices with simplistic mathematical formulas?
- How similar are diversity scores obtained with different indices of diversity when they are applied to age-, ethnic-, and gender-based employment data?

These issues are addressed in this book because little empirical research has been devoted to examining the validity and reliability of diversity indices used to measure and analyze demographic (or social) diversity in organizations. Although the issues and questions addressed in this and its companion books are fundamental to carrying out empirical research, practitioners and researchers alike often ignore or take the analytical or measurement issues and questions for granted.

As stated in previous companion books, the book series consists of 9 books (Guajardo, 2023a, 2023b, 2023c, 2023d, 2023e, and 2024a). They are the following:

- *Assessing Organizational Diversity with the Simpson Index* applies the Simpson diversity index to demographic and employment data reported by New York City (NYC) departments for FY 2019. This book focuses on the application and analysis of Simpson diversity formulas for calculating biased and unbiased measures of demographic heterogeneity.
- *Assessing Organizational Diversity with the Shannon Index* applies the Shannon diversity index to the same demographic and employment data used in the first book. This book focuses exclusively on the application and analysis of Shannon diversity formulas for calculating biased and unbiased measures of demographic heterogeneity.
- *Assessing Organizational Diversity with the Heip Index* applies the Heip, Sheldon, and other Shannon-based diversity indices to the data used in the first and second books. The Heip, the Sheldon, and the other Shannon-based diversity indices presented in the book are modifications of the Shannon index of diversity. From a statistical standpoint, the Heip and Sheldon indices possess statistical properties that are superior to the original Shannon index. Like the first and second books, this book focuses on the application and analysis of the indices in terms of measuring demographic heterogeneity in organizations.
- *Assessing Organizational Diversity with the Smith and Wilson Indices* applies the Smith and Wilson (SW) indices to the same data used in the previous companion books. In addition to applying the SW indices, other Simpson-based indices such as the Ray and Singer (RS) index of concentration are presented in the book. The SW and RS indices are modifications of the Simpson ($D = 1- \sum p^2$) diversity index and assess demographic heterogeneity as well. This book applies the Simpson-based indices to the same data used in previous books to measure demographic heterogeneity in organizations.
- *Assessing Organizational Diversity with the McIntosh Index* applies the McIntosh evenness index to the same demographic and employment data used in the preceding companion books. Because the index incorporates the number of groups used to categorize a demographic (or social) characteristic of interest and the size of the workforce simultaneously, the diversity scores contain less measurement bias and have a greater degree of compatibility in comparison to the other diversity indices covered in the preceding companion books.
- *Assessing Organizational Diversity with the Index of Qualitative Variation* (IQV) applies the Mueller and Schuessler IQV (Wilcox, 1967) to the same demographic and employment data used in previous companion books. Because the IQV may not be invariant to ordering

sequences, this book focuses on the application and analysis of heterogeneity scores obtained from the different ordering sequences of the data. Like the McIntosh evenness index presented in the 5th book, the IQV incorporates jointly the number of groups used in the categorization of the demographic (or social) characteristic of interest and the size of the workforce.

- *Assessing the Validity of Diversity Indices* compares the indices used in each book jointly and uses factor analysis to determine whether they assess the same (or different) aspects of demographic (or social) diversity. Pearson pairwise correlation analyses are also performed to assess the statistical associations amongst the diversity indices. Statistical analyses for equality of means are performed as well.
- *Assessing Organizational Diversity with Quantile Regression* applies quantile regression analysis to each of the diversity indices presented in this book series. This book performs quantile regression analyses at the 25th, 50th, 75th, and 90th percentiles for age, ethnic, and gender diversity.
- *Assessing Organizational Diversity with Structural Equation Modeling* (SEM) focuses exclusively on causal modeling. This book focuses on the development and analysis of a SEM for specific diversity indices discussed in the book series. In so doing, the analyses treat age, ethnic, and gender diversity as intervening (or mediating) variables of organizational performance.

For purposes of continuity and compatibility, each diversity index is subjected to the same statistical analyses. The IQV, McIntosh evenness, Shannon, Simpson, and SW indices are of special focus in this book series because they have been used in previous research on the demographic (or social) diversity in nonprofit, private, or public organizations.

This book series is written for practitioners and researchers in human resources and other fields that are interested in measuring and analyzing demographic, occupational, or social heterogeneity in organizations. The purpose of the book series is to addresses measurement and analytical issues that practitioners and researchers alike are likely to face when they apply a particular diversity index to demographic and employment data provided by a nonprofit, private, or public organization. As such, this book series should serve as a reference for selecting the diversity index that is best suited for measuring and analyzing heterogeneity in an organizational setting. This book series also should serve as a reference for selecting the statistical method that is best suited for analyzing the distribution of scores obtained by the diversity index of choice.

References

Guajardo, Salomón A. 1996. “Representative bureaucracy: An estimation of the reliability and validity of the Nachmias-Rosenbloom MV Index”. *Public Administration Review*, Vol. 56: 467 – 477.

Guajardo, Salomón A. 2013. “Workforce diversity: An application of diversity and integration indices to small agencies”. *Public Personnel Management*, Vol. 41: 27 – 40.

Guajardo, Salomón A. 2015. “Measuring diversity in police agencies”. *Journal of Ethnicity in Criminal Justice,* Vol. 13: 1 – 15.

Guajardo, Salomón A. 2023a. *Assessing organizational diversity with the Simpson index*. UK: Cambridge Scholars Publishing.

Guajardo, Salomón A. 2023b. *Assessing organizational diversity with the Shannon index*. UK: Cambridge Scholars Publishing.

Guajardo, Salomón A. 2023c. *Assessing organizational diversity with the Heip index*. UK: Cambridge Scholars Publishing.

Guajardo, Salomón A. 2023d. *Assessing organizational diversity with the Smith and Wilson indices*. UK: Cambridge Scholars Publishing.

Guajardo, Salomón A. 2023e. *Assessing organizational diversity with the McIntosh index*. UK: Cambridge Scholars Publishing.

Guajardo, Salomón A. 2024a. *Assessing organizational diversity with the index of qualitative variation*. UK: Cambridge Scholars Publishing.

McIntosh, Robert P. 1967. "An index of diversity and the relation of certain concepts to diversity." *Ecology*, Vol. 48: 392 – 404.

Mercil, Steven Bray and Williams, John Delane. 1984. “Measures of school integration: Comparing Coleman's index to measures of species diversity.” New Orleans, LA: Annual Meeting of the American Educational Research Association.

Mueller, John H., Schuessler, Karl F., and Costner, Herbert L. 1970. *Statistical reasoning in sociology*. Second edition. New York, NY: Houghton Mufflin Company.

Simpson, Edward Hugh. 1949. “Measurement of diversity”. *Nature*, Vol. 163: 688.

Wilcox, Allen. R. 1967. *Indices of Qualitative Variation*. Oak Ridge, TN: Oak Ridge National Laboratory, U.S. Atomic Energy Commission.

CHAPTER 1

INTRODUCTION

Although numerous unstandardized and standardized indices of *diversity* have been discussed and analyzed statistically in previous companion books, the indices were not assessed concomitantly or thoroughly with respect to *measurement efficiency*, *validity*, and *reliability* (Guajardo, 2023a, 2023b, 2023c, 2023d, 2023e, and 2024a). While *factor analysis* methods were undertaken to assess whether some of the diversity indices had the same statistical structure and whether they measured the same dimension of diversity, a comprehensive statistical analysis of the *construct* and measurement validity of the indices was not undertaken previously. Similarly, the *Pearson correlation* analyses that were performed previously assessed some of the diversity indices with respect to how well the heterogeneity scores aligned with each other. In so doing, the diversity indices that were covered in some of the companion books were excluded from the correlation analyses. Consequently, it is unclear how well the heterogeneity scores of the diversity indices that were excluded from the Pearson correlation analyses align with the scores obtained with other diversity indices. Previous companion books also did not conduct *distance correlation* analyses to assess the measurement validity amongst the standardized (or evenness) heterogeneity scores obtained with the various diversity indices. In addition, previous companion books did not conduct statistical assessments of the measurement reliability of heterogeneity scores obtained by the diversity indices. Because the extant research on the demographic (or social) diversity in nonprofit, private, and public organizations has given little attention to the assessment of measurement validity and reliability of the indices used to measure diversity, this book undertakes a comprehensive and systematic statistical analysis of the measurement validity and reliability of the indices covered in previous companion books (Guajardo, 2023a, 2023b, 2023c, 2023d, 2023e, and 2024a).

While previous companion books have analyzed systematically each diversity index covered with respect to the statistical properties of the

distribution of scores, the indices have not been classified with respect to whether they are probability- or logarithm-based. The book series began with *Assessing Organizational Diversity with the Simpson Index* (Guajardo, 2023a) which focused exclusively on Simpson's (1949) diversity index, an index developed to assess the *probability* that two or more members of a community would possess the same (or different) characteristics. Simpson-based diversity indices such as the Ray and Singer (1973) index and the Smith and Wilson (1996) indices were discussed in *Assessing Organizational Diversity with the Smith and Wilson indices* (Guajardo, 2023d). Each index discussed in the book is probability based. The McIntosh (1967) evenness diversity index, which uses the sum of squared percentages (or probabilities) of each category ($M = 1 - \sqrt{\sum p^2}$), was discussed and analyzed exclusively in *Assessing Organizational Diversity with the McIntosh Index* (Guajardo, 2023e). However, the index was not presented as a probability-based index. In *Assessing Organizational Diversity with the Index of Qualitative Variation* (IQV; Guajardo, 2024a), the book presented the IQV as producing a series of cross products (i.e., joint probabilities) of the percentage (or probability) of one group multiplied by the percentage (or probability) of each other group in a community that are summarized in a matrix containing *joint* and *marginal probabilities* (or observed differences). Because the IQV uses the sum of the joint probabilities ($\Sigma p_i p_j$), it is classified as a probability-based index.

Assessing Organizational Diversity with the Shannon Index (Guajardo, 2023b) focused exclusively on the Shannon (1948) index. As presented in the book, the Shannon index uses logarithm-based formulas to obtain unstandardized and standardized (or evenness) diversity scores. Shannon-based diversity indices were presented and analyzed systematically in *Assessing Organizational Diversity with the Heip Index* (Guajardo, 2023c). Each index presented in the book is logarithm based. Although Smith and Wilson (1996) developed a hybrid evenness index which incorporates aspects of the Shannon (1948) and Simpson (1949) indices ($D_{SW3} = \frac{-\ln \sum p^2}{\ln n}$), D_{SW3} was not included in the book because it was covered in *Assessing Organizational Diversity with the Smith and Wilson Indices* (Guajardo, 2023d). In this book, D_{SW3} is classified as a logarithm-based index, but it is included in the statistical analysis of probability-based indices because of its use of probabilities (i.e., Σp^2).

This introductory chapter reviews the diversity indices discussed and analyzed in previous companion books (Guajardo, 2023a, 2023b, 2023c, 2023d, 2023e, and 2024a). In doing so, the diversity indices are categorized

based on the nature of the index. Indices based on the Shannon (1948) index are categorized as logarithm-based. By contrast, indices based on the percentage (or probability) of individuals in various categories (or groups) are classified as probability-based. After reviewing and classifying each index, the chapter discusses the statistical techniques which are applied to the diversity indices to assess their validity and reliability. The statistical techniques for assessing construct and measurement validity are discussed before discussing the techniques that are used to assess measurement reliability. For consistency with previous companion books, a summary is provided to highlight the key points of each chapter.

Overview of diversity Indices

Several evenness indices that quantify *demographic* (or *social*) *diversity* (or heterogeneity) in communities or organizations were presented in previous companion books (Guajardo, 2023a, 2023b, 2023c, 2023d, 2023e, and 2024a). After applying each index to the qualitative employment data reported by NYC departments for the fiscal year (FY) 2019, the indices were analyzed statistically with respect to the characteristics of the distribution of the scores. The statistical analyses found that most of the indices produced standardized (or evenness) diversity scores that had negatively skewed and heavy-tailed distributions, indicating the distributions were nonnormal (i.e., asymmetrical). In addition, the statistical analyses revealed that most of the indices produced distributions of diversity scores that were clustered around the mean so that they had low levels of *measurement variation* (or variability). Although previous companion books conducted a comprehensive and systematic statistical analysis of the diversity indices, the companion books did not address the underlying nature of the indices. Stated simply, the companion books did not classify the indices with respect to the statistical basis of the scores. Consistent with McDonald and Dimmick (2003), the standardized (or evenness) diversity indices used in the book series are categorized as probability- or logarithm-based.

Probability- based diversity indices

When developing the index of diversity ($D = 1 - \Sigma p^2$), Simpson (1949, 688) defined diversity as the "probability that two individuals chosen at random and independently from the population will be found to belong to the same group." Based on Simpson's definition, probability-based standardized evenness diversity indices are rooted in the probability that two or more individuals selected randomly from a community or organization will differ

with respect to their demographic (or social) characteristics (McDonald and Dimmick, 2003). Exhibit 1-1 categorizes the standardized diversity indices that are based on Simpson's index of diversity.

Exhibit 1-1. Probability-based standardized diversity indices

Index	Formula	Empirical Maximum Value (EMV)
Hussain and Khan (2019) index	$D_{HK} = \frac{1}{n-1}\left(\frac{1}{f_m} - 1\right)$	
McIntosh (1967) index for unequal groups	$M_E = \frac{1-\sqrt{\sum p^2}}{1-\frac{1}{\sqrt{n}}}$	$M_{EMax} = 1 - \frac{1}{\sqrt{n}}$
Muller and Schuessler index (Wilcox, 1967)	$IQV = \frac{\sum(p_i p_j)}{\frac{n(n-1)}{2}\left(\frac{1}{n}\right)^2}$	$IQV_{Max} = \frac{n(n-1)}{2}\left(\frac{1}{n}\right)^2$
Ray and Singer (1973) index	$D_{RS} = 1 - \sqrt{\frac{\sum p^2 - \frac{1}{n}}{1-\frac{1}{n}}}$	$D_{RSMax} = 1 - \frac{1}{n}$ or $\frac{n-1}{n}$
Simpson (1949) index (biased)	$S_E = \frac{1-\sum p^2}{\frac{n-1}{n}}$	$S_{Max} = 1 - \frac{1}{n}$ or $\frac{n-1}{n}$
Simpson (1949) index (unbiased)	$S_{EU} = \frac{S_u}{\frac{n-1}{n}}$	$S_{Max} = 1 - \frac{1}{n}$ or $\frac{n-1}{n}$
	$S_U = \sum \frac{n(n-1)}{N(N-1)}$	
Simth and Wilson (1996) indices	$D_{SW1} = \frac{1-\sum p^2}{1-\frac{1}{n}}$	$D_{SW1Max} = 1 - \frac{1}{n}$ or $\frac{n-1}{n}$
	$D_{SW2} = \frac{\left(\frac{1}{\sum p^2}\right)}{n}$	$D_{SW2Max} = n$
Wilcox (1967) index	$D_W = 1 - \sqrt{\frac{\sum p^2 - \frac{1}{n}}{\left(\frac{n-1}{n}\right)}}$.	$D_{WMax} = 1 - \frac{1}{n}$ or $\frac{n-1}{n}$

As discussed in *Assessing Organizational Diversity with the Smith and Wilson Indices* (Guajardo, 2023d), D_{HK} obtains an evenness diversity score by using the *reciprocal value* ($\frac{1}{f_m}$) of the category (or group) with the largest percentage (see Exhibit 1-1). According to Hussain and Khan (2019), the index produces evenness scores ranging from 0 to 1. A score of 0 indicates the distribution of individuals is concentrated in 1 category (or group). When individuals are distributed equally across all categories (or groups), D_{HK} obtains a score of 1. Although Hussan and Khan (2019) used 15 tests to examine the mathematical features of the index (see Smith and Wilson, 1996), they did not assess the index with respect to construct and measurement validity. They also did not assess the measurement reliability of the index.

The McIntosh (1967) evenness index for unequal groups was discussed and analyzed statistically in *Assessing Organizational Diversity with the McIntosh Index* (Guajardo, 2023e; see Exhibit 1-1). When developing the evenness index, McIntosh (1967) squared and summed the proportions (or probabilities) of each group (Σp^2) and then took the square root of Σp^2 to obtain an unstandardized score of diversity by subtracting $\sqrt{\sum p^2}$ from 1 as follows: $M = 1 - \sqrt{\sum p^2}$. In addition, McIntosh (1967) estimated the *maximum empirical value* (EMV) of M_E based on the number of categories (or groups) identified in a community. According to McIntosh, the EMV is obtained with the following formula when the frequencies (or probabilities) are unequal in each category (or group): $M_{EMax} = 1 - \frac{1}{\sqrt{n}}$. An evenness diversity score is obtained by dividing the unstandardized score by the EMV. In doing so, M_E obtains a distribution of scores ranging from 0 to 1. A score of 0 indicates the individuals in a community are concentrated in 1 category (or group). A standardized score of 0 also indicates a community did not attain any proportion of the EMV. Conversely, a standardized score of 1 indicates the individuals in a community are distributed equally amongst the groups and that the community attained 100% of the EMV.

When developing the evenness indices of diversity, McIntosh (1967) compared the indices to the Shannon (1948) and Simpson (1949) indices (see Exhibit 1-1 and 1-2). However, McIntosh did not test the indices with respect to construct or measurement validity. McIntosh also did not test the indices with respect to measurement reliability.

As developed by Mueller and Schuessler (Mueller, Schuessler, and Costner, 1970; Wilcox, 1967), the Index of Qualitative Variation (IQV) uses the

proportions (or probabilities) of each group to calculate cross products (or joint probabilities) and to obtain the "total number of observed differences" in a community or organization (see Exhibit 1-1). Based on Wilcox's (1967) formula, the IQV obtains the total number of observed differences by summing each of the cross products (or joint probabilities) as follows: $\Sigma p_i p$. Like M_E, the IQV has an EMV based on the number of categories (or groups) in a community or organization. According to Wilcox (1967), the maximum number of possible differences is obtained with the following formula: $IQV_{Max} = \frac{n(n-1)}{2}\left(\frac{1}{n}\right)^2$. When $\Sigma p_i p_j$ is divided by the EMV, a standardized diversity score is obtained. Like M_E, the IQV produces a distribution of scores ranging from 0 to 1. A score of 0 indicates a community or organization did not attain any proportion of the EMV; a score of 1 indicates 100% of the EMV was attained by a community or organization. When treated as a measure of evenness, an IQV score of 1 indicates the individuals in a community or organization are distributed equally amongst the categories (or groups).

When the formula for the IQV was created, Wilcox (1967) did not apply data to the IQV. Wilcox also did not assess the construct or measurement validity of the IQV with respect to measuring heterogeneity (or variation). Wilcox also did not assess the measurement reliability of the IQV with respect to yielding consistent scores. Similarly, Mueller, Schuessler, and Costner (1970) did not assess the measurement validity or reliability of the IQV when they applied the index to demographic data.

Although the IQV was analyzed statistically in *Assessing Organizational Diversity with the Index of Qualitative Variation* (Guajardo, 2024a) with respect to the characteristics of the distribution of diversity scores, the construct and measurement validity of the IQV was not assessed. The measurement reliability of the IQV was not assessed as well. Guajardo (1996) conducted a cursory reliability assessment of the IQV. In so doing, however, Guajardo (1996) did not compare the IQV to other diversity indices such as the McIntosh (1967), Shannon (1948), or Simpson (1949) indices.

Ray and Singer (1973) developed the index in concentration to assess whether individuals are "concentrated' in one group or distributed equally amongst groups. The formula for the index of concentration is as follows: $C_{RS} = \sqrt{\frac{\Sigma p^2 - \frac{1}{n}}{1 - \frac{1}{n}}}$. As the formula indicates, the index uses the proportions (or

probabilities) of each category (or group) to obtain a concentration score. The EMV is obtained with the following formula: $C_{RSMax} = 1 - \frac{1}{n}$. An equivalent formula for obtaining the EMV is as follows: $C_{RSMax} = \frac{n-1}{n}$. When C_{RS} is subtracted from 1, the index yields a diversity score equivalent to the IQV (Taagepera and Ray, 1977) and to the Wilcox (1967) index of diversity (D_W; see Exhibit 1-1). Like the IQV and M_E, D_{RS} has a distribution of scores ranging from 0 to 1. Briefly, a score of 0 indicates individuals are clustered in 1 category (or group), and a score of 1 indicates individuals are distributed equally amongst the categories (or groups).

When developing C_{RS}, Ray and Singer (1973) compared the index to other indices. However, they did not assess the index's construct or measurement validity. They also did not assess the index's measurement reliability. Taagepera and Ray (1977) also compared C_{RS} to other indices of variation but did not assess the index's measurement validity and reliability when doing so.

Along with the Shannon (1948) index of diversity (or entropy), the Simpson (1949) index ($S = 1 - \Sigma p^2$) is used frequently to measure demographic (or social) diversity in organizations. As created by Simpson, the unstandardized score yields the probability that 2 individuals chosen at random belong to the same (or different) group. According to Simpson (1949), the unstandardized index produces a distribution of scores ranging from 0 to 1. Actually, the unstandardized index has an EMV score based on following formula: $S_{Max} = 1 - \frac{1}{n}$ or $\frac{n-1}{n}$. When an unstandardized score is divided by the EMV for a demographic (or social) characteristics of interest, a biased standardized (or evenness) diversity score is obtained, indicating the proportion of the EMV that was attained (see Exhibit 1-1). Because the standardized scores represent the proportion of the EMV that is attained, the distribution of the scores ranges from 0 to 1, where a score of 1 indicates the complete attainment of the EMV. When treated as an evenness index, a biased standardized score indicates how well individuals are distributed amongst the groups. As such, a score of 0 indicates individuals are assembled in 1 group, and a score of 1 indicates individuals are distributed equally amongst the groups.

Because unstandardized and standardized Simpson indices do not adjust for the size of a community, group, or organization, Biemann and Kearney (2010) recommend the use of the following formula to obtain unbiased diversity scores (see Exhibit 1-1): $S_U = 1 - \sum \frac{n(n-1)}{N(N-1)}$. An unbiased

standardized score is obtained by dividing S_U by the EMV. Like a biased standardized Simpson score, an unbiased standardized score ranges from 0 to 1, where a score of 1 indicates the complete attainment of the EMV.

Despite its frequent use to assess demographic (or social) diversity in organizations, the construct and measurement validity of the Simpson index has not been assessed thoroughly. The measurement reliability of the Simpson index also has not been assessed thoroughly. For instance, Biemann and Kearney (2010) did not assess the construct or measurement validity of the biased and unbiased Simpson indices concomitantly or with other diversity indices discussed in their research. They also did not assess the measurement reliability of the diversity scores obtained with the biased and unbiased indices. Additionally, the differences amongst the simulated *biased* and *unbiased diversity scores* were not tested statistically for equality. While several studies have assessed the mathematical features of the Simpson index (Hussain and Khan, 2019; Smith and Wilson, 1996; Stirling, 2007), Mercil and Williams (1984) correlated the Simpson index with the McIntosh, Shannon, and other indices of variation. They also conducted a factor analysis with a principle components solution and *varimax rotation* to assess whether the indices measured the same dimension of variation. However, Mercil and Williams (1984) did not test specifically for construct or statistical validity or for measurement reliability.

Smith and Wilson (1996) developed several diversity indices based on the Simpson dominance index ($S_D = \Sigma p^2$; see Exhibit 1-1). The first index, D_{SW1}, is a restatement of the biased standardized Simpson index, which has a distribution of scores ranging from 0 to 1. Their second Simpson-based index, D_{SW2}, uses the reciprocal of Simpson dominance index which is divided by n, the number of categories (or groups). Like D_{SW1}, D_{SW2} has a distribution of scores ranging from 0 to 1. As standardized measures, D_{SW1} and D_{SW2} scores indicate the proportion of the EMV that is attained. When treated as evenness scores, they indicate how well individuals are distributed amongst the categories (or groups). For instance, a score of 1 indicates an equal distribution of individuals across each category.

As stated previously, Smith and Wilson (1996) compared 14 indices of evenness on 15 mathematical features and qualities. However, despite obtaining intercorrelation coefficients for the indices, Smith and Wilson did not test specifically for measurement validity. Unlike Mercil and Williams (1984), Smith and Wilson (1996) did not perform factor analyses to assess