

Development of a Scale for Measuring Perceptions of Trustworthiness for Digitized
Archival Documents

by

Devan Ray Donaldson

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Information)
in the University of Michigan
2015

Doctoral Committee:

Associate Professor Paul Conway, Chair
Sponsored Affiliate William G. Jacoby
Professor James M. Lepkowski
Associate Professor Soo Young Rieh
Professor Elizabeth Yakel

© Devan Ray Donaldson 2015

To Mom, Dad, Mimi, Papa, Gran, and Grandpa

ACKNOWLEDGEMENTS

I first want to thank Jesus for always loving, guiding, and protecting me. As the church folk say, “If it had not been for the Lord on my side,” I would never have made it through this process!

I would like to express my deepest gratitude to my adviser, Dr. Paul Conway. I can still remember meeting him at a NHPRC conference in Chapel Hill, NC. I knew then that I wanted to work with him. I am so happy that he took a chance on me and allowed me the privilege of being his student. Paul spent countless hours talking to me, exposing me to new ideas, helping me learn how to frame ideas into research problems, and teaching me a lot about writing. He gave me the confidence in the direction I wanted to take regarding my dissertation, assuring me that I would make a substantive and meaningful contribution to the field. And so many times, he helped me to see “the forest” when I could not see “the trees.”

I want to thank Dr. Elizabeth Yakel for being a mentor, colleague, and friend. I have learned so much from her because of her willingness to share her expertise with me and provide support. I really cannot imagine my experience at UMSI without her. I have taken classes with her, I have published journal articles with her, and she has been a part of every major milestone of my doctoral studies, from the first-year paper, to the field prelim, to the dissertation proposal, to the dissertation. Throughout my time as a student at Michigan, Beth has provided very fast, but also

very detailed feedback and constructive criticism to help me fine-tune my research skills. I don't know how she did it all, especially given all of her other responsibilities. Beth has deeply invested in my education and for that I am truly grateful. If I can become half the professor that she is, I will surely be very successful!

I want to thank Dr. Soo Young Rieh for sharing her wisdom and perspectives with me over the years. I took cues from her regarding how to conduct good research. I also got advice from her about steps to take regarding life after the dissertation. Soo is such a smart and meticulous researcher. I have always admired her work and appreciate all that she has done to help me with mine.

I also want to thank Dr. James Lepkowski. It is from him that I overcame my fear of statistics. It was also from him that I learned the difference between running statistics on datasets prepared for a class versus real-world data. It was so helpful to have an expert that I could go to who could help me understand that, if I got a result that was different from something that I learned in class, it didn't necessarily mean that I did something wrong. It didn't matter how many times I asked to meet with him or emailed him, he was always there. He also asked questions to help me more deeply understand my work and its implications for future research in my field. Thanks Jim, for pushing me and stretching my boundaries as a researcher.

I also want to thank Dr. William Jacoby. By the time I figured out that I wanted to build a scale for my dissertation, I had already completed all of my coursework. I seriously wondered if I would be able to do the project that I wanted to do, given what I wanted to accomplish with my dissertation. Bill made the

difference. When I asked him to serve on my dissertation committee, he said yes without any hesitation. Since that day, we were communicating via email constantly. In fact, I think we communicated via email on just about every major holiday during the year of 2014! I know how busy he was, so to receive that level of attention from him really meant a lot.

I want to thank Mr. Terence S. Badger, Deputy State Archivist of the Washington State Archives, for all of his help and support. Anything I ever asked him for, he always went above and beyond the call of duty to provide. I would not have a dissertation without him! I could never thank Terry enough. I would also like to thank the users of the Washington State Digital Archives (WADA) who graciously volunteered to participate in my study.

I want to thank Marcia McCrary, President of the Genealogical Society of Washtenaw County (GSWC), for her support of this project. Also, I would like to thank the members of the GSWC who volunteered their time to participate in my study.

I also want to thank Mrs. Darlene Ray-Johnson. She was there for me throughout my entire doctoral experience—all the ups and all the downs. Without her support, I truly don't know what I would have done. Words cannot express my gratitude for all she did for me.

I want to thank Dr. Jeff DeGraff for being my informal adviser. Whenever I called on him, I really needed him, and he was always there to provide a listening ear, substantive feedback, and thoughtful advice.

There are so many people to thank at the University of Michigan School of Information. I want to thank the administration of the school, past and present, including: Dean Jeff MacKie-Mason, Martha Pollack, Thomas Finholt, Stephanie Teasley, Paul Resnick, and Douglas Van Houweling. Also, I want to thank several UMSI professors, including: Margaret Hedstrom, Paul Edwards, Mark Ackerman, Yan Chen, Lionel Robert, Pedja Klasnja, Tiffany Veinot, and Tawanna Dillahunt. Also, I want to thank members of the UMSI staff, past and present, including: Mike Doa, Sue Schoun, Laura Elgas, Lorraine Robert, Heather Carpenter, Alissa Talley-Pixley, Judy Lawson, Veronica Falandino, Allison Wachter, Glenda Bullock, Dennis Hogan, and Jay Jackson.

My fellow doctoral students (past and present) at UMSI have been a constant source of support, including: Charles Senteio, Dharma Akmon, Ji-Yeon Yang, Ted Hanss, Adam Kriesberg, Carrie Xu, Tracy Liu, Karina Kervin, Jude Yew, Brian Hilligoss, Rayoung Yang, Eric Cook, Sangseok You, Andrea Barbarin, Rebecca Frank, Allen Flynn, Melissa Chalmers, Xin Rong, Morgan Daniels, Pablo Quinones, and Libby Hempbill.

I would also like to thank the Archives Research Group (ARG) at UMSI. The ARG helped me tremendously by providing feedback on my written work over the years.

I would also like to thank the Archival Education and Research Institute (AERI). AERI allowed me to share my research and get very helpful feedback throughout my time as a doctoral student. I gained so many benefits as a result of my participation in AERI over the years, including establishing contacts that led to

my collaboration with the Washington State Digital Archives (thanks Adam Jansen!) and correctly naming the instrument I built as a result of my dissertation research (thanks Dr. Heather Ryan!).

I would not have made it through this process without the very generous support of the Horace H. Rackham School of Graduate Studies at the University of Michigan. For example, I never would have been able to attend any conferences to disseminate my research, there are a couple of summers for which I would not have had any financial support, I would not have had the financial resources to conduct the research for this dissertation, and my health and general well-being would have been compromised were it not for the efforts and programming of the Rackham Graduate School. There are so many to thank at Rackham, past and present, including: Ellen Meader, Mark Kamimura-Jimenez, Emma Flores-Scott, Larry Rowley, Pat McCune, Debbie Willis, Debby Mitchell, and Doug Keasal. In particular, I would like to thank the Rackham Graduate School for their support of the summer institute. The summer institute provided me with a network of graduate students of color outside of my department. The graduate students I met as a result of the summer institute are my friends and sources of support to this very day, including: Karryll Winborne, Lori Hoggard, Jamein Cunningham, and Kim Lijana.

The University of Michigan administration, in particular Janet Weiss and Abigail Stewart, have been constant supporters and I really appreciate all they have done to help me as well as their service and support of the university.

I would like to thank my Students of Color of Rackham Family (SCOR) family, including: Leslie Upton, Portia Hempbill, Chantal Cotton-Gaines, Stacey Frazier, Elan Hope, Chris Tounsel, Timeka Tounsel, and Gbenga Olumolade.

I would also like to thank my Society of Minority Engineers and Scientists-Graduate Component (SMES-G) family, including: Clinique Brundidge, Iverson Bell, Taeyjuana Curry, Erica Morgan, Richard Ezike, Nancy Douyon, Afi Goss, Stan Goss, Kyla McMullen, Angelique Johnson, and Jorge Pernillo.

Towards the end of my dissertation, several people and organizations were instrumental to my completion, including the Center for the Education of Women, the Dissertation Writing Retreat program, founded by Dr. Wendy Carter-Veale, and my freshman year roommate from the College of William and Mary, Ryan Davidson, who read my dissertation from cover to cover! I also want to thank Dr. Ruth Person for taking time out of her busy schedule to read both my dissertation proposal (at the same time that she was Chancellor of the University of Michigan—Flint!) and my dissertation!

I also want to thank my family. All of my aunts, uncles, and cousins have told me how proud they are of me and my accomplishments, and I appreciate their well wishes and words of encouragement over the years. I want to thank my Brother and Sister-in-law for their love and support. My Dad has been awesome. Of course, he has provided support throughout my life, but he continued to go above and beyond in terms of supporting me during my doctoral program.

Finally, I want to thank my Mother for her love, guidance, and support throughout my entire life and during the doctoral program in particular. My Mother

always encouraged me and sacrificed for me so that I could have every opportunity that she did not have. She put my needs before hers at all times throughout my life. Her love for me is truly extraordinary. I can honestly say that I have never seen anyone love anyone else more than she loves me. I am very fortunate to have someone who cares so much about me and my future! Thank you Mom!

TABLE OF CONTENTS

DEDICATION.....	ii
ACKNOWLEDGEMENTS	iii
LIST OF TABLES.....	xviii
LIST OF FIGURES.....	xxi
LIST OF APPENDICES.....	xxii
ABSTRACT.....	xxiii
CHAPTER 1 INTRODUCTION.....	1
1.1 Definitions.....	5
1.2 Conceptual Framework	12
1.3 Motivations of the Study and Research Questions.....	13
1.4 Study Site in Brief	14
1.5 Research Design in Brief.....	15
1.6 Significance of the Study	17
1.7 Outline of the Chapters	19
1.8 References.....	21
CHAPTER 2 LITERATURE REVIEW	26
2.1 The Importance of Trustworthiness as a Concept in the Field of Digital Curation	27
2.1.1 What is a Trustworthy Digital Repository?.....	27

2.1.2	Chronology and Development of Trustworthy Digital Repositories	29
2.1.3	Standardizing Trustworthiness Through Criteria in Standards for Trustworthy Digital Repositories	34
2.1.4	Asserting Trustworthiness of Digital Repositories	35
2.2	Empirical Research on Trustworthiness Perception in the Field of Digital Curation	37
2.2.1	Repository Trustworthiness Perception.....	37
2.2.2	Document Trustworthiness Perception in the Context of a Digital Repository	39
2.2.3	Summary of Empirical Research on Trustworthiness Perception in the Field of Digital Curation	40
2.3	Trustworthiness in the Field of Web Credibility	40
2.3.1	Qualitative Research on Trustworthiness.....	41
2.3.2	Quantitative Research on Trustworthiness	42
2.3.3	Summary on Trustworthiness in the Field of Web Credibility.....	43
2.4	A Conceptual Framework for Trustworthiness Perception	44
2.5	Defining Trustworthiness.....	47
2.6	Conclusion	49
2.7	References.....	51
CHAPTER 3 METHODOLOGY: THE FOUR STEPS OF SCALE DEVELOPMENT.....		58
3.1	Selecting Scale Development.....	59
3.1.1	The First Approach	59
3.1.2	The Second Approach.....	61

3.2	The Methodology of Scale Development	63
3.2.1	Step 1 of Scale Development—Construct Definition	65
3.2.2	Step 2 of Scale Development—Generating an Item Pool	67
3.2.3	Step 3 of Scale Development—Designing the Scale	72
3.2.4	Step 4 of Scale Development—Full Administration and Item Analysis..	75
3.3	Ethical Considerations	80
3.4	References.....	81
CHAPTER 4 RESULTS OF STEPS 1 AND 2 OF SCALE DEVELOPMENT: FOCUS GROUPS STUDY		84
4.1	Study Participants	86
4.1.1	Size of the Focus Groups and Number of Focus Groups.....	89
4.1.2	Participant Recruitment.....	89
4.1.3	Locations	91
4.1.4	Demographic information	91
4.1.4.1	Participants by Age	92
4.1.4.2	Participants by Gender.....	92
4.1.5	Participants' WADA Document Usage.....	93
4.1.5.1	Primary Reason for Using WADA Documents.....	93
4.1.5.2	Frequency of Using WADA Documents.....	94
4.1.5.3	Frequency of Use by Document Type	96
4.1.5.4	Trust in WADA Documents	97
4.1.6	Study Participants Section Summary	98
4.2	Focus Groups Results: Identified Themes.....	98

4.2.1	Authenticity	102
4.2.2	Accuracy	104
4.2.3	Primary or First-hand Evidence.....	106
4.2.4	Believability	110
4.2.5	Validity	112
4.2.6	Proper Form	113
4.2.7	Stability.....	115
4.2.8	Legibility.....	117
4.2.9	Coverage.....	120
4.2.10	Objectivity.....	122
4.3	Conclusion	123
4.4	References.....	128
CHAPTER 5 RESULTS OF STEPS 3 AND 4 OF SCALE DEVELOPMENT: MEASURING TRUSTWORTHINESS		131
5.1	Study Participants	134
5.1.1	Participant Recruitment and Sampling Strategies.....	135
5.1.2	Sample Size	137
5.1.3	Demographic Information	138
5.1.3.1	Participants by Age	138
5.1.3.2	Participants by Gender.....	139
5.1.4	Participants' WADA Document Usage.....	140
5.1.4.1	Primary Reason for Using WADA Documents.....	140
5.1.4.2	Frequency of Using WADA Documents.....	141

5.1.4.3	Frequency of Use by Document Type	142
5.1.4.4	Time Spent Using Documents on a Typical Visit to WADA.....	144
5.1.5	Study Participants Section Summary	146
5.2	Initial Examination of Items' Performance	146
5.2.1	Items' Variances.....	147
5.2.2	Item-Total Correlations	150
5.2.3	Items' Means and Standard Deviations.....	151
5.2.4	Initial Evaluation of Items' Performance Section Summary	154
5.3	Exploratory Common Factor Analysis Results	155
5.3.1	Results of Tests for Appropriateness of the Data for EFA	157
5.3.1.1	Kaiser-Meyer-Olkin Measure of Sampling Adequacy Results	158
5.3.1.2	Bartlett's Test of Sphericity Results	158
5.3.1.3	Results of Tests for Appropriateness of the Data for EFA Summary	159
5.3.2	Results of Tests for Selecting the Number of Factors to Retain	159
5.3.2.1	Kaiser's Criterion Results.....	160
5.3.2.2	Parallel Analysis Results.....	161
5.3.2.3	Scree Test Results.....	162
5.3.2.4	Results of Tests for Selecting the Number of Factors to Retain Summary	164
5.3.3	Percentage of Variance Explained by the Factor	166
5.3.4	Factor Interpretation	168
5.3.4.1	Identification of Significant Loadings for Each Item Results.....	168

5.3.4.1.1	Items' Pattern Coefficients	169
5.3.4.1.2	Items' Structure Coefficients	171
5.3.4.2	Assessment of Items' Communalities Results	173
5.3.4.3	Factor Model Respecification	174
5.3.4.4	Factor Labeling Results.....	175
5.3.4.5	Factor Interpretation Summary	178
5.4	Reliability	179
5.4.1	Cronbach's Alpha Results	179
5.4.2	Item-Deleted Alpha Analysis Results.....	180
5.5	Validation Study	182
5.5.1	Research Questions	182
5.5.2	Study Participants.....	183
5.5.3	Study Details.....	185
5.5.3.1	Instrumentation.....	185
5.5.3.1.1	The Digitized Birth Certificate	185
5.5.3.1.2	The Digitized Marriage Certificate	189
5.5.4	Study Procedures	190
5.5.5	Results.....	191
5.6	Conclusion	195
5.7	References.....	197
CHAPTER 6	DISCUSSION & CONCLUSION	204
6.1	Trustworthiness is Unidimensional with Multiple Nuances.....	208
6.2	Trustworthiness in a TDR Context	213

6.2.1	Motivation for Revisions to Kelton et al.'s Specification of Trustworthiness	215
6.2.2	Revisions to Kelton et al.'s (2008) Specification of Trustworthiness....	217
6.2.2.1	Advantages and Disadvantages of Conceptual Trustworthiness Model	218
6.2.3	The Digitized Archival Document Trustworthiness Scale (DADTS)	219
6.2.3.1	Advantages and Disadvantages of Trustworthiness Measurement Model	221
6.2.3.2	Recommendations.....	222
6.3	The Digitized Archival Document Trustworthiness Scale in Relation to the Literature	222
6.3.1	The Best Performing Item: "The document is authentic"	223
6.3.2	The Second Best Performing Item: "The document is factual"	224
6.3.3	The Third Best Performing Item: "The document includes documentation of where it came from"	224
6.3.4	The Fourth Best Performing Item: "The document was created using responsible and accepted practices"	225
6.3.5	The Fifth Best Performing Item: "The digitized document is an actual picture of the original physical document"	226
6.3.6	The Sixth Best Performing Item: "The document is credible"	226
6.3.7	The Seventh Best Performing Item: "The document appears free from error"	227

6.3.8	The Eighth Best Performing Item: “The document is what it claims to be”	228
6.3.9	The Ninth Best Performing Item: “The document is a primary source”	228
6.3.10	The Tenth Best Performing Item: “The document accurately reflects what happened”	229
6.3.11	The Eleventh Best Performing Item: “The document is official”	229
6.3.12	The Twelfth Best Performing Item: “The document was written at the time of the event”	230
6.4	Implications: The Relationship Between Document and Repository Trustworthiness	230
6.5	Limitations and Delimitations of the Research	235
6.5.1	Limitations	235
6.5.1.1	Homogeneity of Study Participants	235
6.5.1.2	Sample Size	236
6.5.2	Delimitations	236
6.5.2.1	Actual versus Perceived Trustworthiness	237
6.5.2.2	Data Saturation	239
6.6	Recommendations for Future Research	241
6.7	Conclusion	243
6.8	References	245
APPENDICES		251

LIST OF TABLES

Table 3.1 Step 1 of Scale Development—Construct Definition.....	66
Table 3.2 Step 2 of Scale Development—Generating an Item Pool	68
Table 3.3 Step 3 of Scale Development—Designing the Scale.....	73
Table 3.4 Step 4 of Scale Development—Full Administration and Item Analysis.....	75
Table 4.1 Micro-interlocutor Analysis Results for Authenticity Theme.	103
Table 4.2 Micro-interlocutor Analysis Results for Accuracy Theme.....	105
Table 4.3 Micro-interlocutor Analysis Results for Primary or First-hand Evidence Theme.	107
Table 4.4 Micro-interlocutor Analysis Results for Believability Theme.....	111
Table 4.5 Micro-interlocutor Analysis Results for Validity Theme.....	112
Table 4.6 Micro-interlocutor Analysis Results for Proper Form Theme.	114
Table 4.7 Micro-interlocutor Analysis Results for Stability Theme.....	116
Table 4.8 Micro-interlocutor Analysis Results for Legibility Theme	118
Table 4.9 Micro-interlocutor Analysis Results for Coverage Theme.....	120
Table 4.10 Micro-interlocutor Analysis Results for Objectivity Theme.....	122
Table 4.11 Items Derived from the Focus Groups Study Findings.....	126
Table 5.1 Participants by Age.	138
Table 5.2 Participants by Gender.....	139
Table 5.3 Participants' Frequency of Using WADA Documents	141

Table 5.4 Participants' Frequency of Use by Document Type.....	143
Table 5.5 Participants' Time Spent Using Documents on a Typical Visit to WADA.	144
Table 5.6 Items' Recoded Minimums and Maximums	149
Table 5.7 Items' Corrected Item-Total Correlations.....	151
Table 5.8 Items' Means.....	153
Table 5.9 Items' Standard Deviations	154
Table 5.10 Factors' Initial Eigenvalues.....	160
Table 5.11 Parallel Analysis Results.....	162
Table 5.12 Eigenvalues and Percentages of Explained Variance for Analysis of an Exploratory 3-Factor Model of Document Trustworthiness Perception.....	167
Table 5.13 Items' Pattern Coefficients.....	171
Table 5.14 Items' Structure Coefficients.....	172
Table 5.15 Items' Final Communalities.....	174
Table 5.16 Items' Pattern Coefficients with Highest Loadings on the First Factor that are Also Greater than .65	177
Table 5.17 Items' Pattern Coefficients that also Loaded Strongly on the First Factor	178
Table 5.18 Item-Deleted Alpha Analysis Results Regarding the Scale Derived from the Trustworthiness Factor.....	181
Table 5.19 Items in the Digitized Archival Document Trustworthiness Scale (DADTS)	182
Table 5.20 Recoded Means of Validation Study Participants' Ratings for DADTS items (Birth Certificate)	192

Table 5.21 Recoded Means of Validation Study Participants' Ratings for DADTS items (Marriage Certificate).....	192
Table 5.22 Recoded Minimums and Maximums of Validation Study Participants' Ratings for DADTS items (Birth Certificate)	193
Table 5.23 Recoded Minimums and Maximums of Validation Study Participants' Ratings for DADTS items (Marriage Certificate)	194
Table 6.1 The Digitized Archival Document Trustworthiness Scale (DADTS).....	220
Appendix A Table 1. Summary of Revisions to Items as a Result of Cognitive interviewees' Feedback.....	252
Appendix A Table 2. Items Deleted Based on Cognitive Interviewees' Feedback....	253
Appendix A Table 3. Items Added Based on Cognitive Interviewees' Feedback.....	253
Appendix A Table 4. Items Revised Based on Cognitive Interviewees' Feedback....	254
Appendix A Table 5. Items Kept As Is After Cognitive Interviews.	255
Appendix I Table 1. List of Items Not Included in Exploratory Factor Analysis.....	291

LIST OF FIGURES

Figure 2.1 Kelton, Fleischmann, & Wallace's (2008) Integrated Model of Trust in Information.....	46
Figure 3.1 Scale Development Procedure	64
Figure 4.1 Contribution of the Focus Groups Study Findings to Scale Development.	86
Figure 5.1 Contribution of the Chapter 5 Study Findings to Scale Development.	132
Figure 5.2 Scree Plot of Factors.....	163
Figure 6.1 Conceptual Trustworthiness Model	218
Appendix H Figure 1. Histogram of Computed Scores for Factor 1	288
Appendix H Figure 2. Histogram of Computed Scores for Factor 2.....	288
Appendix H Figure 3. Histogram of Computed Scores for Factor 3.....	289

LIST OF APPENDICES

Appendix A Catalogue of Changes to Item Pool as a Result of Cognitive Interviews.....	252
Appendix B Sample Recruitment Email for Focus Groups.....	256
Appendix C Sample Web Survey Form for Recruiting Focus Group Participants	257
Appendix D Focus Groups Pre-survey	258
Appendix E Focus Groups Guide	261
Appendix F Codebook for Focus Groups Data	262
Appendix G Final Survey Used During Step 4 of Scale Development.....	264
Appendix H Outlier Detection Process	287
Appendix I List of Items Not Included in Exploratory Factor Analysis	290
Appendix J SPSS Syntax for Item Analysis and Factor Analysis	292
Appendix K SPSS Syntax for Parallel Analysis	293
Appendix L Validation Study Survey	300
Appendix M Digitized Birth Certificate.....	313
Appendix N Digitized Marriage Certificate	314

ABSTRACT

Trustworthiness is the most fundamental but least well understood property of digital repositories that hold and preserve archival documents. As these digital repositories scale in size and complexity, they are becoming essential sources for increasingly diverse populations of users. Scholarship across multiple disciplines has demonstrated that the trustworthiness of a digital repository tends to originate with organizational branding, surrounds and envelops the “control zone” of the managed digital space, and so resides primarily at the collective level of the repository. In spite of its conceptual centrality, little research has investigated trustworthiness of the documentary contents of repositories as conceived by the designated communities of users that the repository is intended to serve.

This dissertation investigates users’ perceptions of trustworthiness for archival documents housed in a large, heterogeneous, government-run digital repository. This dissertation utilizes the methodology of scale development, which involves four steps: 1) Construct Definition, 2) Generating an Item Pool, 3) Designing the Scale, and 4) Full Administration and Item Analysis. To address Steps 1 and 2 of scale development, I conducted a focus group study to elicit perspectives on trustworthiness and identify items for measurement of trustworthiness based upon actual users’ articulation of the concept; twenty-two genealogists who regularly utilize documents preserved by the Washington State Digital Archives

participated. To address Steps 3 and 4 of scale development, I conducted quantitative survey research and evaluated the responses of 233 genealogists, including constructing and testing an original Digitized Archival Document Trustworthiness Scale (DADTS). I also validated DADTS with a sample of users beyond the participants who were used to develop it. DADTS specifies the components of trustworthiness and also demonstrates the measurability of the concept within a digital repository context at the document level.

This dissertation advances scholarship on trustworthiness in three ways. First, it revises an existing conceptual model for trustworthiness perception. Second, it creates an original measurement model for digitized archival document trustworthiness perception—the Digitized Archival Document Trustworthiness Scale (DADTS). Third, it contributes to a deeper understanding of the concept of trustworthiness by providing measurement of the concept in a way that is sensitive to its nuances.

CHAPTER 1

INTRODUCTION

Trustworthiness is the most fundamental but perhaps least well understood property of digital repositories that hold and preserve archival documents. For at least fifteen years, digital curation researchers, information scientists, digital archivists, and computer scientists have worked successfully to design and construct robust, standards-oriented storehouses for digital archival documents. As these “Trustworthy Digital Repositories” (International Organization for Standardization, 2012) scale in size and complexity, they are becoming “sources of first resort” for increasingly diverse populations of users, ranging from scholars, students, government and corporate administrators, investigators from the private sector, genealogists, and the general curious public.

Scholarship across multiple disciplines has demonstrated that repository trustworthiness tends to originate with organizational branding, surrounds and envelops the “control zone” of the managed digital space (Atkinson, 1996), and so resides primarily at the collective level of the repository (Waters & Garrett, 1996). In spite of its conceptual centrality, little research has investigated the trustworthiness of the documentary contents of repositories as conceived by the

“Designated Communities” (Consultative Committee for Space Data Systems, 2002) of users that repositories are intended to serve.

Some digital curation researchers have defined trustworthiness at the document level based upon a foundation of archival ideas within the discipline of traditional archival science. For example, MacNeil (2000) explicitly defines trustworthiness in terms of reliability and authenticity:

A trustworthy record is one that is both an accurate statement of facts and a genuine manifestation of those facts. Record trustworthiness thus has two qualitative dimensions: reliability and authenticity. Reliability means that the record is capable of standing for the facts to which it attests, while authenticity means that the record is what it claims to be. (p. xi)

While a clear understanding of some digital curation researchers’ definitions of trustworthiness at the document level exists, users’ definitions of document trustworthiness are not known. For instance, users’ definitions of document trustworthiness could correspond to digital curation researchers’ conceptualizations of the term. On the whole, users’ definitions of document trustworthiness are elusive and underspecified.

Users’ definitions of document trustworthiness might correspond to concepts related to documents as described in international standards for digital repositories. For example, international standards for digital repositories link the concepts of authenticity and integrity at the document level to trustworthiness at the repository level (Data Seal of Approval Board, 2013; Deutsches Institut für Normung, 2012; International Organization for Standardization, 2012). Specifically,

these standards require repositories to address the authenticity and integrity of documents as part of certifying the repositories with “Trustworthy” status.

If in fact users define document trustworthiness in terms of some or all of the concepts that pertain to documents in the context of international standards for digital repositories, then we can begin to understand the relationship between document trustworthiness and repository trustworthiness. As an initial step toward this type of research, I focused on the concept of trustworthiness at the document level within a digital repository context from the vantage of user perception. Specifically, I sought to understand what the concept of trustworthiness meant to actual users of digitized archival documents that were preserved by a digital repository that, in theory, ought to be considered as trustworthy. I also sought to measure document trustworthiness based on users’ definition of the concept as empirical proof of its existence.

This study creates a new measure of user document trustworthiness perception—The Digitized Archival Document Trustworthiness Scale (DADTS) (See Table 6.1). The items in DADTS relate to concepts including authenticity, accuracy, reliability, and credibility, each of which has different traditions, definitions and research bases within the digital curation, information quality, and web credibility literatures. While none of these concepts are new, considering items that relate to all of these concepts together in a single scale to measure document trustworthiness as a variable is novel. I take this approach based upon overwhelming evidence that I gathered during this study that, even though researchers have drawn conceptual distinctions among the concepts of authenticity, accuracy, reliability, and credibility,

the distinctions do not hold from a statistical, empirical standpoint. Instead, the empirical evidence collected during this study demonstrates that trustworthiness should be measured using a single scale with items that represent the concepts of authenticity, accuracy, reliability, and credibility.

This study bridges approaches to trustworthiness in the digital curation, information quality, and web credibility research literatures. In particular, I draw upon approaches in information quality and web credibility to investigate the meaning and measurability of users' perceptions of trustworthiness for digitized archival documents housed in a large, heterogeneous, government-run digital repository. Even though empirical studies have examined users' perceptions of trustworthiness for digital repositories (e.g., Yakel, Faniel, Kriesberg, & Yoon, 2013; Yoon, 2014), users' perceptions of trustworthiness of the documents in repositories are not known. Examining the meaning and measurability of trustworthiness perception for digitized archival documents contributes to the body of literature on trustworthiness by providing a better understanding of:

- The concept of trustworthiness within a digital repository context, and
- The adaptability of trustworthiness as a measurable variable that might explain the use of documents in a repository.

The following sections present 1) definitions; 2) the conceptual framework; (3) motivations of the study and research questions; 4) the study site in brief; 5) the research design in brief; and 6) the significance of the study.

1.1 Definitions

Accuracy: The concept of accuracy applies in digital curation and information quality literatures. In the digital curation literature, accuracy has two definitions. The first definition refers to the factual accuracy of the information. This definition refers to the correctness and truthfulness of the information (Duranti & Preston, 2008). In this respect, documents that are created soon after the event that they are about are considered more accurate than documents that were created at a later time (Association for Information and Image Management, 1992).

The second definition of accuracy in the digital curation literature refers to the extent to which information is free from error or distortion, which corresponds to a more technical definition of accuracy (Duranti & Preston, 2008).

The goal of digital repositories is to ensure that documents do not become less accurate in a factual accuracy sense of the term, e.g., in presentation of facts, as a result of efforts to preserve that information—a more technical sense of the term. This aspect of accuracy refers to the concept of integrity (Hedstrom, 1995). In this respect, it is important to minimize error. What information a digital repository preserves may be completely comprised of falsehoods, but this aspect of accuracy is irrelevant, primarily because digital repositories do not create the intellectual content of documents; they are only responsible for preserving the intellectual content and making it accessible. Hence the primary concern of digital repositories is the technical definition of accuracy. On the other hand, users of documents found within digital repositories may be more concerned with factual accuracy and only

question the more technical definition of accuracy when a document appears suspicious to them.

In the information quality literature, researchers commonly position accuracy as a dimension of information quality (Knight, 2008; Lee, Strong, Kahn, & Wang, 2002; Rieh, 2002; Wang & Strong, 1996) For example, Knight (2008) found information quality researchers consistently listed accuracy as a dimension of information quality in thirteen out of twenty different information quality and data quality models.

In this study, trustworthiness is defined in part in terms of a user's perception of a digitized archival document as accurate. In particular, accuracy in this study can refer to both factual accuracy and technical accuracy.

Authenticity: According to Smith (2000), ““authenticity” in recorded information connotes precise, yet disparate, things in different contexts and communities. It can mean being original but also being faithful to an original; it can mean uncorrupted but also of clear and known provenance, “corrupt” or not. [...] Behind any definition of authenticity lie assumptions about the meaning and significance of content, fixity, consistency of reference, provenance, and context” (p. vi). This study borrows Duranti's (1995) conceptualization of authenticity, i.e., when a document is “what [it] purports to be,” for two reasons (p. 8). First, Duranti (1995) considers this particular conceptualization of authenticity as a component of document trustworthiness, which is essential to explore in a study of document trustworthiness within a digital archival context. Second, Duranti's (1995)

conceptualization of authenticity most closely aligns with the study participants' conceptualization of the term.

Data: In this study, data and documents are used interchangeably, borrowing from Buckland's (1991) conception of "information as thing," which casts data and documents similarly because they have the common feature of being informative.

Designated Community: According to the Consultative Committee for Space Data Systems (CCSDS) (2002), a designated community is "an identified group of potential Consumers who should be able to understand a particular set of information. The Designated Community may be composed of multiple user communities" (p. 1-10).

Digital Archive: See Digital Repository. Used synonymously with the term Digital Repository.

Digital Curation: According to Yakel (2007), "digital curation is the active involvement of information professionals in the management, including the preservation, of digital data for future use" (p. 335). In her review of official definitions of digital curation, she concludes that "digital curation is becoming the umbrella term for digital preservation, data curation, and digital asset and electronic records management" (p. 338).

Digital Repository: According to Heery and Anderson (2005), digital repositories share the following characteristics: (1) content is deposited in a repository, whether by the content creator, owner or third party; (2) the repository architecture manages content as well as metadata; (3) the repository offers a minimum set of basic services (e.g. put, get, search, access control); and (4) the repository must be sustainable and trusted, well-supported and well-managed (pp. 1-2).

Digitized Archival Document: For purposes of this study, a digitized archival document refers to a digitized genealogical record found in the Washington State Digital Archives, such as a digitized birth, death, or marriage record. What makes these documents archival is not just the fact that they are digitized images of non-current records, but that they are actively preserved within the context of a digital archive. See *Information*.

Document: Use of the term document refers to what Levy (1998) calls a digital document. According to him, digital documents “are split between an intangible digital object (which is ineffective outside of a complex, technical context) and a set of perceptible manifestations” (p. 154). Data and documents are used interchangeably in this study, borrowing from Buckland’s (1991) conception of “information as thing,” which casts data and documents similarly because they have the common feature of being informative.

End-user: One not responsible for the creation or preservation of the content they use (Donaldson, 2011; St. Jean, Rieh, Yakel, & Markey, 2011); the “pure consumer” of digitized archival documents.

In this study, end-users refer to people who perceive the trustworthiness of digitized archival documents that they find within the context of a digital repository. End-users neither create nor preserve digitized archival documents. They only use digitized archival documents in pursuit of their information needs. See *User*.

Information: According to Buckland (1991), “the term ‘information’ is ... used attributively for objects, such as data and documents, that are referred to as ‘information’ because they are regarded as being informative, as ‘having the quality of imparting knowledge or communicating information; instructive’” (p. 351). It is also important to note O’Toole’s (1989) definition of permanence. He argues for archivists to refocus their attention “on the permanence of the information in records rather than on the documents themselves” (p. 24). This perspective is important because the digitized images of archival documents in this study are not “the originals,” regardless of whether users perceive them as such. Thus, digitized archival documents in this study represent what O’Toole (1989) considers “the possibilities and usefulness of preserving information in formats other than the original” (p. 24).

Perception: “*The organization, identification, and interpretation of a sensation in order to form a mental representation*” (emphasis in original) (Schacter, Gilbert, &

Wenger, 2009, p. 123). In this study, perception refers to the key idea of user judgment.

Scale: A measurement instrument that is a collection of items “combined into a composite score and intended to reveal levels of theoretical variables not readily observable by direct means” (DeVellis, 2012, p. 11).

Trustworthy Digital Repository (TDR): According to the RLG/OCLC Working Group on Digital Archive Attributes (2002), a TDR is “one whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future” (p. 5). By this definition, some, but not all, digital archives, digital libraries, digital repositories, and institutional repositories apply (Garrett & Waters, 1996; Lynch, 2003; Ross, 2012; RLG/OCLC Working Group on Digital Archive Attributes, 2002; Wheatley, 2004).

Trustworthiness: In the digital curation literature, some researchers define trustworthiness in terms of reliability and authenticity. For example, Duranti (1995) defines the reliability of a record in terms of trustworthiness: “reliability refers to the authority and trustworthiness of the records as evidence, the ability to stand for the facts they are about” (p. 6). MacNeil (2000) explicitly defines record trustworthiness in terms of reliability and authenticity: “reliability means that the record is capable of standing for the facts to which it attests, while authenticity means that the record is what it claims to be” (p. xi). Similarly to MacNeil (2000),

Duff, Cherry, and Craig (2004) define the trustworthiness of archival sources in terms of reliability and authenticity, but employ different definitions for both concepts: reliability—“the degree to which the record accurately reflects what happened”; authenticity—“a record has not been altered or changed since its original creation” (p. 67). Overall, some digital curation researchers’ definitions of trustworthiness at the document level correspond to concepts of reliability and authenticity, if they define trustworthiness at all.

In this study, trustworthiness refers to a user’s perception of four concepts, authenticity, accuracy, reliability, and credibility. It includes specific definitions of authenticity that refer to the extent to which a digitized archival document is what it claims to be (Duranti, 1995). It also includes specific definitions relating to the factual and technical accuracy of the content (Association for Information and Image Management, 1992). It also includes specific definitions of reliability that refer to appropriate procedures of creation and form of the digitized archival documents (Duranti, 1995).

In this study, trustworthiness is a unidimensional concept from a statistical, empirical standpoint. Items pertaining to the concepts of authenticity, accuracy, reliability, and credibility come together to formulate this one dimension: trustworthiness. This study highlights the nuances of the dimension of trustworthiness, in particular, the concepts that contribute to it from the vantage of user perception within a digital repository context.

User: Used interchangeably with end-user; one not responsible for the creation or preservation of the content they use (Donaldson, 2011; St. Jean, Rieh, Yakel, & Markey, 2011); the “pure consumer” of digital archival documents. See *end-user*.

1.2 Conceptual Framework

The conceptual framework for this study is Kelton, Fleischmann, and Wallace’s (2008) “Integrated Model of Trust in Information.” This dissertation selects Kelton et al.’s (2008) model, because, while other models engage the concept of trustworthiness from the user’s point of view (e.g., Hilligoss & Rieh, 2008; Lucassen & Schraagen, 2012; Pirson & Malhotra, 2011), Kelton et al. (2008) provide the largest framework for trustworthiness perception specifically at the document or information level. For example, according to Kelton et al., trustworthiness is the extent to which a user perceives information as accurate, current, complete, believable, objective, valid, and stable. In this respect, Kelton et al.’s framework provides a point of departure for study of the concept of trustworthiness.

While promising, Kelton et al.’s framework, in particular its specification of trustworthiness, has not yet been empirically tested within a digital repository context. Thus, which parts of Kelton et al.’s specification of trustworthiness correspond to users of digitized archival documents and their trustworthiness perceptions is an open question.

1.3 Motivations of the Study and Research Questions

The problem with the Kelton model, one that motivates the research reported here, concerns the specification of trustworthiness. The model maps judgments of trustworthiness to no fewer than four complex and somewhat conflated concepts under the term “accuracy,” including accuracy itself, believability, coverage, and currency. These terms are but a small subset of the many characteristics of information sources covered in the information quality literature (Knight, 2008). Scholars across a number of associated disciplines also have wrestled with how to operationalize the components of this abstract notion. Any measure of trustworthiness should seek to reconcile this diverse research, if for no other reason than increasing the usefulness of a trustworthiness measure.

This study builds a scale of trustworthiness because:

1. A scale is a very strong, defensible, and comprehensive mechanism for measuring a phenomenon (Jacoby, 1991), and
2. The existing research literature asserts trustworthiness or defines it within the context of a research instrument, but the literature does not provide instrumentation that actually measures the phenomenon within a digital preservation context.

To explore these issues, the study addresses the following research questions:

- Research Question 1: How do members of a designated community conceptualize trustworthiness for documents they find in a digital repository?

- Research Question 2: To what extent are designated community members' perceptions of document trustworthiness measurable?

1.4 Study Site in Brief

This study centers on users' perceptions of trustworthiness for digital archival documents. In order to conduct the investigation, the primary site for the study is the Washington State Digital Archives (WADA). WADA is a state-of-the-art digital repository; it is a heavily utilized digital cultural heritage resource, developed and maintained at taxpayer expense as a mechanism for providing open, public access to archives and records of the State of Washington.

WADA's collections are diverse in terms of document type and file format. For example, WADA's collections include audio recordings of city council meeting minutes in mp3 file format as well as annual reports of local library systems in pdf file format. WADA preserves surveyor plat maps accessible in jpeg format. WADA also includes records of value for conducting genealogical research, including digitized marriage, death, and birth records in jpeg file format.

WADA also has diversity in terms of its designated communities. Users of WADA include, but are not limited to: genealogists, local and state government employees, title company employees, historians, and primary, secondary, and post-secondary educators (International Organization for Standardization, 2012; T. S. Badger, personal communication, March 8, 2013).

Approximately 500,000 people visit the home page of WADA per year with thousands of unique visitors per month. WADA has a strong and explicit mission statement that focuses on making preserved digital information accessible to users (Washington State Archives – Digital Archives, 2015).

Importantly, WADA conforms in principle to the requirements of a TDR. It abides by leading best practices and standards for organizational infrastructure, digital object management, and technical infrastructure, including security issues, consistent with the International Organization for Standardization’s specifications, despite not being formally certified as a TDR as of the time of this study (T. S. Badger, personal communication, March 8, 2013).

1.5 Research Design in Brief

I executed the methodology of scale development during this study. Scale development involves four steps: 1) Construct Definition, 2) Generating an Item Pool, 3) Designing the Scale, and 4) Full Administration and Item Analysis. I rigorously applied the methodology of scale development, using multiple methods to execute each step.

This study focuses on one specific designated community of users—genealogists—because, according to WADA staff, genealogists comprise WADA’s largest designated community (T. S. Badger, personal communication, March 8, 2013). The study also focuses on the designated community members’ experiences with digitized marriage, death, birth, census, and land records, as they are among

the most heavily downloaded documents according to WADA staff (T. S. Badger, personal communication, March 8, 2013).

I created two samples of the designated community for the study. Both samples were similar in terms of demographic characteristics; they consisted of mostly older, Caucasian females. The main difference between the samples was their size. The sample for the focus groups study contained twenty-two participants; the sample for the item analysis study contained 233 participants.

I developed a focus group guide that I reviewed with two experts on focus groups methodology for clarity and appropriateness. The focus group guide included questions concerning the nature and purposes of the documents participants use; user perceptions of trustworthiness for the documents; examples of documents that participants perceived as trustworthy; and any circumstances for questioning the trustworthiness of documents in WADA. The focus groups lasted approximately two hours. I conducted the focus groups to inform development of items for a scale to measure trustworthiness perception.

Building upon the focus groups study findings, I developed a scale. I generated a large pool of items for measurement of trustworthiness from a variety of sources, including the literature, trustworthiness subject matter experts, and actual users of digitized archival documents (i.e., the focus groups study participants). Each item described a circumstance one might encounter while using a digitized archival document. I included the items in a web-administered survey. Participants answered whether the circumstance described by each item would cause them to perceive a digitized archival document as either untrustworthy or

trustworthy on a 7-point scale: very untrustworthy, untrustworthy, slightly untrustworthy, neither untrustworthy nor trustworthy, slightly trustworthy, trustworthy, or very trustworthy. I also included an eighth option, 99 or “Not Applicable,” for participants to choose if the circumstance an item described was not relevant to their experience of using digitized archival documents. The best performing items constituted a scale for measurement of trustworthiness. For validation, I administered the scale to a different sample of users who also had experience utilizing digitized archival documents. These methods are described in greater detail in Chapter 3 and their implementation are discussed in Chapters 4 and 5.

1.6 Significance of the Study

The digital curation community has long acknowledged that a repository’s claim for the basic capacity to preserve digital materials is insufficient grounds for trustworthiness; repositories must demonstrate they are trustworthy. As a result, various working groups have formed over the years to create definitions of what a TDR is and ought to be able to do. The digital curation community has even come up with standards and best practices that, if adhered to, can permit repository managers to assert that their repositories are trustworthy under certain conditions. However important these efforts may be, they do not constitute measurement of trustworthiness. Measuring trustworthiness is important because the concept is at the heart of the justification for TDRs, but the concept has only been defined as a

concept, and not in a way that is amenable to verifying the presence or absence of the concept.

This dissertation drills deeply into the notion of repository trustworthiness by taking the view that the perception of trustworthiness in individual documents in a TDR might be a variable affecting repository trustworthiness. Toward this end, this dissertation is the first empirical study to measure the trustworthiness perceptions of users of digital documents within a digital repository context. This study assumes that what users think of repositories and their content is at least as important as what repository managers or third party auditors think. The study is significant because it tackled one of the most important questions in the digital curation research domain (when is a digital repository trustworthy?) by:

- Exploring what trustworthiness means to actual users,
- Focusing on trustworthiness perception at the level where users actually interact with the repository—the document level, and
- Measuring trustworthiness.

The study provides clarity on the ability to identify trustworthiness perception at the document level within a TDR context. It demonstrates that trustworthiness is more than just an abstract concept. It has meaning for the users of digital repositories and the content that they contain.

1.7 Outline of the Chapters

The dissertation proceeds from the argument that trustworthiness is a fundamental principle in digital preservation, yet not well understood from the user's point of view. This discussion begins in Chapter 2, which outlines scholarship on trustworthiness in the fields of digital preservation, digital curation, and web credibility to examine the extent of our understanding regarding trustworthiness. Chapter 2 (section 2.5) also discusses the most relevant literature related to "Construct Definition," which is the first of four steps in scale development. Chapter 3 describes the methodology of scale development. It describes the four steps of scale development: 1) Construct Definition, 2) Generating an Item Pool, 3) Designing the Scale, and 4) Full Administration and Item Analysis. I take this methodological approach to provide a fuller understanding of trustworthiness within a digital repository context at the document level.

I describe the bulk of the study's findings in Chapters 4 and 5. These chapters cover users' perceptions of trustworthiness for digitized archival documents. These chapters also shed light on the measurability of these users' perceptions.

Chapter 4 provides in-depth analysis of qualitative findings from focus group interviews with genealogists of the Washington State Digital Archives (WADA). These findings demonstrate the range of designated community members' perceptions of trustworthiness for digitized archival documents, but also pinpoint particular perceptions of trustworthiness that are most important to the

participants. These findings address Steps 1 and 2 of scale development: 1) Construct Definition, and 2) Generating an Item Pool.

Chapter 5 is the crux of the dissertation. The chapter focuses on the results of executing Steps 3 and 4 of scale development: 3) Designing the Scale, and 4) Full Administration and Item Analysis. First, Chapter 5 discusses the analysis of data resulting from participants' responses after administering a pool of items to participants regarding the concept of trustworthiness. The results come from analysis of the quantitative data of 233 genealogists who rated trustworthiness items in terms of importance. The data analysis, including item analysis and exploratory factor analysis, identify items for a scale to measure trustworthiness perception.

Chapter 5 also includes findings of a separate validation study to assess the validity of the scale resulting from analysis of the 233 participants' responses to other genealogists besides those whose responses I used to build the scale. The validation study also demonstrates the ability of the scale to apply to measurement of trustworthiness of specific digitized archival documents.

The dissertation concludes in Chapter 6, where I discuss the implications of the findings for understanding:

- Trustworthiness as a concept
- Trustworthiness in a TDR context, and
- Document versus repository trustworthiness.

1.8 References

- Association for Information and Image Management. (1992). *Performance guideline for the legal acceptance of records produced by information technology systems*. Silver Spring, MD: The Association for Information and Image Management.
- Atkinson, R. (1996). Library functions, scholarly communication, and the foundation of the digital library: Laying claim to the control zone. *The Library Quarterly*, 66(3), 239-265.
- Buckland, M. K. (1991). Information as thing. *Journal of the American Society for Information Science*, 42(5), 351-360.
- Consultative Committee for Space Data Systems. (2002). Reference model for an open archival information system (OAIS) recommendation for space data system standards ; blue book (Blue Book. Washington, DC: CCSDS Secretariat. doi:<http://public.ccsds.org/publications/archive/650x0b1s.pdf>
- Deutsches Institut für Normung. (2012). *Information and documentation—Criteria for trustworthy digital archives*. Berlin: Deutsches Institut für Normung.
- DeVellis, R. F. (2012). *Scale development: Theory and applications*. Thousand Oaks, Calif.: SAGE.
- Donaldson, D. R. (2011). Users' trust in trusted digital repository content. 8th International Conference on Preservation of Digital Objects (iPres 2011), Singapore. pp. 20-23.
- Duff, W., Craig, B., & Cherry, J. (2004). Finding and using archival resources: A cross-canada survey of historians studying canadian history. *Archivaria*, 58, 51-80.

- Duranti, L. (1995). Reliability and authenticity: The concepts and their implications. *Archivaria*, 39, 5-10.
- Hedstrom, M. (1995). Electronic archives: Integrity and access in the network environment. *The American Archivist*, 58(3), 312-324.
- Heery, R., & Anderson, S. (2005). Digital repositories review. Joint Information Systems Committee.
- Hilligoss, B., & Rieh, S. Y. (2008). Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context. *Information Processing & Management*, 44(4), 1467-1484.
- International Organization for Standardization. (2012). Space data and information transfer systems : Audit and certification of trustworthy digital repositories [Systèmes de transfert des informations et données spatiales : audit et certification des référentiels numériques de confiance] (1st ed.). Geneva, Switzerland: International Organization for Standardization.
- Jacoby, W. G. (1991). Data theory and dimensional analysis. Newbury Park, Calif.: Sage Publications.
- Kelton, K., Fleischmann, K. R., & Wallace, W. A. (2008). Trust in digital information. *Journal of the American Society for Information Science and Technology*, 59(3), 363-374.
- Knight, S. (2008). User perceptions of information quality in world wide web information retrieval behaviour. (Doctoral Dissertation, Edith Cowan University), 1-362.

- Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). AIMQ: A methodology for information quality assessment. *Information & Management*, 40(2), 133-146.
- Levy, D. M. (1998). Heroic measures: Reflections on the possibility and purpose of digital preservation. Proceedings of the Third ACM Conference on Digital Libraries, Pittsburgh, Pennsylvania, USA. pp. 152-161.
- Lucassen, T., & Schraagen, J. M. (2012). Propensity to trust and the influence of source and medium cues in credibility evaluation. *Journal of Information Science*, 38(6), 566-577.
- Lynch, C. A. (2003). Institutional repositories: Essential infrastructure for scholarship in the digital age. *Portal: Libraries and the Academy*, 3(2), 327-336.
- MacNeil, H. (2000). *Trusting records: Legal, historical and diplomatic perspectives*. Dordrecht; London: Kluwer Academic.
- O'Toole, J. M. (1989). On the idea of permanence. *The American Archivist*, 52(1), 10-25.
- Pirson, M., & Malhotra, D. (2011). Foundations of Organizational Trust. *Organization Science*, 22(4), 1087-1104.
- Rieh, S. Y. (2002). Judgment of information quality and cognitive authority in the web. *Journal of the American Society for Information Science and Technology*, 53(2), 145-161.

- RLG/OCLC Working Group on Digital Archive Attributes. (2002). Trusted digital repositories: Attributes and responsibilities. Mountain View, CA: RLG.
- Retrieved from <http://www.oclc.org.proxy.lib.umich.edu/research/activities/past/rlg/trustedrep/repositories.pdf>;
- Ross, S. (2012). Digital preservation, archival science and methodological foundations for digital libraries. *New Review of Information Networking*, 17(1), 43-68.
- Schacter, D. L., Gilbert, D. T., & Wegner, D. M. (2009). *Introducing psychology*. Macmillan.
- Smith, A. (2000). Introduction. *Authenticity in a digital environment* (pp. vi-viii). Washington D.C.: Council on Library and Information Resources. Retrieved from <http://clir.org/pubs/reports/pub92/pub92.pdf>
- St. Jean, B., Rieh, S. Y., Yakel, E., & Markey, K. (2011). Unheard voices: Institutional repository end-users. *College & Research Libraries*, 72(1), 21-42.
- Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*, 12(4), 5-33.
- Washington State Archives - Digital Archives. (2015). Background and history. Retrieved 1 23, 2015, from <http://www.digitalarchives.wa.gov/StaticContent/background>

- Waters, D., & Garrett, J. (1996). Preserving digital information : Report of the task force on archiving of digital information. Washington, D.C.: Council on Library and information Resources. Retrieved from <http://www.clir.org/pubs/reports/pub63watersgarrett.pdf>
- Wheatley, P. (2004). Technology watch report: Institutional repositories in the context of digital preservation No. 04-02)Digital Preservation Coalition. Retrieved from <http://www.dpconline.org/docs/DPCTWf4word.pdf>
- Yakel, E. (2007). Digital curation. *OCLC Systems & Services*, 23(4; 1065-075), 335-340.
- Yakel, E., Faniel, I., Kriesberg, A., & Yoon, A. (2013). Trust in digital repositories. *International Journal of Digital Curation*, 8(1), 143-156.
- Yoon, A. (2014). End users' trust in data repositories: Definition and influences on trust development. *Archival Science*, 14(1), 17-34.

CHAPTER 2

LITERATURE REVIEW

To lay the groundwork for an in-depth study of user perceptions of trustworthiness for digitized archival documents found within a digital repository context, this chapter presents scholarship on trustworthiness. First, this chapter explores literature regarding Trustworthy Digital Repositories (TDRs), including discussion of the origins and development of the concept of a TDR as well as international standards for digital repositories. Second, the chapter discusses prior empirical research on trustworthiness perception at both the repository and document levels within the context of a digital repository. Third, as this dissertation is particularly focused on measurement of trustworthiness at the document level, this chapter also expounds selected studies in the research domain of web credibility, where researchers have built a tradition of measuring user trustworthiness perception at the document or information level. Finally, this chapter discusses an existing conceptual model for trustworthiness perception, and concludes by examining definitions of trustworthiness at the document level that are relevant to the current study.

2.1 The Importance of Trustworthiness as a Concept in the Field of Digital Curation

For nearly two decades, one of the biggest questions in the field of digital curation has been: “When is a digital repository trustworthy?” Emphasis on this question arises from the fact that digital materials do not preserve themselves; they require active preservation and management to enable future use (Lee & Tibbo, 2007). The literature on Trustworthy Digital Repositories (TDRs) best illustrates discussion of the concept of trustworthiness within the field of digital curation. For this reason, Section 2.1.1 describes the notion of a Trustworthy Digital Repository (TDR), Section 2.1.2 catalogues the chronology and development of TDRs, Section 2.1.3 focuses on the international standards for TDRs, and Section 2.1.4 explains the assertion of trustworthiness inherent in international standards for TDRs.

2.1.1 What is a Trustworthy Digital Repository?

Members of the digital curation community have produced several definitions of a TDR based upon what attributes it should have and what tasks it should be able to accomplish. Even though the creators of these definitions may not use the specific term “Trustworthy Digital Repository,” analysis of the definitions suggests that those who use the term “Digital Repository” or even “Open Archival Information System” are often referring to the same type of repository. In fact, close

inspection of definitions for TDRs and digital repositories reveals that they draw heavily upon the definition of an Open Archival Information System (OAIS).

The Consultative Committee for Space Data Systems (CCSDS) (2002) defines an Open Archival Information System (OAIS) as “an archive, consisting of an organization of people and systems, that has accepted the responsibility to preserve information and make it available for a Designated Community” (p. 1-11). Similarly, The RLG/OCLC Working Group on Digital Archive Attributes (WGDAAs) (2002) defines a TDR as “one whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future” (p. 5). The NESTOR Working Group on Trusted Repositories Certification (NWGTRC) (2006) defines a digital repository as “as an organisation [sic] (consisting of both people and technical systems) that has assumed responsibility for the long-term preservation and long-term accessibility of digital objects, ensuring their usability by a specified target group, or ‘designated community’” (p. 2). Three years later, the NWGTRC (2009) slightly modifies its definition of a digital repository as “an organisation [sic] (consisting of people and technical systems) which has assumed responsibility for the long-term preservation and long-term accessibility of digital objects, and also for their interpretability, for the purpose of their being used by a specific designated community” to focus on interpretability rather than usability (p. 4). With both the 2006 and 2009 definitions of a digital repository, the NWGTRC acknowledges the definitions are based on the CCSDS’s (2002) definition of an OAIS. The RLG/NARA Digital Repository Certification Task Force’s (DRCTF) (2007)

discussion of a TDR expands prior definitions by stating that “a trusted digital repository will understand threats to and risks within its systems” (p. 3).

Dow (2009) provides perhaps the most comprehensive definition of a TDR:

a sustainable, trustworthy, well-supported, and well-managed digital repository needs hardware, software, policies, processes, services, and people to assure long-term retention and, perhaps, access to its content and metadata. It needs the ability to integrate a significant number of document and file types created through word processors, spreadsheets, e-mail systems, database management systems, website systems, imaging systems, computer-assisted design/computer-assisted publishing systems, and a host of other media, known and unknown, creating both structured and unstructured data. (p. 81)

Given the definitions of what a TDR, it seems that answering the question “what is a TDR?” turns on the strictness of one’s interpretation.

2.1.2 Chronology and Development of Trustworthy Digital Repositories

In the seminal report *Preserving Digital Information*, Waters and Garrett (1996) mark the dawn of the modern era for digital repositories. They are the first authors to articulate the need for Trustworthy Digital Repositories (TDRs). They motivate the need for TDRs by raising two key issues. First, they cite examples of organizations with proven track records of successfully preserving analog materials and yet failing to preserve digital information. In some of these cases, access to important digital cultural heritage information was lost forever, and in others, nothing short of very expensive and risky digital archaeology efforts were necessary

to regain access to the digital information. Second, they acknowledge the fact that the skills and technologies necessary for preserving digital information are much different and more complex than those needed to support analog preservation. Using both of these motivations, Waters and Garrett underscore the fact that a proven track record for preserving analog materials is not sufficient grounds for trusting those same organizations with the responsibility of preserving digital information. Essentially, any organization assuming responsibility for preserving digital information must be able to prove that they actually have the ability to preserve it. Toward this end, they propose the establishment of a formal, third-party certification process for digital archives to assure users, “that a digital archives is what it says that it is and that the information stored there is safe for the long term,” and “establish an overall climate of value and trust about the prospects of preserving digital information” (Waters & Garrett, 1996, pp. 23-24).

One of the most critical and influential parts of the Waters and Garrett (1996) report is the section on digital information integrity. First, the authors define integrity as “those features of an information object that distinguish it as a whole and singular work” (Waters & Garrett, 1996, p. 12). Second, the authors identify and define five specific features of integrity: content, context, fixity, provenance, and reference. And finally, the authors argue that what distinguishes a TDR from other, less capable repositories is its ability to preserve the integrity of its holdings.

Building on Waters and Garrett (1996), the Consultative Committee for Space Data Systems (CCSDS) (2002) developed a reference model for an Open Archival Information System (OAIS). Specifically, the CCSDS (2002) use Waters and Garrett

(1996) typology of information integrity (i.e., content, context, fixity, provenance, and reference information) as part of the OAIS model's Preservation Description Information (CCSDS, 2002, p. 4-34; p. B-1). As a whole, the OAIS model represents the outcome of efforts "to create a consensus on what is required for an archive to provide permanent, or indefinite long-term preservation of digital information" (CCSDS, 2001 quoted in RLG/OCLC WGDA, 2002, p. 2). The thoughtfulness of the model and the broad range of players involved in its development led to its robustness and helped to further shape notions of what a TDR should be at a high level. According to the RLG/NARA Digital Repository Certification Task Force (DRCTF) (2007), repositories were declaring themselves as OAIS-compliant as an attestation of their trustworthiness without providing any evidence; thus, declaring trustworthiness as OAIS-compliance was too open to interpretation and not measureable.

The RLG/OCLC Working Group on Digital Archive Attributes (WGDA) (2002) contributes to the development of a TDR by providing a definition for it. According to the WGDA, a TDR is "one whose mission is to provide reliable, long-term access to managed digital resources to its designated community, now and in the future" (p. 5). The WGDA adds that TDRs "may take different forms: some institutions may choose to build local repositories while others may choose to manage the logical and intellectual aspects of a repository while contracting with a third-party provider for its storage and maintenance" (p. 5). Regardless of what form TDRs take, the WGDA specifies that all TDRs must (p. 5):

- Accept responsibility for the long-term maintenance of digital resources on behalf of its depositors and for the benefit of current and future users;
- Have an organizational system that supports not only long-term viability of the repository, but also the digital information for which it has responsibility;
- Demonstrate fiscal responsibility and sustainability;
- Design its system(s) in accordance with commonly accepted conventions and standards to ensure the ongoing management, access, and security of materials deposited within it;
- Establish methodologies for system evaluation that meet community expectations of trustworthiness;
- Be depended upon to carry out its long-term responsibilities to depositors and users openly and explicitly;
- Have policies, practices, and performance that can be audited and measured; and
- Meet several high-level organizational and curatorial responsibilities as well as operational responsibilities.

Essentially, the WGDAA advances the notion of a TDR by providing a definition, including attributes and responsibilities, at a more precise level than before.

Taking the next step forward, the RLG/NARA Digital Repository Certification Task Force (DRCTF) (2007) operationalizes the WGDAA's definition of a TDR as a

set of criteria for audit and certification (better known as Trustworthy Repositories Audit and Certification or TRAC). Specifically, the DRCTF includes criteria for organizational infrastructure, digital object management, and technical infrastructure, including security. According to the DRCTF, repository administrators should use its criteria as a guide to determine the extent to which their repositories adhere to the standards and best practices of the digital preservation community. Further, repository administrators can use the criteria for purposes of self-audit and third parties can also use the criteria as a metric for determining whether a given repository should or should not be certified as trustworthy. In sum, TRAC was designed to operationalize the OAIS model and move toward the creation of metrics for establishing and certifying the trustworthiness of digital repositories.

Building upon DRCTF (2007), the CCSDS (2011) created an international standard for TDRs – ISO 16363 (International Organization for Standardization, 2012). According to the Center for Research Libraries (CRL), CCSDS (2011) is a revision of DRCTF (2007). Since the creation of the DRCTF’s audit and certification checklist, the CRL is one body that has assumed responsibility for third-party certification of TDRs using the DRCTF’s audit and certification checklist as a guide. As of January 2015, the CRL has formally certified five repositories—Chronopolis, HathiTrust, Scholars Portal, CLOCKSS, and Portico (Center for Research Libraries, n.d.).

Most recently, the International Organization for Standardization established the Primary Trustworthy Digital Repository Authorisation Body (ISO-PTAB) to

handle training for auditors of digital repositories and prepare repository managers for audit (PTAB, n.d.). Specifically, to facilitate audit of digital repositories, ISO-PTAB relies upon *Requirements for Bodies Providing Audit and Certification of Candidate Trustworthy Digital Repositories* (ISO 16919), which specifies the process for audit (International Organization for Standardization, 2014).

2.1.3 Standardizing Trustworthiness Through Criteria in Standards for Trustworthy Digital Repositories

For over a decade, the digital curation community has worked to build consensus regarding what is necessary for digital repositories to be considered trustworthy. The net result of these efforts is recent standards for Trustworthy Digital Repositories (TDRs). Examples of these standards and guidelines include *Data Seal of Approval* (Data Seal of Approval Board, 2013), *Information and Documentation – Criteria for Trustworthy Digital Archives* (DIN 31644) (Deutsches Institut für Normung, 2012), and *Audit and Certification of Trustworthy Digital Repositories* (ISO 16363) (International Organization for Standardization, 2012). These standards and guidelines specify similar requirements for repositories in terms of their organizational infrastructure, digital object management, technical infrastructure, and security (Data Seal of Approval Board, 2013; Deutsches Institut für Normung, 2012; International Organization for Standardization, 2012). If these requirements are met, digital repository managers can assert that their repositories are trustworthy.

Of particular importance to my study are concepts that apply to documents that digital repositories are responsible for preserving in order to achieve and sustain trustworthy status. In particular, concepts of understandability, interpretability, usability, authenticity, and integrity at the document level are directly related to criteria in standards for Trustworthy Digital Repositories (Data Seal of Approval Board, 2013; Deutsches Institut für Normung, 2012; International Organization for Standardization, 2012). The fact that these concepts relate to criteria in standards for digital repositories links considerations of documents housed within digital repositories to trustworthiness at the repository level.

If users define document trustworthiness in terms of some or all of the concepts that standards for digital repositories employ with regard to documents housed within digital repositories, then user perception of document trustworthiness would link to repository trustworthiness as defined in international standards for digital repositories. What international standards for digital repositories lack is the incorporation of an understanding of trustworthiness from the user's perspective. To address this issue, my study investigates users' perception of trustworthiness at the document level, but starts with trying to understand users' definitions of document trustworthiness perception.

2.1.4 Asserting Trustworthiness of Digital Repositories

Examining the titles of significant international standards for digital repositories demonstrates that, by and large, the digital curation community

conceptualizes trustworthiness as a property that repository managers can assert over their repositories if they meet certain criteria. The naming conventions of these standards for digital repositories include the term “trustworthy.” For example, the international standard *Audit and Certification of Trustworthy Digital Repositories* (ISO 16363) includes “trustworthy” in the name of the standard (International Organization for Standardization, 2012). As another example, the international standard *Information and Documentation – Criteria for Trustworthy Digital Archives* (DIN 31644) also includes “trustworthy” in its name (Deutsches Institut für Normung, 2012).

Overall, recent standards for digital repositories are very similar in the criteria they employ. These standards represent what members of the international digital curation community consider necessary for a digital repository to be considered trustworthy. They reflect objective measures for determining digital repositories’ trustworthiness. While users, or designated communities members, are actively mentioned across these standards, little is known about users’ perspectives on repository trustworthiness in the context of these standards. Recent empirical studies on digital repository trustworthiness perception shed light on this topic.

2.2 Empirical Research on Trustworthiness Perception in the Field of Digital Curation

Empirical studies aimed at understanding trustworthiness from the perspective of user perception tend to focus on experienced users of different types of repositories, including science data repositories, social science data repositories, institutional repositories, and digital repositories which house archival collections of digitized documents and photographs. The researchers typically collected data from their study participants using semi-structured interviews, in which they asked participants questions about trustworthiness. The questions required participants to think retrospectively about their experiences. These studies tend to focus on trustworthiness primarily at the repository level, but are starting to engage discussion of trustworthiness at the document level.

2.2.1 Repository Trustworthiness Perception

Prior research on trustworthiness perception in the field of digital curation tends to focus on the repository level. Research has shown that several factors affect users' perception of a repository as trustworthy, such as:

- Users' past experience with a repository (e.g., positive past experience with a repository leads to perception of that repository as trustworthy) (Conway, 2010; Yoon, 2014),
- The accurate (or faithful) representation of data found within the context of a digital repository (Conway, 2010; Yoon, 2014),

- Positive recommendations about the repository from other people that the users know (Yakel et al., 2013; Yoon, 2014),
- A repository's processes (e.g., adherence to standards, data processing, metadata creation, data cleaning, providing documentation) (Conway, 2010; Niu, 2009; Yakel et al., 2013; Yoon, 2014), and
- A repository's integrity (e.g., honesty and lack of deception) (Yakel et al., 2013; Yoon, 2014).

A limitation of research on repository trustworthiness perception is that it tends to focus more on the factors that affect repository trustworthiness perception rather than specifying trustworthiness perception. A noteworthy exception to this trend is the work of St. Jean, Rieh, Yakel, and Markey (2011), who found that their study participants described trustworthiness at the repository level in terms of a series of attributes, such as factual, legitimate, reliable, reputable, professional, comprehensive, updated, and verifiable.

Of particular importance to my study is what research on trustworthiness perception at the repository level might suggest regarding research on trustworthiness perception at the document level. Results vary regarding the extent to which repository and document trustworthiness perceptions interact. Conway (2010) found that, for his study participants, trustworthiness “ascends to the organizational level and, as a consequence, pervades the resources delivered digitally” (p.455). In contrast, Yakel et al. (2013) found that “trust in the repository is a separate and distinct factor from trust in the data” (p. 11). Yoon (2014) calls for more research on the association between repository trustworthiness and

trustworthiness at the document level: “trust in data itself plays a distinctive and important role for users to reuse data, which may or may not be related to the trust in repositories” (p. 32).

Before we can attempt to understand the relationship between repository trustworthiness perception and document trustworthiness perception, we need a better understanding of both phenomena. As an initial step toward this type of research, my dissertation attempts to specify trustworthiness at the document level.

2.2.2 Document Trustworthiness Perception in the Context of a Digital Repository

Similar to research on repository trustworthiness perception, research on trustworthiness perception for documents or data found within digital repositories also tends to focus on the factors that affect trustworthiness perception. Research has shown that several factors affect users’ perceptions of documents or data found within a digital repository as trustworthy, such as:

- The author/creator/producer of the information (Fear & Donaldson, 2012; St. Jean, Rieh, Yakel, & Markey, 2011; Van House, 2002; Van House, 2003; Van House, Butler, & Schiff, 1998),
- Peer review of the information (Fear & Donaldson, 2012; St. Jean, Rieh, Yakel, & Markey, 2011),
- A document or dataset’s presence in a digital repository (Fear & Donaldson, 2012), and

- The intended use to which a dataset or document is put (Fear & Donaldson, 2012).

2.2.3 Summary of Empirical Research on Trustworthiness Perception in the Field of Digital Curation

At present, research on trustworthiness perception is focused more on the factors that affect trustworthiness perception rather than specifying definitions of trustworthiness perception as concepts at both the repository and document levels. A consequence of this focus, digital curation research treats the concepts of repository trustworthiness and document trustworthiness as monolithic “black boxes,” the contents of which are unknown. Future research should seek to specify definitions of trustworthiness at both repository and document levels to better understand the connection between repository and document trustworthiness perception. As a start, my dissertation specifies trustworthiness at the document level within the context of a digital repository that ought to be considered trustworthy.

2.3 Trustworthiness in the Field of Web Credibility

In contrast to the digital curation research literature, the field of web credibility has a more established history of studying trustworthiness. Part of what it means for any person, website, or information to be credible is to be trustworthy. Since trustworthiness is a core construct of credibility, web credibility research has

specifically explored user perception of trustworthiness. For comprehensive reviews of credibility, see Metzger, Flanagin, Eyal, Lemus, and McCann (2003), Rieh and Danielson (2007), and Metzger (2007).

Some of the research on user trustworthiness perception in the field of web credibility focuses at the document or information level. This research employs both qualitative and quantitative methods to examine users' perceptions of online information as trustworthy. These findings demonstrate the existence and importance of trustworthiness to users of online information, but also underscore the impact of context on trustworthiness assessments.

2.3.1 Qualitative Research on Trustworthiness

Qualitative research conducted in the field of web credibility suggests that trustworthiness, from the user's perspective, is a multifaceted concept. In Rieh (2002), participants expressed their views of the trustworthiness of information they encountered on the web using phrases such as "I trust it, trustworthy, believe in, confidence that this is true, seems real, [and] faith in the quality," suggesting that users conceptualize trustworthiness in terms of believability, perceived truthfulness, authenticity, and high (or at least sufficient) quality (p. 153). Metzger, Flanagin, and Medders (2010) found that participants in a focus group study considered information they perceive as objective as much more trustworthy than information they considered subjective.

The qualitative research studies in web credibility also underscore the

importance of context regarding users' assessment of online information as trustworthy. For example, Rieh and Hilligoss (2007) found that some information might be trustworthy enough for everyday life purposes, but not trustworthy enough to use in class assignments, based upon the sources from which the information were derived. Hargittai, Fullerton, Menchen-Trevino, and Thomas (2010) found that the context in which users find information online affects how trustworthy they perceive the information. Their study participants perceived information as trustworthy because they found that information in the context of Google—a search engine that they trusted.

2.3.2 Quantitative Research on Trustworthiness

Several quantitative studies have measured user trustworthiness perception for various types of online information. Flanagin and Metzger (2000) and Metzger, Flanagin, and Zwarun (2003) measured participants' trustworthiness perceptions for commercial, entertainment, news, and reference information found on the Internet. In both studies, participants rated how trustworthy they thought commercial, entertainment, news, and reference information found on the Internet was on a 7-point scale ranging from 1="not at all" to 7="extremely." Similarly, Liu (2004) and Liu and Huang (2005) measured users' trustworthiness perceptions using "The document content is trustworthy" as a measurement item. Flanagin and Metzger's (2007) study participants indicated how trustworthy they thought a news story on the topic of the potentially harmful effects of radiation on pregnant women

who fly in airplanes was on a 7-point scale ranging from 1=“not at all” to 7=“extremely.” Rieh, Kim, Yang, and St. Jean (2010) used the item “The information is trustworthy,” to measure their participants’ trustworthiness perceptions on a 7-point scale ranging from 1=“not at all” to 7=“very much” for information encountered while using various traditional web sites, user-generated content sites, and multimedia sites.

In all of these studies, the web credibility researchers collected data from participants suggesting that they had high levels of trustworthiness perception. The findings generated by the trustworthiness measurement items in these studies all fell above the mean. These findings indicate that trustworthiness is important to users of various types of online information and is measurable.

2.3.3 Summary on Trustworthiness in the Field of Web Credibility

Taken together, the web credibility literature demonstrates a dichotomy regarding qualitative data collected as a result of asking questions about trustworthiness versus quantitative measurement of the concept of trustworthiness. The qualitative research literature demonstrates that users are capable of articulating their perceptions of trustworthiness at the document or information level. The qualitative findings also underscore the role of context in shaping users’ perceptions of online information as trustworthy. By contrast, the quantitative research literature demonstrates the importance and measurability of

trustworthiness for users of online information, but treats trustworthiness as a monolithic concept that can be measured using a single item.

Since the qualitative and quantitative findings demonstrate the importance of trustworthiness, the measurability of trustworthiness, and the role of context regarding trustworthiness, there seems an opportunity to investigate trustworthiness more deeply in other contexts besides the open web by adapting web credibility researchers' approaches toward understanding and measuring trustworthiness perception at the document level using more than a single item. In particular, there seems an opportunity to ask users about their perceptions of trustworthiness in qualitative research studies and build upon that research with measurement of the concept of trustworthiness using multiple items that correspond to all the important aspects of trustworthiness from the user's perspective.

2.4 A Conceptual Framework for Trustworthiness Perception

Frameworks for trustworthiness perception in the existing literature of management and information science tend to focus on the technical expertise of the organization providing the information or the constructs of the individuals judging trustworthiness. For example, Pirson and Malhotra's (2011) model of trustworthiness situates technical competence, an "organization's ability to deliver high-quality products and services," as part of what affects people's perceptions of that organization as trustworthy (p. 1092). To evaluate technical competence,

people external to the organization consider their “satisfaction with the quality of the product or service being offered” (Pirson & Malhotra, 2011, p. 1092). Although Pirson and Malhotra’s (2011) model of trustworthiness was not developed specifically with digital repositories and digital repository users in mind, Yakel et al. (2013) have adapted Pirson and Malhotra’s framework to the study of trust in digital repositories. While Yakel et al.’s work demonstrates the promise of Pirson and Malhotra’s model for understanding trust in digital repositories, Yakel et al. also demonstrate that more work, specifically adaptation of Pirson and Malhotra’s model is necessary to apply within a digital repository context. In particular, specification of perception of the product or service, in this case, the document or the information, is underspecified.

More specific than Pirson and Malhotra (2011) regarding definition of trustworthiness at the document or information level is Hilligoss and Rieh’s (2008) specification of trustworthiness in their Unifying Framework for Credibility Assessment. They argue that “information is trustworthy when it appears to be reliable unbiased, and fair” (Hilligoss & Rieh, 2008, p. 1469).

In contrast to Pirson and Malhotra (2011) and Hilligoss and Rieh (2008), the most precise conceptual model for trustworthiness at the document or information level is found in Kelton, Fleischmann, and Wallace’s (2008) “Integrated Model of Trust in Information” (see Figure 2.1). The model posits trustworthiness as one of four clusters of factors (along with disposition, relevance judgments, and the authority of recommendations) that lead a user from an encounter with an information source to the confidence and willingness to make use of that source. In

Figure 2.1, the labels and arrows pointing to and from *Trustworthiness* denote the relationships between *Trustworthiness* and other constructs, suggesting that trustworthiness perception is a process, dependent upon multiple related factors that reinforce one another.

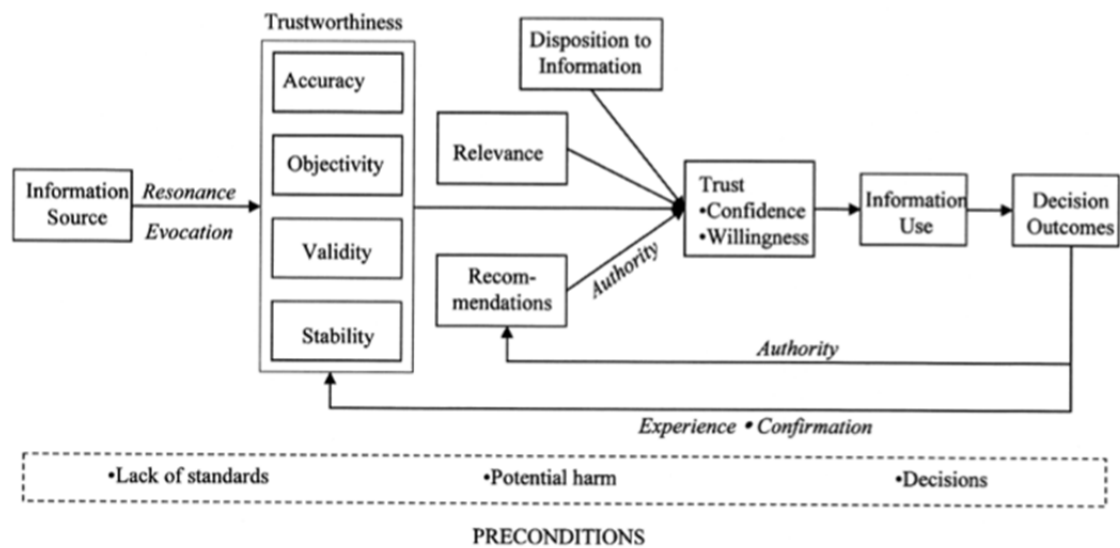


Figure 2.1 Kelton, Fleischmann, & Wallace's (2008) Integrated Model of Trust in Information.

The *Trustworthiness* component in the Kelton model (p. 370) is the collective outcome of user perceptions of four properties of an information source: accuracy, objectivity, validity, and stability, defined as:

- *Accuracy*: the extent to which information is free from error (plus believability, coverage, and currency).
 - *Believability*: the extent to which information appears to be plausible.
 - *Coverage*: the completeness of information.
 - *Currency*: the degree to which information is up-to-date rather than obsolete.

- *Objectivity*: the balance of content.
- *Validity*: the use of responsible and accepted practices, such as:
 - The soundness of the methods used,
 - The inclusion of verifiable data, and
 - The appropriate citation of sources.
- *Stability*: the persistence of information, both its presence and contents.

Unlike other efforts in the information science literature to define trust in information (Lucassen & Schraagen, 2011; Lucassen & Schraagen, 2012; Lucassen, Muilwijk, Noordzij, & Schraagen, 2013), the Kelton model is an admirable attempt to construct a model that can generate testable hypotheses.

Interestingly, Kelton et al.'s (2008) model does not take full advantage of the literature on trustworthiness to provide empirical support of its specification of trustworthiness. For example, besides citing an early empirical study of web credibility (Rieh & Belkin, 1998), Kelton et al. fail to mine the web credibility literature more deeply to provide a broad and solid foundation for their concept of trustworthiness.

2.5 Defining Trustworthiness

Just as conceptual models for trustworthiness range in their scope and level of specificity, so do definitions of trustworthiness. Analysis of definitions of trustworthiness in digital curation, information science, web credibility, and

communication studies demonstrates the range and scope of various definitions of trustworthiness in the literature.

Duranti (1995), MacNeil (2000), and Duranti, Suderman, and Todd's (2008) definitions of trustworthiness were relevant to this study because they engaged the document level. For example, Duranti (1995) and MacNeil (2000) define trustworthiness in terms of a document's reliability, authenticity, and genuineness. A reliable document is "capable of standing for the facts to which it attests" (MacNeil, p. xi) and is assessed in terms of the document's "completeness and controlled procedure of creation" (Duranti, p. 8). An authentic document is, ontologically speaking, "what it purport[s] to be" (Duranti) or "is what it claims to be" (MacNeil). A genuine document, according to Duranti, is one whose content is truthful. Related to Duranti's and MacNeil's conceptualizations of document and record trustworthiness is Duranti, Suderman, and Todd's (2008) definition. Somewhat circuitously, they define a digital record as one that is both reliable and accurate. "Reliability is the trustworthiness of a record's content," they claim, "because by definition, the content of a reliable record is trustworthy, and trustworthy content is, in turn, predicated on accurate data, it follows that a reliable record is also an accurate record" (p. 667). Duranti (1995), MacNeil (2000), and Duranti, Suderman, and Todd's (2008) definitions partition trustworthiness into multiple components, some having to do with a document "as a document" and others having to do with a document "as content or information." Altogether, Duranti (1995), MacNeil (2000), and Duranti, Suderman, and Todd's (2008) definitions specify trustworthiness at the document level for archival documents

found in the digital realm, but are also limiting because they do not specify document trustworthiness in terms of user perception.

Broader definitions of trustworthiness are found in information science, web credibility, and communication studies literatures. These definitions refer to trustworthiness at the document level from the perspective of user perception. For example, Kelton et al.'s (2008) definition considers trustworthiness as the extent to which a user perceives information as accurate, current, complete, believable, objective, valid, and stable. Also relevant is Hilligoss and Rieh's (2008) definition, which casts reliable, objective, and fair information as trustworthy.

Perhaps the most flexible definition of trustworthiness that also specifies trustworthiness at the document level from the perspective of user perception is Flanagin and Metzger's (2008) definition. They define trustworthiness as "a receiver judgment based primarily on subjective factors" (p. 8).

2.6 Conclusion

Taken together, the literature demonstrates that trustworthiness is central to justification for digital repositories, but it has only been asserted as a concept. Trustworthiness has not been defined in a way that is amenable to verifying its presence or absence in a repository context from the user's perspective. The research on users has identified factors that affect their perception of trustworthiness at both repository and document levels. These findings provide insight into assumptions users make about the type of trustworthiness repositories

assert. Existing empirical research also suggests that users' conceptualization of trustworthiness is multi-faceted (e.g., St. Jean et al., 2011). For any repository, a fuller understanding of how their designated communities conceptualize trustworthiness would involve assessing those members' perceptions of trustworthiness. Toward this end, there is an opportunity for research that leverages approaches within the web credibility literature to inform a fuller understanding of user trustworthiness perception for archival documents delivered digitally.

The following chapter describes the methodology for development of a scale for measuring designated communities' concept of trustworthiness within a digital repository context at the document level.

2.7 References

Center for Research Libraries. (n.d.). TRAC and TDR Checklists. Retrieved from:

<http://www.crl.edu/archiving-preservation/digital-archives/metrics-assessing-and-certifying-0>

Consultative Committee for Space Data Systems. (2002). *Reference model for an open archival information system (OAIS) recommendation for space data system standards ; blue book* (Blue Book. Washington, DC: CCSDS Secretariat.

doi:<http://public.ccsds.org/publications/archive/650x0b1s.pdf>

Consultative Committee for Space Data Systems. (2011). *Audit and certification of trustworthy digital repositories; magenta book*. No. CCSDS 652.0-M-1).

Washington, DC: CCSDS Secretariat.

Conway, P. (2010). Modes of seeing: Digitized photographic archives and the experienced user. *American Archivist*, 73(2), 425-462.

Data Seal of Approval Board. (2013). *Data seal of approval guidelines* No. Version 2).

Netherlands: Data Archiving and Networked Services.

Deutsches Institut für Normung. (2012). *Information and documentation—Criteria for trustworthy digital archives*. Berlin: Deutsches Institut für Normung.

Dow, E. H. (2009). *Electronic records in the manuscript repository*. Lanham, Md.: Scarecrow Press.

Duranti, L. (1995). Reliability and authenticity: The concepts and their implications. *Archivaria*, 39, 5-10.

- Duranti, L., Suderman, J., & Todd, M. (2008) Part seven-structuring the relationship between records creators and preservers: Policy cross-domain task force report. In L. Duranti, & R. Preston (Eds.), *International research on permanent authentic records in electronic systems (InterPARES) 2: Experiential, interactive and dynamic records* (). Padova, Italy: Associazione Nazionale Archivistica Italiana. Retrieved from http://www.interpares.org/display_file.cfm?doc=ip2_book_part_7_policy_task_force.pdf
- Fear, K., & Donaldson, D. R. (2012). Provenance and credibility in scientific data repositories. *Archival Science*, 12(3), 319-339.
- Flanagin, A. J., & Metzger, M. (2008). Digital media and youth: Unparalleled opportunity and unprecedented responsibility. In M. J. Metzger, & A. J. Flanagin (Eds.), *Digital media, youth, and credibility* (pp. 5-28). Cambridge: MIT Press.
- Flanagin, A. J., & Metzger, M. J. (2000). Perceptions of internet information credibility. *Journalism and Mass Communication Quarterly*, 77(3), 515-540.
- Flanagin, A. J., & Metzger, M. J. (2007). The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. *New Media & Society*, 9(2), 319-342.
- Hargittai, E., Fullerton, L., Menchen-Trevino, E., & Thomas, K. Y. (2010). Trust online: Young adults' evaluation of web content. *International Journal of Communication*, 4(1), 468-494.

- Hilligoss, B., & Rieh, S. Y. (2008). Developing a unifying framework of credibility assessment: Construct, heuristics, and interaction in context. *Information Processing & Management*, 44(4), 1467-1484.
- International Organization for Standardization. (2012). *Space data and information transfer systems : Audit and certification of trustworthy digital repositories* [Systèmes de transfert des informations et données spatiales : audit et certification des référentiels numériques de confiance] (1st ed.). Geneva, Switzerland: International Organization for Standardization.
- . (2014). *Space data and information transfer systems : Requirements for bodies providing audit and certification of candidate trustworthy digital repositories* [Systèmes de transfert des informations et données spatiales : audit et certification des référentiels numériques de confiance] (1st ed.). Geneva, Switzerland: International Organization for Standardization.
- Kelton, K., Fleischmann, K. R., & Wallace, W. A. (2008). Trust in digital information. *Journal of the American Society for Information Science and Technology*, 59(3), 363-374.
- Lee, C. A., & Tibbo, H. R. (2007). Digital curation and trusted repositories: Steps toward success. *Journal of Digital Information*, 8(2)
- Liu, Z. (2004). Perceptions of credibility of scholarly information on the web. *Information Processing & Management*, 40(6), 1027-1038.
- Liu, Z., & Huang, X. (2005). Evaluating the credibility of scholarly information on the web: A cross cultural study. *The International Information & Library Review*, 37(2), 99-106.

- Lucassen, T., Muilwijk, R., Noordzij, M. L., & Schraagen, J. M. (2013). Topic familiarity and information skills in online credibility evaluation. *Journal of the American Society for Information Science and Technology*, 64(2), 254-264.
- Lucassen, T., & Schraagen, J. M. (2011). Factual accuracy and trust in information: The role of expertise. *Journal of the American Society for Information Science and Technology*, 62(7), 1232-1242.
- . (2012). Propensity to trust and the influence of source and medium cues in credibility evaluation. *Journal of Information Science*, 38(6), 566-577.
- MacNeil, H. (2000). *Trusting records: Legal, historical and diplomatic perspectives*. Dordrecht; London: Kluwer Academic.
- Metzger, M. J., Flanagin, A. J., Eyal, K., Lemus, D. R., & McCann, R. M. (2003). Credibility for the 21st century: Integrating perspectives on source, message, and media credibility in the contemporary media environment. *Communication Yearbook*, 27, 293-336.
- Metzger, M. J. (2007). Making sense of credibility on the web: Models for evaluating online information and recommendations for future research. *Journal of the American Society for Information Science and Technology*, 58(13), 2078-2091.
- Metzger, M. J., Flanagin, A. J., & Medders, R. B. (2010). Social and heuristic approaches to credibility evaluation online. *Journal of Communication*, 60(3), 413-439.
- Metzger, M. J., Flanagin, A. J., & Zwarun, L. (2003). College student web use, perceptions of information credibility, and verification behavior. *Computers & Education*, 41(3), 271-290.

- NESTOR Working Group on Trusted Repositories Certification. (2006). *Catalogue of criteria for trusted digital repositories* No. Version 1). Deutsche Nationalbibliothek: NESTOR Working Group.
- NESTOR Working Group on Trusted Repositories Certification. (2009). *Catalogue of criteria for trusted digital repositories* No. Version 2). Deutsche Nationalbibliothek: NESTOR Working Group.
- . (2013). *Explanatory notes on the nestor seal for trustworthy digital archives*. Germany: Deutsche Nationalbibliothek.
- Niu, J. (2009). Overcoming inadequate documentation. *Proceedings of the American Society for Information Science and Technology*, 46(1), 1-14.
- Pirson, M., & Malhotra, D. (2011). Foundations of Organizational Trust. *Organization Science*, 22(4), 1087-1104.
- PTAB. (n.d.). Home. Retrieved from: <http://www.iso16363.org>
- Rieh, S. Y. (2002). Judgment of information quality and cognitive authority in the web. *Journal of the American Society for Information Science and Technology*, 53(2), 145-161.
- Rieh, S. Y., & Danielson, D. R. (2007). Credibility: A multidisciplinary framework. *Annual Review of Information Science and Technology*, 41(1), 307-364.
- Rieh, S. Y., & Hilligoss, B. (2007). College students' credibility judgments in the information-seeking process. *The John D. and Catherine T. MacArthur Foundation Series on Digital Media and Learning*, -, 49-71.

- Rieh, S. Y., Kim, Y., Yang, J. Y., & St. Jean, B. (2010). A diary study of credibility assessment in everyday life information activities on the web: Preliminary findings. *Proceedings of the American Society for Information Science and Technology*, 47(1), 1-10.
- Rieh, S. Y., & Belkin, N. J. (1998). Understanding judgment of information quality and cognitive authority in the WWW. Paper presented at the *ASIS Annual Meeting*, Pittsburgh, PA. Retrieved from <http://rieh.people.si.umich.edu/papers/asis98.pdf>
- RLG/OCLC Working Group on Digital Archive Attributes. (2002). *Trusted digital repositories: Attributes and responsibilities*. Mountain View, CA: RLG. Retrieved from <http://www.oclc.org/research/activities/past/rlg/trustedrep/repositories.pdf>
- RLG-NARA Digital Repository Certification Task Force. (2007). *Trustworthy repositories audit and certification: Criteria and checklist* OCLC and CRL. Retrieved from http://www.crl.edu/sites/default/files/attachments/pages/trac_0.pdf
- St. Jean, B., Rieh, S. Y., Yakel, E., & Markey, K. (2011). Unheard voices: Institutional repository end-users. *College & Research Libraries*, 72(1), 21-42.
- Van House, N. A., Butler, M. H., & Schiff, L. R. (1998). Cooperative knowledge work and practices of trust: Sharing environmental planning data sets. *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work*, pp. 335-343.

- Van House, N. A. (2002). Digital libraries and practices of trust: Networked biodiversity information. *Social Epistemology*, 16(1), 99-114.
- . (2003). Digital libraries and collaborative knowledge construction. In A. P. Bishop, N. A. Van House & B. P. Battenfield (Eds.), *Digital library use: Social practice in design and evaluation* (pp. 271-295). Cambridge, Mass.: MIT Press.
- Waters, D., & Garrett, J. (1996). *Preserving digital information : Report of the task force on archiving of digital information*. Washington, D.C.: Council on Library and information Resources. Retrieved from <http://www.clir.org/pubs/reports/pub63watersgarrett.pdf>
- Yakel, E., Faniel, I., Kriesberg, A., & Yoon, A. (2013). Trust in digital repositories. *International Journal of Digital Curation*, 8(1), 143-156.
- Yoon, A. (2014). End users' trust in data repositories: Definition and influences on trust development. *Archival Science*, 14(1), 17-34.

CHAPTER 3

METHODOLOGY: THE FOUR STEPS OF SCALE DEVELOPMENT

In the previous chapter, I examined the concept of trustworthiness as described in international standards for Trusted Digital Repositories (TDRs). I also explored the empirical studies that focus on trustworthiness perception within the context of digital repositories, finding that a deeper understanding of trustworthiness perceptions at the document level is still unknown. This gap in the research literature leads to the formulation of two research questions:

1. How do designated community members conceptualize trustworthiness for documents they find in a digital repository?
2. To what extent are designated community members' perceptions of document trustworthiness measurable?

To address both research questions, I employed the established methodology of scale development as defined in Spector (1992) and DeVellis (2012). The premise for using scale development for this study was that, if trustworthiness does in fact exist, then it ought to be definable and measurable. After discussing the process of selecting scale development for this study, this chapter describes the methodology

of scale development as it applies to an exploratory study of the concept of trustworthiness at the document level within a digital repository context.

3.1 Selecting Scale Development

I considered two different approaches to addressing my research questions before settling on a third approach: scale development:

- Interviews followed by measurement of trustworthiness perception using a single-item measure (The first approach), and
- Index construction (The second approach).

3.1.1 The First Approach

The first approach I considered would have involved use of one-on-one interviews to address my first research question, How do members of a designated community conceptualize trustworthiness for documents they find in a digital repository?, followed by use of a single-item measure to address my second research question, To what extent are designated community members' perceptions of document trustworthiness measurable? One-on-one interviews are useful because they provide in-depth accounts of individuals' experiences, attitudes and perceptions (Kitzinger, 1995). Applied to this study, I considered asking participants specifically about their perceptions of trustworthiness for digitized archival

documents with which they had recent experiences. While using one-on-one interviews had the potential to address my first research question, it may well have addressed my first research question at too granular of a level. Often, one-on-one interviews provide the specific details of individuals' points of view (Kitzinger, 1995), which may have led to identification of idiosyncratic definitions of trustworthiness perception in this study. I wanted to identify definitions of trustworthiness perception that corresponded to a broad sweep of designated community members; thus, I judged the one-on-one interview technique as not the most appropriate means of collecting data regarding my first research question.

To address my second research question, To what extent are designated community members' perceptions of document trustworthiness measurable?, I considered using single-item measures. According to Kim (2009), a single-item measure can be appropriate for measuring perceptions of a concept. For example, I considered using the single item, "The document is trustworthy," on a 7-point scale to measure users' perceptions of digitized archival documents as trustworthy. However, measurement of people's perceptions of a concept only works if that concept is both "easily and uniformly imagined" in the minds of the individuals who provide responses to that measure (Kim, 2009, p. 1179). I decided against using a single-item measure of trustworthiness. While there is a precedent for measuring trustworthiness perception using a single-item measure in other research domains (e.g., Liu, 2004; Liu & Huang, 2005), there is not enough empirical research on trustworthiness perception at the document level within a digital repository context to know if actual users uniformly imagine the concept. Research on trustworthiness

perception at the document or data level within the context of a digital repository is only beginning to understand the phenomenon (Conway, 2010; St. Jean et al., 2011; Yakel et al., 2013; Yoon, 2014).

In sum, using one-on-one interviews and single-item measures would have addressed both of my research questions. However, use of both of these approaches would have done so sub-optimally.

3.1.2 The Second Approach

I also considered the methodology of index construction as a means of addressing my research questions. Indexes are “sets of items that are cause indicators—that is, items that determine the level of a construct” (DeVellis, 2012, p. 12). Indexes include items that are formative; in aggregate, index items compose (i.e., form) the construct of interest (Bollen & Lennox, 1991). Index construction refers to the building, testing, and assessment of indexes (Diamantopoulos & Winklhofer, 2001). According to Diamantopoulos and Winklhofer (2001), index construction involves four steps (pp. 271- 273): 1) content specification (i.e., specifying the domain of the content the index is supposed to capture); 2) indicator specification (i.e., identifying the complete census of indicators (i.e., items) that are necessary to formulate the construct of interest); 3) indicator collinearity (i.e., assessing the influence of each formative variable on the construct of interest); and 4) external validity (i.e., examining the relationship of the index to other reflective indicators of the construct of interest for purposes of validation).

By applying these four steps, I could have addressed both of my research questions. During Step 1, I could have specified the domain of content that was necessary to capture trustworthiness perception. During Step 2, I could have identified items that in aggregate form the concept of trustworthiness perception emerging from Step 1. During Step 3, I could have administered the items resulting from Step 2 to a sample of participants, collected their responses, and assessed the collinearity of the concepts that in aggregate form trustworthiness perception. During Step 4, I could have compared the correlation of the trustworthiness perception index, which would focus on trustworthiness perception at the document level, to other items that are supposed to reflect the concept of trustworthiness perception at the repository level. If in fact repository trustworthiness and document trustworthiness should correlate, then the index emerging from index construction should correlate with items that could measure repository trustworthiness perception.

While, if successful, index construction would have supported both of my research questions, the methodology of scale development is more aligned with the goals of the research study. In contrast to indexes, scales include items that are reflective; they reflect the construct of interest (Diamantopoulos & Winklhofer, 2001; DeVellis, 2012). My research identifies the structure underlying a set of items that are related to trustworthiness and ascertains which items are most important for measurement of trustworthiness, Scale development could better facilitate these particular research aims. The remainder of this chapter provides more discussion

and explanation of the value of scale development and describes the methodology as applied to this study.

3.2 The Methodology of Scale Development

Researchers employ scale development when they want to measure a phenomenon that they believe exists but is not directly observable (DeVellis, 2012). A scale is a measurement instrument that is a collection “of items combined into a composite score and intended to reveal levels of theoretical variables not readily observable by direct means” (DeVellis, 2012, p. 11). The process of scale development is not merely assembly of items for measurement of a concept; the development aspect of scale development denotes a careful and meticulous methodology for arriving at a scale that is both reliable (i.e., it consistently measures a phenomenon) and valid (i.e., it measures what it claims to measure). The most highly cited literature on scale development states that scale development involves four primary steps (DeVellis, 2012; Spector, 1992):

- Step 1—Construct Definition
- Step 2—Generating an Item Pool
- Step 3—Designing the Scale
- Step 4—Full Administration and Item Analysis.

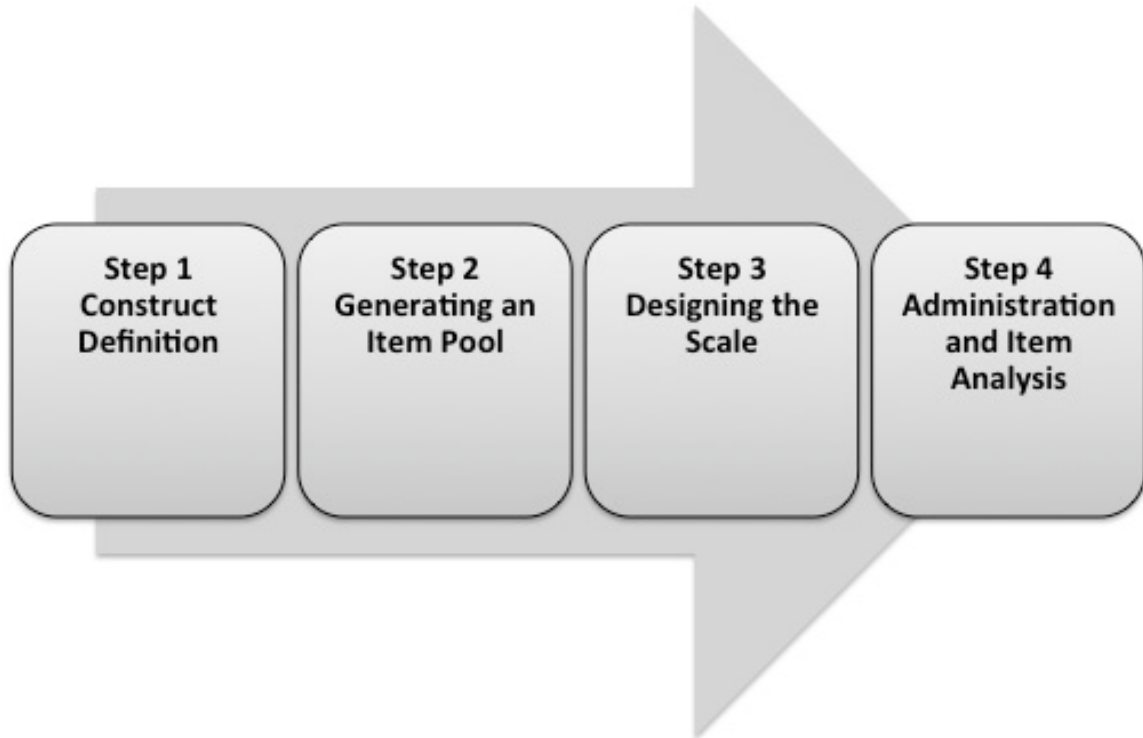


Figure 3.1 Scale Development Procedure
Adapted from Spector (1992) and DeVellis (2012).

As shown in Figure 3.1, the first step of scale development is Construct Definition. Step 1 involves review of the literature to identify the scope of the concept for the purpose of empirical investigation. Step 1 also involves focus groups to understand how members of a particular population talk about a concept. As shown in Figure 3.1, the second step of scale development is Generating an Item Pool. Step 2 involves identifying items for measurement of a concept from multiple sources, including the literature, subject matter experts, and focus groups data. As shown in Figure 3.1, the third step of scale development is Designing the Scale. Step 3 involves transforming the item pool resulting from Step 2 into a web survey for pretesting and refinement. As shown in Figure 3.1, the fourth step of scale development is Full Administration and Item Analysis. Step 4 involves

administering the final item pool comprising items gathered from earlier steps of scale development to a large sample of participants for their evaluation. Step 4 also involves analyzing participants' responses to identify the variables underlying the set of items and identify the items that are most essential for measurement of the concept. Step 4 also involves validation of the scale resulting from the statistical tests during Step 4 by administering the scale to a different sample of participants. In this study, I applied all four steps of scale development to the concept of trustworthiness—from the perspective of actual users of digitized archival documents found within a digital repository.

Scale development is a rigorous methodology. Each of the four steps of scale development has its own set of methods. The methods of each step build on one another and, in turn, each step builds on the one preceding it. Hence the arrow pointing from left to right containing the steps of scale development in Figure 3.1. The remainder of this chapter describes the four steps of scale development in detail, including description of the methods that are necessary for execution of each step.

3.2.1 Step 1 of Scale Development—Construct Definition

Step 1 of scale development involves completing four tasks related to defining the construct of interest (See Table 3.1). First, according to Choemprayong and Wildemuth (2009), Step 1 involves focusing attention on “clearly defining what it is he or she wants to measure and the goals of that measurement” (p. 281).

Second, Step 1 involves developing a brief definition of the construct of interest, “including its scope (i.e., what it is and what it is not) and any subcomponents that are to be included” (p. 281). Third, Step 1 involves drawing upon definitions of the construct that exist in prior research studies or more theoretical literature to further develop the construct definition. Fourth, Step 1 involves operationalizing the construct definition by considering the different types of questions or rating scales to which study participants can respond, and asking oneself what kinds of responses would be clear indicators of the respondents’ levels or amounts of the construct of interest.

-
- Focus on clearly defining what you want to measure and the goals of that measurement.
 - Develop a brief definition of the construct of interest, including its scope and subcomponents.
 - Draw upon definitions of the construct that exist in prior research studies or more theoretical literature to further develop construct definition.
 - Operationalize construct definition.
 - Conduct focus groups.
-

Table 3.1 Step 1 of Scale Development—Construct Definition
(DeVellis, 2012; Spector, 1992).

In addition, focus groups can provide clarity during Step 1. According to DeVellis (2012):

It is often important to assess whether the theoretical constructs we as researchers identify correspond with the actual perceptions and experiences of the people we plan to study. Focus groups ... can be a means of determining whether ideas that underlie constructs of interest make sense to participants. (p. 187)

Focus groups can also provide insight into participants’ own terminology regarding a concept (DeVellis, 2012). Overall, focus groups can clarify the scope and meaning

of the concept from the perspective of the participants and can aid construct definition.

In this study, I applied the aforementioned tasks to the concept of trustworthiness perception. Section 2.5 includes analysis of the literature I reviewed for Step 1. Section 1.1 includes the construct definition of trustworthiness perception for this study. Chapter 4 discusses the results of the focus groups.

3.2.2 Step 2 of Scale Development—Generating an Item Pool

DeVellis (2012) explains that Step 2 of scale development involves a number of tasks and considerations related to generating a large pool of items that are candidates for eventual inclusion in the scale for the construct of interest, nineteen of which are highlighted in this section (See Table 3.2).

First, according to Spector (1992) and Choemprayong and Wildemuth (2009), Step 2 involves examining any existing instruments that measure the construct of interest.

Second, Spector (1992) states that the scale developer may choose items from several instruments “as a starting point in writing an initial item pool” (p. 16). If instruments to measure the construct of interest do not exist, then utilizing other ways of generating an item pool is necessary. For example, Choemprayong and Wildemuth (2009) point out that existing instruments may not measure the construct of interest specifically, but may be related to the construct of interest such that some of their items are acceptable to include in an item pool.

Third, scale developers should consider deriving items from related instruments. If no items in existing or related instruments are appropriate, the scale developer still has a number of options for generating items. For example, DeVellis (2012) states that a scale developer can create items – the fourth option for scale developers regarding generating items. When creating items, he or she should think creatively about the construct of interest concerning how many ways an item can be worded “so as to get at the construct,” and then write those items down (p. 77).

Fifth, according to Bernard (2000), “ideas for [items] can come from reading the literature on whatever research problem has captured [the scale developer], from personal experience, from ethnography, from reading newspapers, [sic] from interviews with experts” (p. 295).

-
- Examine existing instruments that measure the construct of interest
 - Choose items from existing instruments
 - Create items
 - Get ideas for items from reading the literature based on the research problem, from personal experience, from ethnography, from reading newspapers, from interviews with experts, etc.
 - Make sure each item expresses only one idea
 - Make sure the item pool consists of an exhaustive list of items that appear to fit the construct of interest’s definition (i.e., face validity)
 - Avoid exceptionally lengthy items
 - Make items as short and uncomplicated as possible
 - Make reading difficulty level match that of the intended respondents
 - Avoid colloquialisms, expressions, and jargon
 - Select items that match the specificity of the construct of interest
 - If there are concerns about bias, include validation items (if any exist)
 - Consider exclusion or inclusion of negatively and positively worded items
 - Consider amount of items to include in item pool
 - Consider expert involvement (for deriving items, assessing selected items, etc.)
 - Consider number of items for item pool
 - Consider acquiescence bias issues
 - Derive items from focus groups
-

Table 3.2 Step 2 of Scale Development—Generating an Item Pool

(Babbie, 2010; Bernard, 2000; Choemprayong & Wildemuth, 2009; DeVellis, 2012; Spector, 1992).

Regardless of where items originate, a scale developer must consider a number of issues during Step 2:

- Sixth - Each item should express one idea (i.e., no double-barreled questions) (Bernard, 2000; DeVellis 2012; Spector, 1992)
- Seventh - The “thing” the items have in common should be a construct and not merely a category. As DeVellis (2012) points out, “just because items relate to a common category, that does not guarantee that they have the same underlying latent variable” (p. 77).
- Eighth - The item pool should consist of an exhaustive list of items that appear to fit the construct of interest’s definition (Choemprayong & Wildemuth, 2009; DeVellis 2012)
- Ninth - Exceptionally lengthy items should be avoided (DeVellis, 2012)
- Tenth - Items should be as short and as uncomplicated as possible (Bernard, 2000; DeVellis, 2012)
- Eleventh - The reading difficulty level should match that of the respondents (DeVellis, 2012; Spector, 1992)
- Twelfth - Colloquialisms, expressions, and jargon should be avoided (Spector, 1992)
- Thirteenth - Items should match the specificity of the construct of interest (i.e., make sure items aren’t too specific or too general based on the construct of interest) (Babbie, 2010; DeVellis, 2012)
- Fourteenth - Concerns about bias should be addressed by including validation items, if any exist (DeVellis, 2012; Spector, 1992)

- Fifteenth - If prior research asserts that the phenomenon the scale developer is measuring relates to other constructs, “then the performance of the scale vis-à-vis measures of those other constructs can serve as evidence of its [construct] validity” (DeVellis, 2012, p. 102)

Positively and negatively worded items: Sixteenth, although Spector (1992) and Bernard (2000) recommend using an even number of positively and negatively worded items to avoid acquiescence bias, DeVellis (2012) states that, in practice, the disadvantages could outweigh the benefits, in particular, if reversals in item polarity confuse respondents. Thus, inclusion of negatively and positively worded items is another issue of concern.

Amount of items: Seventeenth, another issue is the amount of items an item pool should consist of. Some recommend a specific amount of items. For example, Bernard (2000) recommends that “you should have four or five times the number of items you think you’ll need in your final scale” (p. 296). Less prescriptive, DeVellis (2012) states that, “it is impossible to specify the number of items that should be included in an initial pool. Suffice it to say that you want considerably more than you plan to include in the final scale” (p. 80). DeVellis (2012) points to numerous advantages of having a large item pool, such as insurance against poor internal consistency and being able to be “fussier” about the items he or she chooses to include in an actual instrument (p. 80).

Expert Involvement: Eighteenth, the extent to which experts on the construct of interest should be involved is another consideration. For example, Choemprayong and Wildemuth (2009) recommend conferring with experts in the field to derive items, while DeVellis (2012) suggests that experts review the entire initial item pool. Common to all of these recommendations is the idea is that expert involvement maximizes content validity with respect to items chosen for an item pool.

Focus Groups: Nineteenth, focus groups can be a source of items for an item pool. Focus groups provide insights into the “natural, everyday language that people use to talk about a concept” (DeVellis, 2012, p. 187). Thus, scale developers can use the responses of participants as scale items to describe the concept in a way that resonates with how people actually consider the concept.

In this study, I applied all of the aforementioned tasks to generating a pool of items for measurement of trustworthiness perception. Those items come from three sources: the literature; trustworthiness subject matter experts; and participants’ responses during focus groups on the topic of document trustworthiness perception. I extracted items from the literature by paying close attention to definitions and/or findings in archival science, communication studies, digital curation, digital preservation, information science, and web credibility literatures regarding the concept of trustworthiness. I derived items from those definitions and findings. I gathered items from trustworthiness subject matter experts by asking

them to provide recommendations for items to measure the concept of document trustworthiness perception in a web-administered survey. I derived items from the focus groups by turning statements that participants provided regarding their concept of trustworthiness into items for measurement of trustworthiness perception.

3.2.3 Step 3 of Scale Development—Designing the Scale

Step 3 involves transforming the item pool resulting from Step 2 into a survey instrument for pretesting and refinement (see Table 3.3). The survey facilitates participants' evaluation of the item pool items—a critical part of scale development. Step 3 involves seven activities, including:

- Deciding on the order of the items.
- Selecting the response categories and choices.
- Writing item stems.
- Writing instructions for participants regarding the survey.
- Selecting the mode of administration for the survey.
- Pretesting the survey by conducting cognitive interviews with a sample of participants, and
- Revising the survey items and instructions based upon the cognitive interviewees' feedback.

-
- Deciding on the order of the items.
 - Selecting the response categories and choices.
 - Writing item stems.
 - Writing instructions for participants regarding the survey.
 - Selecting the mode of administration for the survey.
 - Pretesting the survey by conducting cognitive interviews with a sample of participants, and
 - Revising the survey items and instructions based upon the cognitive interviewees' feedback.
-

Table 3.3 Step 3 of Scale Development—Designing the Scale

(Collins, 2003; Choemprayong & Wildemuth, 2009; DeVellis, 2012; Spector, 1992)

Order of Items: Randomizing the order in which items appear to participants can address order effects.

Response Categories: According to Spector (1992), evaluation response categories ask for an evaluative rating for each item. Since evaluation categories measure people's attitudes and perceptions, evaluation response categories as opposed to other commonly used response categories, such as agreement or frequency response categories, were appropriate for this study.

Response Choices: Nunnally (1978) recommends between five and nine response categories per item in a psychometric instrument—with the inclusion of more rather than fewer choices when developing a scale. Per Nunnally's suggestion, I chose eight response choices since I was developing a scale.

Item Stems: According to Spector (1992), the phrasing of items should depend upon "the type of judgment[s] or response[s] people are asked to make" (pp. 22-23).

Writing Instructions: Spector (1992) and Choemprayong and Wildemuth (2009) recommend that instructions provide details about how to use the instrument, while also providing a “frame of reference” concerning the construct of interest. Providing a frame of reference for participants helps them appropriately engage in evaluating the items. I used the critical incident technique to evoke participants’ frame of reference (Flanagan, 1954). In the survey instructions, I asked participants to consider their recent experiences with digitized archival documents in their evaluation of the trustworthiness items.

In addition, Spector (1992) also recommends that instructions explain the response choices to participants.

Mode of Administration: According to Choemprayong and Wildemuth (2009), there are two primary advantages to computer administration of a scale, including automatic recording of participants’ responses and the “capability of administering the scale remotely” (p. 282). Computer administration of the scale facilitated analysis of the participants’ data and also enabled participants who were highly geographically dispersed to participate in this study.

Pretesting: Collins (2003) recommends pretesting survey instruments, arguing that “cognitive testing should be *a standard part* [emphasis added] of the development process of any survey instrument” (p. 229). DeVellis (2012) also recommends use of cognitive interviews as a way to identify items that participants might not understand as well as instructions that might confuse participants. I conducted

cognitive interviews with a small sample of WADA genealogists. They thought aloud about each of the trustworthiness items. Afterwards, I revised, added, and deleted items from the survey. Appendix A catalogues the changes to the trustworthiness items as a result of the cognitive interviews.

3.2.4 Step 4 of Scale Development—Full Administration and Item Analysis

Step 4 involves administering the instrument resulting from Step 3 to a large sample of participants and analysis of the resulting data (see Table 3.4). According to Spector (1992), Bernard (2000), Babbie (2010), Choemprayong and Wildemuth (2009), and DeVellis (2012), the scale developer selects a development sample and administers the instrument to that sample. Then the scale developer conducts item analysis and factor analysis on the resulting data.

-
- Decide on sample population for study
 - Decide on number of participants for study
 - Administer instrument
 - Conduct item analysis (item scoring, item-total correlations, item variances, item means, and standard deviations)
 - Decide upon criteria for retaining and discarding items (if necessary)
 - Examine item content closely in drawing conclusions about what is being measured
 - Conduct exploratory or confirmatory factor analysis
-

Table 3.4 Step 4 of Scale Development—Full Administration and Item Analysis
(Bernard, 2000; Choemprayong & Wildemuth, 2009; DeVellis, 2012; Spector, 1992).

Sample Population: According to Spector (1992), Step 4 requires that the item pool instrument resulting from Step 3 be administered to a sample of respondents.

Choemprayong and Wildemuth (2009) contend that scale developers should

administer the instrument “in conditions that are as similar to the real data collection conditions as possible” in terms of the sampling plan and data collection procedures (p. 283). Similarly, Spector (1992) states, “it is helpful if the respondents are as representative as possible of the ultimate population for which the [instrument] is intended,” but he also tempers his statement by saying that “this is not always possible, and many [instruments] are developed initially on college students because they are readily available” (p. 29). Ultimately, the decision of which participants should be involved in the development of a scale depends upon the feasibility of capturing those respondents. In this study, actual users of digitized archival documents, i.e., genealogists, comprised the sample of participants for Step 4.

Number of Participants: A scale developer must also decide how many participants to include in Step 4. Some authors of scale development texts recommend a specific number or range of participants. For example, Nunnally (1978) recommends 300 people as an adequate number of respondents. Spector (1992) and Bernard (2000) recommend a range of 100 to 200 participants. According to DeVellis (2012), the rationale for having large samples of respondents is to try to “eliminate subject variance as a significant concern” (p. 102). Less prescriptive than Spector (1992) and Bernard (2000) about the number of respondents a scale developer should have during Step 4, DeVellis (2012) does not recommend a certain number of respondents, but warns that having too small of a sample is problematic in many respects. I exceeded Spector’s (1992) recommendation of 100 to 200 participants

during this study so as to avoid data analysis problems that could arise from having too few study participants.

Item Analysis: After administering the instrument to a sample population, a scale developer evaluates several characteristics of individual items. Item analysis includes initial evaluation of items' performance (i.e., item variances, item-total correlations, item means, and item standard deviations) (DeVellis, 2012; Spector, 1992). Item analysis also includes factor analysis (DeVellis, 2012; Spector, 1992).

Item Variances: Assessment of item variances involves looking at the range of responses for each item. Items for which participants provide the broadest range of responses are most capable of discriminating against participants with different levels of the construct being measured (DeVellis, 2012).

Item-Total Correlations: Item-total correlations represent the correlation of a particular item with all other items in an item set (DeVellis, 2012). Items with high item-total correlations are considered most useful in measuring the concept under investigation during a scale development project (DeVellis, 2012). According to DeVellis (2012), there are two types of item-total correlations: uncorrected and corrected. Uncorrected correlations compare the correlation of the item to the set of items including that item. In contrast, corrected item-total correlations compare the correlation of the item to the set of items without including that item. Essentially, the uncorrected item-total correlation counts any particular item twice thereby

inflating reliability of any individual item's item-total correlation. Given this, I calculated corrected item-total correlations during this study. In addition, I compared all item-total correlations against Nunnally's (1978) benchmark of .30, retaining all items above this benchmark. Section 5.2.2 presents results of these tests.

Items' Means and Standard Deviations: According to DeVellis (2012), another favorable attribute of an item is having "a mean close to the center of the range of possible scores" (p. 107). For example, if using a 7-point scale, a mean of 4 is ideal. According to DeVellis, a mean of 4 would indicate that participants used a range of responses to arrive at that mean. In actuality, there are two ways to arrive at a mean which lies at the midpoint of possible responses to an item: 1) people use a range of responses (i.e., some low responses and some high responses), such that the average response is the midpoint of possible responses, or 2) everyone provides the same rating, given that rating is the midpoint of possible responses. DeVellis assumes positive implications from the former way of attaining a midpoint mean. Section 5.2.3 presents results of these tests. In addition, DeVellis (2012) recommends examination of items' standard deviations to determine which items are capable of discriminating among individuals with different levels of the construct being measured. Section 5.2.3 presents results of these tests.

Factor Analysis: In factor analysis, scales and subscales are derived from factors. The items that load onto the factors constitute measurement of the phenomenon under investigation. According to DeVellis (2012), factor analysis is “an essential tool in scale development” (p. 158). He suggests that factor analysis can reveal important properties of a scale. In particular, factor analysis can help “determine *empirically* how many constructs, or latent variables, or factors underlie a set of items” (DeVellis, 2012, p. 116; Jacoby, 1991).

There are two types of factor analysis: confirmatory factor analysis and exploratory factor analysis. Researchers perform confirmatory factor analysis when there is prior research on the appropriate course of action to take regarding measurement of a phenomenon (Kline, 2013). For example, researchers perform confirmatory factor analysis when prior research indicates that a phenomenon should be measured using a scale comprised of two or three subscales. In contrast, when there is no prior research on if or how candidate items should be grouped together to measure a phenomenon, exploratory factor analysis is the appropriate course of action (Kline, 2013). In this study, I had no a priori hypothesis about how the trustworthiness items identified during earlier stages of scale development would group together to measure trustworthiness perception. Thus, exploratory factor analysis was the appropriate course of action to take. Section 5.3 provides details regarding the implementation and results of factor analysis in this study.

3.3 Ethical Considerations

Prior to conducting the study, the University of Michigan Health Sciences and Behavioral Sciences Institutional Review Board (IRB-HSBS) reviewed the research design for the dissertation. IRB-HSBS granted my study (i.e., HUM00069195) exempt status under 45 C.F.R. 46.101.(b)(2) because they determined that: 1) the information I would obtain during the study would not be recorded in such a way that the participants could be identified, and 2) the study posed no more than minimal risk to the participants (Protection of Human Subjects, 2009). For these reasons, I did not have to obtain written informed consent from the study participants.

3.4 References

- Babbie, E. R. (2010). *The practice of social research*. Belmont, Calif: Wadsworth Cengage.
- Bernard, H. R. (2000). *Social research methods : Qualitative and quantitative approaches*. Thousand Oaks, Calif: Sage Publications.
- Bollen, K., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin*, *110*(2), 305-314.
- Choemprayong, S., & Wildemuth, B. M. (2009). Developing new measures. In B. M. Wildemuth (Ed.), *Applications of social research methods to questions in information and library science* (pp. 278-293). Westport, Conn.: Libraries Unlimited.
- Collins, D. (2003). Pretesting survey instruments: An overview of cognitive methods. *Quality of Life Research*, *12*(3), 229-238.
- Conway, P. (2010). Modes of seeing: Digitized photographic archives and the experienced user. *American Archivist*, *73*(2), 425-462.
- DeVellis, R. F. (2012). *Scale development: Theory and applications*. Thousand Oaks, Calif.: SAGE.
- Diamantopoulos, A., & Winklhofer, H. M. (2001). Index construction with formative indicators: An alternative to scale development. *Journal of Marketing Research*, *38*(2), 269-277.
- Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, *51*(4), 327.

- Jacoby, W. G. (1991). *Data theory and dimensional analysis*. Newbury Park, Calif.: Sage Publications.
- Kim, Y. (2009). Validation of psychometric research instruments: The case of information science. *Journal of the American Society for Information Science and Technology*, 60(6), 1178-1191.
- Kitzinger, J. (1995). Qualitative research: Introducing focus groups. *British Medical Journal*, 311(7000), 299-302.
- Kline, R. B. (2013). Exploratory and confirmatory factor analysis. In Y. Petscher, C. and the social sciences (pp. 171-207). New York: Routledge.
- Liu, Z. (2004). Perceptions of credibility of scholarly information on the web. *Information Processing & Management*, 40(6), 1027-1038.
- Liu, Z., & Huang, X. (2005). Evaluating the credibility of scholarly information on the web: A cross cultural study. *The International Information & Library Review*, 37(2), 99.
- Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.
- Protection of Human Subjects, 45 C.F.R. 46.101.(b)(2) (2009)
- Spector, P. E. (1992). *Summated rating scale construction: An introduction*. Newbury Park, Calif.: SAGE Publications.
- St. Jean, B., Rieh, S. Y., Yakel, E., & Markey, K. (2011). Unheard voices: Institutional repository end-users. *College & Research Libraries*, 72(1), 21-42.
- Yakel, E., Faniel, I., Kriesberg, A., & Yoon, A. (2013). Trust in digital repositories. *International Journal of Digital Curation*, 8(1), 143-156.

Yoon, A. (2014). End users' trust in data repositories: Definition and influences on trust development. *Archival Science*, 14(1), 17-34.

CHAPTER 4¹

RESULTS OF STEPS 1 AND 2 OF SCALE DEVELOPMENT: FOCUS GROUPS STUDY

The purpose of this chapter is to present results of the focus group study. Trustworthiness is the construct of interest in this study. One of the primary uses of focus groups is to gain insight into how a particular group of people conceptualizes a concept (Stewart, Shamdasani, & Rook, 2007). Focus groups were used in this study to find out how members of a designated community conceptualize the construct of trustworthiness.

To address the first research question, “How do members of a designated community conceptualize trustworthiness for documents they find in a digital repository?”, a qualitative study was conducted. Focus group data were collected. Themes were identified as a result of analyzing the focus group data set, which consisted of a combination of focus group interview transcripts and video recordings. These themes represent how members of a designated community

¹ Portions of this chapter first appeared as Donaldson, D.R., & Conway, P. (in press). User Conceptions of Trustworthiness for Digital Archival Documents, *Journal of the Association for Information Science and Technology*.

defined document trustworthiness in the context of the data set. The findings demonstrate the breadth and depth of what is salient in the minds of members of a particular designated community with respect to the concept of trustworthiness for archival documents delivered digitally. Overall, the findings highlight the complexity of the concept, as several themes were identified, but also suggest that some perceptions of document trustworthiness are more important than others.

This chapter relates to Steps 1 and 2 of scale development. As shown in Figure 4.1, both Step 1, Construct Definition, and Step 2, Generating an Item Pool, are highlighted in black. This chapter relates to Step 1 of scale development because the focus groups findings provide scope and clarity to the construct definition of trustworthiness by providing insight into the concept from the perspective of actual users of digitized archival documents. This chapter relates to Step 2 of scale development because the focus groups findings provide insight into how actual users of digitized archival documents talk about the concept of trustworthiness; their statements serve as item pool items because they reflect actual users' terminology regarding the concept of trustworthiness.

The structure of this chapter is as follows. First, I present data regarding the study participants whose responses were used to generate the data for the focus groups. Second, I discuss the themes I identified in the context of the data set. Finally, I conclude with implications of the findings for understanding the concept of trustworthiness from the vantage of user perception. I also discuss the influence of these findings on later stages of the scale development methodology employed during this study.

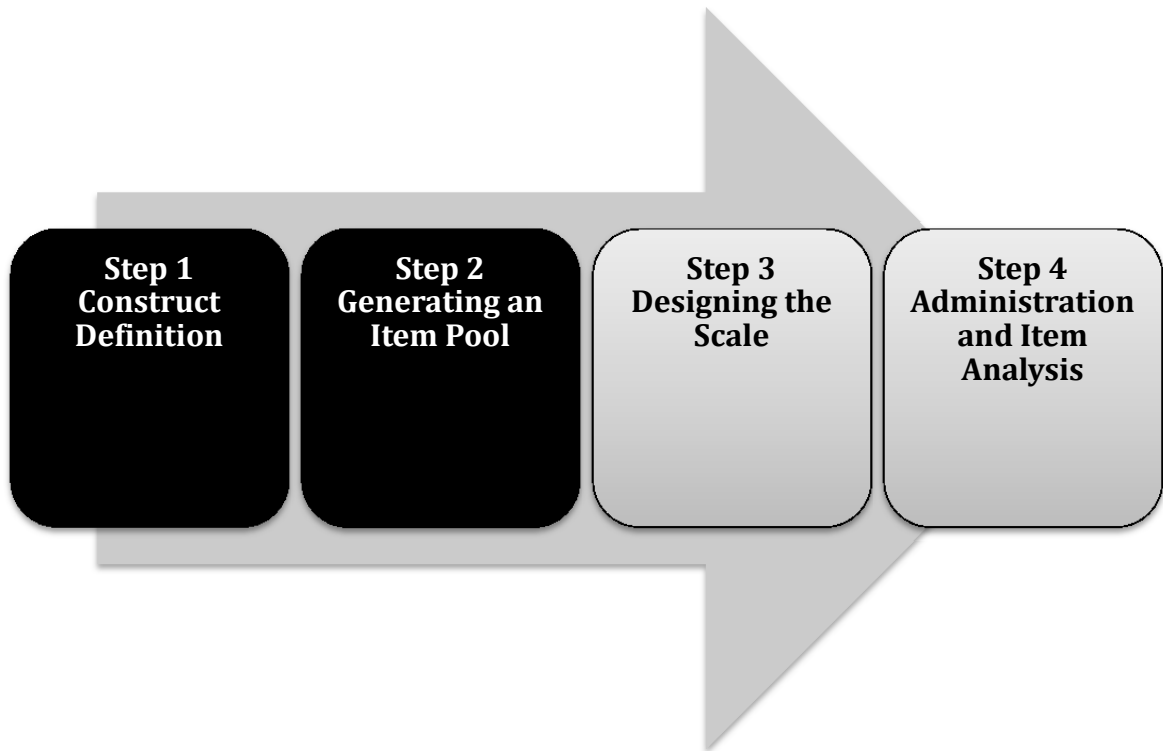


Figure 4.1 Contribution of the Focus Groups Study Findings to Scale Development.

As denoted in black, the focus groups study findings enhance Steps 1 and 2 of scale development. The focus groups study findings provide scope and clarity to the construct definition (Step 1) of trustworthiness perception by providing insight into how actual users of digitized archival documents talk about trustworthiness. The focus groups study findings contribute to generating an item pool (Step 2) of scale development because participants provided statements regarding trustworthiness that were turned into items that reflected actual users' terminology with respect to the concept of trustworthiness perception.

4.1 Study Participants

The dearth of research on the trustworthiness of digital archival documents from the vantage of user perception required data collection from actual users regarding their definition of the concept. According to Stewart et al. (2007), "focus groups ... are particularly useful for exploratory research when rather little is known about the phenomenon of interest. As a result, focus groups tend to be used

very early in a research project” (p. 41). Given the highly exploratory nature of the project, focus groups were the appropriate to conduct before later stages of scale development (see Chapter 5).

Focus groups provide an opportunity to learn how participants talk about the phenomenon of interest (Stewart et al., 2007). Applied to this study, focus groups provided an opportunity to learn about how members of a designated community talk about document trustworthiness.

Focus groups require participants with shared experiences as they pertain to the research questions and topics of interest (Liamputtong, 2011). In this study, the most important experience for participants to have in common was experience using WADA documents. Experience utilizing digital archival documents was a necessary prerequisite for discussing digital archival documents in terms of trustworthiness perception. Without any experience using digital archival documents, it would have been impossible to draw upon those experiences in considering digital archival document trustworthiness, which was the focus of the project as a whole. Prior contact with the Deputy State Archivist of the Washington State Archives confirmed that WADA’s largest designated community of users included genealogists (T.S. Badger, personal communication, March 8, 2013). Thus, an appropriate and representative sample for this study required participants with demographics similar to the larger population of genealogists as well as experience in using digital archival documents.

While there is no sampling frame available to assess genealogists as a population, there are multiple studies of genealogists in archival science and

information science. Most of what is known about genealogists in archival science and information science is based upon samples of predominantly older females (e.g., Case, 2008; Duff & Johnson, 2003; Lucas, 2008; Yakel, 2004; Yakel & Torres, 2007). A profile of demographic characteristics can be derived from studies of genealogists in archival science and information science and used as a proxy for the demographics of the larger population. In this study, participants were asked questions about their age and gender. Afterwards, these findings were compared to findings of previous studies in archival science and information science involving genealogists and were used as a measure of the representativeness of the sample.

In addition to demographic characteristics, the study also required participants with experience using digital archival documents. Verifying the behavior of the study participants beforehand, e.g., observing them as they utilized digital archival documents over time, would have ensured that they had experience using digital archival documents. However, this approach was not practically feasible. Alternatively, asking a set of questions pertaining to participants' WADA document usage was possible. Although this study is limited by no attempts to verify the self-reports, they at least provide some insight into the extent to which the study participants thought they used digital archival documents.

The study participants were asked demographic questions regarding their age and gender as well as questions pertaining to their use of WADA documents. These questions were asked in order to understand participants' demographic characteristics and participants' appropriateness for the study. After discussing the sizes and number of focus groups, participant recruitment, and locations of the

focus groups, this section examines participants' responses to the demographic questions and WADA document use questions, highlighting the trends that were identified as a result of the data analysis.

4.1.1 Size of the Focus Groups and Number of Focus Groups

Both the size of each focus group as well as the total number of focus groups in this study are consistent with recommendations in the focus group literature (Krueger & Casey, 2009; Morgan, 1997; Stewart & Shamdasani, 1990; Stewart et al., 2007). For example, according to Stewart et al. (2007), "most focus groups are composed of six to twelve people. [...] Most focus group applications involve more than one group, but seldom more than three or four groups" (p. 58). In this study, the first focus group included 7 participants, the second focus group contained 7 participants, and the third focus group included 8 participants.

4.1.2 Participant Recruitment

I recruited study participants using two methods. First, WADA staff sent messages on their researcher listservs inviting users to participate in the focus groups, and to contact me directly via email if they were interested in participating (For a sample recruitment email, see Appendix B). I successfully recruited a total of five participants using this method.

Second, WADA staff enabled a pop-up message inviting users to participate in the focus groups to everyone who visited the WADA homepage. This pop-up message appeared as a window overlaying WADA's homepage as opposed to a separate window. Therefore WADA users with pop-up blockers also received invitations to participate in this study. The message included a link to a participant recruitment form administered via Qualtrics, a web-based tool for building and administering surveys (See Appendix C). The form asked participants which physical location they were closest to (e.g., Cheney, WA; Olympia, WA; or Bellevue, WA) and their primary reason for visiting the WADA website (e.g., business, genealogy, historical research, local government, title searches, or other). In addition, the form asked participants to provide their names and email addresses and/or phone numbers so that I could contact them with details about the focus groups. Within a 24-hour period of a participant completing the focus group invitation form, I emailed and/or called each person. A total of 121 people filled out the survey between Wednesday, May 29, 2013 and Saturday June 8, 2013. After screening for people who were genealogists, were within driving distance of one of the focus group locations, and were available on one of the planned days/times of the focus groups, I successfully recruited the remaining eighteen participants using this method.

4.1.3 Locations

Since WADA users are geographically dispersed, each focus group took place at a different location to make it easier for users to participate in the study. I conducted the focus group sessions in private conference rooms within three separate locations over the period of three days during the month of June 2013. The first focus group (FG1) took place at the Washington State Archives – Digital Archives in Cheney, WA. The second focus group (FG2) took place at the Washington State Archives in Olympia, WA. The third focus group (FG3) took place at the Puget Sound Regional Archives in Bellevue, WA.

4.1.4 Demographic information

Prior to the focus groups, participants took pre-surveys that asked participants demographic questions pertaining to their age and gender (see Appendix D). Answers to these questions are discussed below and compared to studies of genealogists in archival science and information science as a measure of the representativeness of the sample of respondents who participated in this study.

4.1.4.1 Participants by Age

The participants were predominantly older. Nineteen participants indicated that they were 50 years of age or older. The remaining 3 participants indicated that they were between the ages of 30 and 39.

Overall, these results provide empirical support for the representativeness of the sample. Most of what is known about genealogists in archival science and information science is based upon samples of older participants (e.g., Case, 2008; Duff & Johnson, 2003; Lucas, 2008; Yakel, 2004; Yakel & Torres, 2007). While demographic characteristics of the population of genealogists are unknown, what can be said with certainty is that results of this study's participants' self-reports are consistent with studies in archival science and information science regarding the average age of genealogists who agree to participate in studies.

4.1.4.2 Participants by Gender

The focus groups were predominantly female. The first focus group included 6 females and 1 male. The second focus group included 5 females and 2 males. The third focus group included 7 females and 1 male. In total, 18 females and 4 males participated in the study.

Overall, these results provide additional empirical support for the representativeness of the sample. Most of what is known about genealogists in archival science and information science is based upon samples of predominantly female participants (e.g., Duff & Johnson, 2003; Lucas, 2008; Yakel, 2004; Yakel & Torres, 2007). In his review of recent studies on genealogists in information science,

Case (2008) concluded that genealogy is a hobby that is perhaps predominantly female. Thus, while demographic characteristics of the population of genealogists are unknown, what can be said with certainty is that results of this study's participants' self-reports are consistent with studies in archival science and information science regarding the breakdown of participants by gender.

4.1.5 Participants' WADA Document Usage

Prior to the focus groups, participants took pre-surveys that asked participants questions about their primary reason for using WADA documents, frequency of use of WADA documents, and their trust in WADA documents. In addition, during the focus groups, I asked participants about the nature of the documents they typically use and their purposes for using them. The following presents the data that were collected regarding these questions.

4.1.5.1 Primary Reason for Using WADA Documents

All participants reported using WADA documents for the purpose of conducting genealogical research. Eighteen participants reported that their purpose for using WADA documents involved tracing their family lineage, including who in their families married whom (e.g., FG1-S5). Participants reported conducting research either to verify information about their family history that other family members provided (e.g., FG2-S6), or to update or otherwise continue working on

family trees to locate and identify as many proverbial branches as possible (e.g., FG1-S3). Still others reported using WADA documents to do research on behalf of libraries (e.g., FG1-S6), museums (e.g., FG1-S7), genealogical societies (e.g., FG1-S2, FG1-S4, FG2-S3, FG2-S4, FG2-S7, FG3-S1), and heritage societies (e.g., FG3-S8). Participants' responses indicate that their use purposes were not mutually exclusive; some participants conducted genealogical research for themselves and also on behalf of others.

Overall, these results indicate additional empirical support regarding the representativeness of the sample. Prior studies on genealogists' information seeking behavior indicate that they conduct research to recount a more complete family history for themselves or others (Darby & Clough, 2013; Duff & Johnson, 2003; Yakel, 2004; Yakel & Torres, 2007). While prior studies in archival science and information science have not focused on genealogists' use of digital archival documents specifically, results from this study suggest that the information needs of genealogists who do rely upon digital archival documents are similar to other genealogists who may or may not rely upon digital archival documents to fulfill their information needs.

4.1.5.2 Frequency of Using WADA Documents

Overall, most of the participants reported using WADA documents frequently. Fifteen or 68% percent of the participants reported using WADA documents either weekly (n=12) or daily (n=3). Five participants indicated that they

used WADA documents monthly. By contrast, only 2 participants reported that they used WADA documents once or twice within the last year.

Given the research question this study was designed to address, it was important that the study participants had experience utilizing digital archival documents. Results from these participants' self-reports suggest that most of the participants had enough experience with WADA documents in order to be in a position to discuss WADA documents in terms of perceived trustworthiness.

At one level, these data cannot be rightfully compared with results from other studies in archival science and information science. Prior studies have not typically asked participants specifically about their frequency of digital archival document use. Instead, researchers have typically asked participants questions concerning how many years they have conducted genealogical research in general (e.g., Darby & Clough, 2013; Lucas, 2008; Yakel, 2004; Yakel & Torres, 2007), their use of online resources (Darby & Clough, 2013; Lucas 2008), and their use of county archives or record office websites (Darby & Clough, 2013). At another level, the data regarding participants' frequency of WADA document use can be compared to results of other studies in archival science and information science. It is plausible that in these previous studies at least some of the participants' genealogical research, use of online resources, and use of county archives websites involved use of digital archival documents. Considering this plausibility, the data collected in this study are consistent with data collected in previous studies regarding frequency of use of digital archival documents.

4.1.5.3 Frequency of Use by Document Type

All participants reported using birth, death, and/or marriage records. In addition, participants reported using criminal records (e.g., FG1-S6), military records (e.g., FG1-S2, FG1-S3, FG1-S4), family history records (e.g., FG1-S4), cemetery records (e.g., FG1-S6; FG3-S6), state census records (e.g., FG1-S3), and frontier justice records (e.g., FG1-S2, FG1-S6, FG1-S7). According to the participants, the marriage records they used included marriage applications and licenses (e.g., FG1-S2, FG1-S6). Participants also indicated that they use newspapers (e.g., FG1-S1) to look up information on their own or others' relatives, particularly when other sources of information on those relatives are lacking.

Overall, these results provide additional empirical support for the appropriateness of the sample for the study. To a certain extent, these findings cross-validate other findings pertaining to participants' frequency of use regarding WADA documents. Specifically, these findings provide insight into which type of documents participants reported using. Had the participants indicated that they used a document type not preserved by WADA, this would have suggested error on the part of the participants or underscored their inappropriateness for the study.

While findings from prior studies suggest that genealogists utilize birth, death, marriage, census, and land records, these studies focus less on whether or not those documents are digital (e.g., Case, 2008; Darby & Clough, 2013; Duff & Johnson, 2003; Lucas, 2008; Yakel, 2004; Yakel & Torres, 2007). This makes the findings from this study difficult to compare with results from previous studies. While specific questions were not asked about participants' use of digital archival documents in

these studies, it is possible that the participants used digitized birth, death, marriage, census, and land records. If so, the findings from this study align with previous studies regarding the representativeness of the sample regarding digital archival document use.

4.1.5.4 Trust in WADA Documents

All participants reported that they either trusted WADA documents “very strongly” (19 or 86%) or “somewhat strongly” (3 or 14%). Given that respondents in this study were able to indicate that they trust WADA documents, they were likely in a position to articulate what it means to them for a WADA document to be trustworthy.

These results are difficult to compare with prior studies because prior studies of genealogists in archival science and information science have not focused on issues of document trustworthiness. However, the most recent “how-to” guides for genealogy underscore the importance of verifying the trustworthiness of documents found during the course of conducting research (e.g., Crowe, 2011), which directly relates to the focus of the overall research project. Since “how-to” guides represent the largest body of knowledge for online genealogy (Veale, 2005), they serve as an indication of the importance of the topic of digital archival document trustworthiness perception.

4.1.6 Study Participants Section Summary

Considering the data presented in Sections 4.1.4.1 through 4.1.5.4 creates a profile of the study participants, which suggests that they comprise an appropriate and representative sample. The study participants:

- Are mostly female and 50 years of age or older,
- Use WADA documents frequently (e.g., weekly or daily),
- Utilize specific types of documents, including digitized marriage, death, birth, census, and land records, and
- Trust WADA documents.

This composite profile suggests that it was possible for the study participants to discuss document trustworthiness during focus groups based on the quantity and quality of their reported experiences with WADA documents. This was critical for this study, which focuses on trustworthiness perception regarding digital archival documents.

4.2 Focus Groups Results: Identified Themes

I conducted a series of focus groups to understand how members of a designated community conceptualize trustworthiness for archival documents delivered digitally. To generate data during the focus groups on the topic of document trustworthiness, the moderator asked the participants three questions related to: (1) participants' perceptions of trustworthiness for the documents they

find using WADA, (2) participants' descriptions of documents they have found in WADA that they think are trustworthy, (3) and any circumstances in which participants would question the trustworthiness of documents they encountered while using WADA (see Focus Groups Guide in Appendix E). Throughout the focus groups, the moderator took on the persona of the "enlightened novice," which according to Krueger (1998), "is bright but lacks knowledge (or gives the impression of lacking knowledge) in the area of expertise possessed by the participants" in order to "get participants to explain more about the topic and causes" (p. 46). Applied to this study, the moderator portrayed the enlightened novice so that participants could explain more about what the concept of document trustworthiness meant to them and why.

This section focuses on consensus that formed among participants regarding particular conceptualizations of document trustworthiness during the focus groups. I identified these conceptualizations as themes as a result of thematic analysis—"a process for encoding qualitative information" (Boyatzis, 1998, p. vi). According to Braun and Clarke (2006), "a theme captures something important about the data in relation to the research question, and represents some level of patterned response or meaning within the data set," and, "the 'keyness' of a theme is not necessarily dependent on quantifiable measures—but rather on whether it captures something important in relation to the overall research question" (p. 82). They also add that "researcher judgment is necessary to determine what a theme is" (p. 82), because the amount of times a theme is or is not mentioned by any particular participant or group of participants may or may not be what determines the importance of a

theme, as much as the identification of a theme in relation to the purposes of the research and the research questions. Part of thematic analysis involves determining the prevalence of themes. According to Braun and Clarke (2006), while there is no right or wrong method for determining the prevalence of a theme during thematic analysis, researchers should make explicit the ways in which they define prevalence.

In this study, I followed procedures for thematic analysis outlined in Braun and Clarke (2006), including: (1) familiarizing myself with the data, (2) generating initial codes, (3) searching for themes, (4) reviewing themes, (5) defining and naming themes, and (6) producing a report of the focus groups. In addition, I used micro-interlocutor analysis to operationalize the prevalence of the themes I identified during thematic analysis both within and across the focus groups. During micro-interlocutor analysis, “meticulous information about which participant responds to each question, the order in which each participant responds, response characteristics, the nonverbal communication used, and the like is collected, analyzed, and interpreted” (Onwuegbuzie, Dickinson, Leech, & Zoran, 2009, p. 1). Applying micro-interlocutor analysis to this study, I:

- Provide verbatim statements made by focus group participants,
- Delineate information about the number of members who appeared to be part of the consensus for each theme, and
- Specify the number of members who appeared to represent a dissenting view (if any) as well as how many participants did not appear to express any view at all.

Micro-interlocutor analysis was conducted to avoid overstating the amount of consensus among participants within and across the focus groups for any particular theme while also being transparent about issues of nonresponse. In addition, I applied micro-interlocutor analysis to the transcribed verbal expressions by extracting from the video recording of each session information about participants' nonverbal communication—including chronemic expressions (i.e., use of pacing of speech and length of silence in conversation), paralinguistic expressions (i.e., all variations in volume, pitch, and quality of voice), and kinesic expressions (i.e., body movements or postures)—and incorporating these data alongside the verbal data (Onwuegbuzie et al., 2009, p. 10). As Stewart et al. (2007) point out, “information represented by nonverbal responses of focus group participants can be useful and complement the information provided via verbal channels of communication” (p. 31).

To ensure descriptive validity of the data, i.e., that the data actually reflect the participants' statements, I checked the transcripts against the videorecordings of the focus groups for accuracy and consistency (Maxwell, 1992).

To ensure reliability of the coding of the identified themes, I calculated inter-rater reliability, i.e. the extent to which two or more coders code data similarly. I hired two University of Michigan School of Information graduate students to code the entire focus groups dataset. These graduate students both had prior qualitative research experience as well as prior experience using NVivo 10.0, the qualitative research data software program that I used to analyze the focus groups data.

Afterwards, I calculated Cohen's kappa to assess inter-rater reliability (Cohen, 1960). Cohen's kappa is a statistical measure the proportion of agreement between two coders corrected by the amount of agreement between coders that could be expected by chance (Multon, 2010). I achieved a Cohen's kappa of .65 with the first coder. I achieved a Cohen's kappa of .61 with the second coder. Both of the hired coders achieved a Cohen's kappa of .62 with each other. Since Cohen's kappa coefficients of .50 or greater are acceptable (Multon, 2010), I considered the reliability of the qualitative data analysis in this study as highly reliable. The codebook for the focus groups data is in Appendix F.

I identified ten themes as a result of thematic analysis and micro-interlocutor analysis: authenticity, accuracy, first-hand or primary evidence, believability, validity, proper form, stability, legibility, coverage, and objectivity. This section utilizes passages from the focus group transcripts to provide illustrative examples of each theme. In addition, I present and discuss selective examples of nonverbal communication pertaining to the themes alongside the verbal data. What follows is a discussion of each theme rank ordered from the theme for which participants expressed the most support to the theme for which participants expressed the least support.

4.2.1 Authenticity

The theme for which the most participants expressed support was the theme of authenticity. In each focus group, participants articulated this particular conceptualization in the form of a question. For example, during the second focus

group, FG2-S5 asked, “it could have been fake? Is that what you’re suggesting or asking?,” and during the third focus group, FG3-S2 said, “now, if you’re asking, is that a forgery?” These participants’ questions about what the moderator meant by the term trustworthiness suggest that they conceptualized trustworthiness in terms of a document’s perceived authenticity, i.e., is it fake?

Table 4.1 presents results of micro-interlocutor analysis regarding the authenticity theme. As shown in Table 4.1, during the first focus group, four participants provided significant statements in agreement with the authenticity theme, and three participants provided expressions of support for the authenticity theme. During the second focus group, five participants provided significant statements in agreement with the authenticity theme, and two participants provided expressions of support for the authenticity theme. During the third focus group, three participants provided significant statements in agreement with the authenticity theme, and five participants provided expressions of support for the authenticity theme. Across the three focus groups, all 22 participants either made strong statements in support of this theme or expressed agreement.

	FG 1	FG 2	FG3
SA	4	5	3
A	3	2	5
SD	0	0	0
D	0	0	0
NR	0	0	0

Table 4.1 Micro-interlocutor Analysis Results for Authenticity Theme.

FG = Focus Group.

SA = Provided significant statement or example suggesting agreement.

A = Indicated agreement (i.e., verbal or nonverbal).

SD = Provided significant statement or example suggesting dissent.

D = Indicated dissent (i.e., verbal or nonverbal).

NR = Did not indicate agreement or dissent (i.e., non-response).

Participants provided several statements in support of authenticity theme, including FG1-S3 who remarked, “I don’t think I ever would question it. I mean, the actual, that that’s not the real thing.” FG1-S6 added, “it may not be the one I’m looking for, but they’re all the real thing.” When the moderator asked why as a follow-up question, participants’ responses indicated that they were considering an authenticity conceptualization of trustworthiness. FG1-S6 responded, “the records are so old that it seems silly to have a fake one in there.” During similar conversation in the second focus group, FG2-S5 asked, “why would we think it was a fake?” In a similar discussion during the third focus group, FG3-S4 responded, “I would think they’d be the real documents,” expressing her authenticity conceptualization of trustworthiness.

4.2.2 Accuracy

The second theme for which participants expressed support was the theme of accuracy. Kelton et al. (2008) define accuracy as “the extent to which information is free from error” (p. 370). Table 4.2 presents results of micro-interlocutor analysis regarding the accuracy theme. Findings indicate that the study participants consistently and strongly conceptualized trustworthiness in terms of accuracy. As shown in Table 4.2, during the first focus group, four participants provided significant statements in agreement with the accuracy theme, one participant provided expressions of support for the accuracy theme, and two participants did not provide a response. During the second focus group, three participants provided

significant statements in agreement with the accuracy theme, three participants provided expressions of support for the accuracy theme, and one participant did not provide a response. During the third focus group, four participants provided significant statements in agreement with the accuracy theme, and four participants did not provide a response. Across the three focus groups, 15 of 22 participants either made strong statements in support of this theme or expressed agreement.

	FG 1	FG 2	FG3
SA	4	3	4
A	1	3	0
SD	0	0	0
D	0	0	0
NR	2	1	4

Table 4.2 Micro-interlocutor Analysis Results for Accuracy Theme.

FG = Focus Group.

SA = Provided significant statement or example suggesting agreement.

A = Indicated agreement (i.e., verbal or nonverbal).

SD = Provided significant statement or example suggesting dissent.

D = Indicated dissent (i.e., verbal or nonverbal).

NR = Did not indicate agreement or dissent (i.e., non-response).

Participants provided several statements in support of accuracy theme. For example, when asked what adjectives participants would use to describe a document they thought was trustworthy, both FG1-S3 and FG1-S4 said, “accurate.” In addition, FG1-S3 and FG1-S4 provided examples suggesting that they conceptualize inaccurate information as untrustworthy information. For example, FG1-S3 stated:

I think that maybe like relatives that have been married six times might put that they’re Miss somebody or other, but that’s according to them at the time. That’s where mine would be untrustworthy of it. (FG1-S3)

Hence, according to FG1-S3, a woman who has been married six times and reports that she is a Miss would be reporting incorrect information about her marital status, rendering that information untrustworthy.

4.2.3 Primary or First-hand Evidence

The third theme for which participants expressed support was the theme of primary or first-hand evidence. Participants conceptualized trustworthiness as the extent to which a document is primary or first-hand. In particular, participants suggested that first-hand or primary documents are more trustworthy than second-hand or other documents. Participants defined primary documents as documents that were written during the time period of the events that they were about as opposed to documents that serve as accounts of what happened that were written at a later time. Other participants defined primary or first-hand as having been written by the actual person(s) the documents were meant to represent. The assumption underlying the high value participants placed on primary documents is that, the closer the document is to the original or the actual event, the less likely the error or the less likely important information has been omitted, changed, or otherwise altered.

Table 4.3 presents results of micro-interlocutor analysis regarding the primary or first-hand evidence theme. As shown in Table 4.3, during the first focus group, two participants provided significant statements suggesting the importance of a document being first-hand or primary, two participants provided expressions in

support of the importance of a document being first-hand or primary, with the remaining three participants not providing any response. During the second focus group, three participants provided significant statements that primary sources are more trustworthy than secondary sources or other sources, two participants provided expressions in support of primary sources as more trustworthy than secondary sources or other sources, with the remaining two participants not providing any response. During the third focus group, four participants provided significant statements in support of primary sources as more trustworthy than secondary sources or other sources, with the remaining five participants not providing any response. Across the three focus groups, 13 of 22 participants either made strong statements in support of this theme or expressed agreement.

	FG 1	FG 2	FG3
SA	2	3	4
A	2	2	0
SD	0	0	0
D	0	0	0
NR	3	2	4

Table 4.3 Micro-interlocutor Analysis Results for Primary or First-hand Evidence Theme.

FG = Focus Group.

SA = Provided significant statement or example suggesting agreement.

A = Indicated agreement (i.e., verbal or nonverbal).

SD = Provided significant statement or example suggesting dissent.

D = Indicated dissent (i.e., verbal or nonverbal).

NR = Did not indicate agreement or dissent (i.e., non-response).

Participants provided several statements in support of the primary or first-hand evidence theme. The following passage illustrates the theme in the context of a discussion among five participants during the second focus group:

FG2-S1—To me, it’s the closest to the source. Closest to the original (0.2)

((gestures air quotes)) “action” that happened.

FG2-S2—((Nods head in agreement)) Yeah.

FG2-S1—If it’s a birth, you want the hospital. And then you want maybe the doctor, and then, maybe, you know, a neighbor, and maybe a census record. If it’s a marriage, you want the marriage document, and then maybe the listing in the book that says, they were married on this date, and then maybe a diary that says they were married, you know. So it’s the closest document to the action, is the most trustworthy.

FG2-S4—I think it’s the document versus maybe say a ledger. You know when you’ve got a ledger of, you know, Suzy had married, and you know, it’s just this big long ledger, whereas, where FG2-S1 is saying, it’s from a hospital, or the doctor, you know, it’s maybe a little bit more trustworthy because whoever is doing the ledger is getting the information second- or third-hand.

FG2-S5—((Nods head in agreement)).

FG2-S1—The less times it’s transcribed-

FG2-S5—((Nods head in agreement)) (overlapping) Right.

FG2-S4—((Points at FG2-S1)) Yes.

FG2-S3—((Nods head in agreement and smiles)).

FG2-S1—the less chance there is to have-

FG2-S5—(Overlapping) Every time it’s recopied or recopied or recopied.

FG2-S1—An error in transcription.

In the passage above, participants used two chronemic expressions for emphasis, and eight kinesic expressions to either draw contrasts or indicate agreement. For example, FG1-S1 used two chronemic expressions to draw attention to the idea that the closer a document is to the actual event that happened, the more trustworthy

she perceives it. Specifically, FG2-S1 took a two-second pause before saying “action,” and she said “action” slower than the other words in her statement to draw attention to the fact that a document’s proximity to the primary action is what influences her trustworthiness perceptions. While pausing, FG2-S1 used hand gestures to put “action” in quotes to further emphasize the importance of a document’s proximity to the action it is about on her trustworthiness perception. In the passage above, the most common kinesic expressions were to indicate agreement with what others were saying. For example, throughout the passage, FG2-S2, FG2-S3, and FG2-S5 nodded their heads in agreement with statements FG2-S1 made about the influence of how primary a document is on her trustworthiness perception. FG2-S4 even pointed at FG2-S1 and said, “yes,” to indicate agreement with one of her statements.

The other use of kinesic expressions during this passage was to illustrate the distinction among primary, secondary, and tertiary sources. Specifically, FG2-S1, FG2-S2, and FG2-S5 all talked about the difference between primary and other sources, and when they did, they made hand gestures closer to themselves when talking about primary sources and deliberately made hand gestures further away from themselves when discussing what they considered secondary or tertiary sources. For example, when FG2-S4 said that he would consider a document more trustworthy than a ledger because, according to him, “whoever is doing the ledger is getting the information second- or third-hand,” he made hand gestures closer to himself when talking about documents and further away from himself when talking about ledgers. Similarly, when FG2-S4 said “every time it’s recopied or recopied or

recopied” she twirled her left hand each time she said recopied, and each time she moved her twirling left hand further away from herself to spatially draw the distinction among primary, secondary, and tertiary instantiations of a document as she was talking.

4.2.4 Believability

The fourth theme for which participants expressed support was the theme of believability. Kelton et al. (2008) define believability as “the extent to which the information appears to be plausible” (p. 370). Kelton and colleagues also treat believability as synonymous with credibility, which is consistent with other researchers who have conducted extensive research on credibility (e.g., Tseng & Fogg, 1999; Fogg, 2003).

Table 4.4 presents results of micro-interlocutor analysis regarding the believability theme. As shown in Table 4.4, during the first focus group, two participants provided significant statements in support of the believability theme, with the remaining five participants not providing any response. During the second focus group, five participants provided significant statements in support of the believability theme, with the remaining two participants not providing any response. During the third focus group, only one participant provided significant statements in support of the believability theme, with the remaining seven participants not providing any response. Across the three focus groups, 8 of 22 participants made strong statements in support of this theme.

	FG 1	FG 2	FG3
SA	2	5	1
A	0	0	0
SD	0	0	0
D	0	0	0
NR	5	2	7

Table 4.4 Micro-interlocutor Analysis Results for Believability Theme.

FG = Focus Group.

SA = Provided significant statement or example suggesting agreement.

A = Indicated agreement (i.e., verbal or nonverbal).

SD = Provided significant statement or example suggesting dissent.

D = Indicated dissent (i.e., verbal or nonverbal).

NR = Did not indicate agreement or dissent (i.e., non-response).

Participants provided several statements in support of the believability theme. For example, during the second focus group, FG2-S2 revealed her credibility conceptualization of document trustworthiness as she explained her process for formulating trustworthiness perceptions:

My first impression, in a split second, 'cause we do these things in milliseconds, is that if I see an original document, the very first thing is I give it a great deal of credibility before I read it. . . . And then I judge it accordingly, but I think, just because I have an original document means a whole lot more to me, and I give it a great deal of credibility. If it's the only thing I have, I give it tremendous credibility, 'cause sometimes you can't compare it to anything. (FG2-S2)

For FG2-S2, the fact that a WADA document is a digitized copy of an original physical document is sufficient grounds for perceiving it as tremendously credible. As FG2-S2 made this statement, FG2-S1, FG2-S3, and FG2-S6 nodded their heads in agreement (kinesic) with FG2-S2, suggesting consensus among multiple participants in the focus group regarding a credibility document trustworthiness conceptualization.

4.2.5 Validity

The fifth theme for which participants expressed support was the theme of validity. Kelton et al. (2008) define validity as “the use of responsible and accepted practices” such as “the soundness of the methods used, the inclusion of verifiable data, and the appropriate citation of sources” (p. 370).

Table 4.5 presents results of micro-interlocutor analysis regarding the validity theme. As shown in Table 4.5, during the second focus group, three participants provided significant statements in support of the validity theme. Two participants provided expressions in support of the validity theme, with the remaining two participants not providing any response. During the third focus group, two participants provided significant statements in support of the believability theme, with the remaining six participants not providing any response. During the first focus group, I did not identify the theme of validity. Across the three focus groups, 7 of 22 participants either made strong statements in support of this theme or expressed agreement.

	FG 1	FG 2	FG3
SA	0	3	2
A	0	2	0
SD	0	0	0
D	0	0	0
NR	0	2	6

Table 4.5 Micro-interlocutor Analysis Results for Validity Theme.

FG = Focus Group.

SA = Provided significant statement or example suggesting agreement.

A = Indicated agreement (i.e., verbal or nonverbal).

SD = Provided significant statement or example suggesting dissent.

D = Indicated dissent (i.e., verbal or nonverbal).

NR = Did not indicate agreement or dissent (i.e., non-response).

Participants provided several statements in support of the validity theme. For example, during the third focus group, FG3-S2 asked, “is it valid?” to describe how she would define document trustworthiness. During the second focus group, FG2-S5 articulated her validity conceptualization of trustworthiness by explaining that a document is still trustworthy regardless of whether its content is accurate:

Even [with] a document, an original document, the person that wrote the information on it could have misheard or got distracted or whatever. I mean that information can be wrong, but it’s still a valid document. But the information, which is a separate thing, could be incorrect. (FG2-S5)

According to FG2-S5, the person recording the information could have been the appropriate person, and could have been receiving the information from the appropriate source; hence the document was created using responsible and accepted practices. And yet, that person could have also misheard or got distracted, and as a result, recorded the information inaccurately. In other words, according to FG2-S5, the document would still be valid because the appropriate person recorded the information from the proper source, even if the person recorded the information in error.

4.2.6 Proper Form

The sixth theme for which participants expressed support was the theme of proper form. Specifically, participants conceptualized document trustworthiness in terms of a document’s perceived proper form.

Table 4.6 presents results of micro-interlocutor analysis regarding the proper form theme. As shown in Table 4.6, during the second focus group, five participants provided significant statements in support of the proper form theme, one participant provided expressions in support of the proper form theme, with the remaining participant not providing any response. During the first or third focus group, I did not identify the proper form theme. Across the three focus groups, 6 of 22 participants either made strong statements in support of this theme or expressed agreement.

	FG 1	FG 2	FG3
SA	0	5	0
A	0	1	0
SD	0	0	0
D	0	0	0
NR	0	1	0

Table 4.6 Micro-interlocutor Analysis Results for Proper Form Theme.

FG = Focus Group.

SA = Provided significant statement or example suggesting agreement.

A = Indicated agreement (i.e., verbal or nonverbal).

SD = Provided significant statement or example suggesting dissent.

D = Indicated dissent (i.e., verbal or nonverbal).

NR = Did not indicate agreement or dissent (i.e., non-response).

Participants provided some statements in support of the proper form theme.

For example, the following interaction among the focus group participants

illustrates this theme:

FG2-S2—I can't imagine anybody not believing that a document wouldn't be trustworthy in that millisecond that you see it's an actual picture of that document.

FG2-S6—And that means you would have to be familiar with what the original-

FG2-S3—((Nods head in agreement)).

FG2-S6—document should look like.

FG2-S3—Exactly. ((nods head in agreement)).

FG2-S2—Well, yeah you'd have to, but I think most people figure that, even in a millisecond, you have a feeling of what an official document should look like.

FG2-S6—((Nods head in agreement)) Yeah.

FG2-S1—You want it to have the proper form-

FG2-S3—(Overlapping) You have some image of it. Yeah, that there would be a proper form.

4.2.7 Stability

The seventh theme for which participants expressed support was the theme of stability. Kelton et al. (2008) define stability as implying that “the information is persistent, in both its presence and its contents” (p. 370). In particular, they define stability as being insusceptible to alteration.

Table 4.7 presents results of micro-interlocutor analysis regarding the stability theme. As shown in Table 4.7, during the first focus group, one participant provided significant statements in support of the stability theme, with the remaining six participants not providing any response. During the third focus group, three participants provided significant statements in support of the stability theme, one participant provided expressions in support of the stability theme, with the remaining four participants not providing any response. During the second focus

group, I did not identify the stability theme. Across the three focus groups, 5 of 22 participants either made strong statements in support of this theme or expressed agreement.

	FG 1	FG 2	FG3
SA	1	0	3
A	0	0	1
SD	0	0	0
D	0	0	0
NR	6	0	4

Table 4.7 Micro-interlocutor Analysis Results for Stability Theme

FG = Focus Group.

SA = Provided significant statement or example suggesting agreement.

A = Indicated agreement (i.e., verbal or nonverbal).

SD = Provided significant statement or example suggesting dissent.

D = Indicated dissent (i.e., verbal or nonverbal).

NR = Did not indicate agreement or dissent (i.e., non-response).

Participants provided some statements in support of the stability theme. For example, during the first focus group, FG1-S4 asked, “Are you hinting to think it was doctored up or something?”, in response to the moderator’s question about whether anyone had ever questioned the trustworthiness of a WADA document. FG1-S4’s question about what the moderator meant by trustworthiness indicated that one way she conceptualized trustworthiness was in terms of whether or not a document has been altered. During the third focus group, FG3-S6 stated that she would never suspect that WADA would alter one of its documents:

I just cannot see [WADA staff] falsifying documents, because this is like their breathing and everything else like that. So I just know the integrity I feel is so high in anyone who’s in [the] archival [profession], because

[they]'re not in it for the money, and [they]'re not in it for anything other than to really try to share all that [they] can gather and put in [the digital archives] and such like that. (FG3-S6)

FG3-S6's perceived integrity of WADA staff and the archival profession in general removed all doubt for her that any document she finds would be altered and therefore unstable due to tampering.

4.2.8 Legibility

The eighth theme for which participants expressed support was the theme of legibility or readability. Table 4.8 presents results of micro-interlocutor analysis regarding the legibility theme. As shown in Table 4.8, during the first focus group, one participant provided significant statements in support of the legibility theme, with the remaining six participants not providing any response. During the second focus group, one participant provided significant statements in support of the legibility theme, one participant provided expressions in support of the legibility theme, one participant provided significant statements dissenting against the legibility theme, with the remaining four participants not providing any response. During the third focus group, one participant provided significant statements in support of the legibility theme; three participants provided expressions in support of the legibility theme, with the remaining four participants not providing any response. Across the three focus groups, 7 of 22 participants either made strong statements in support of this theme or expressed agreement.

Table 4.8. Micro-interlocutor Analysis Results for Legibility Theme.

	FG 1	FG 2	FG3
SA	1	1	1
A	0	1	3
SD	0	1	0
D	0	0	0
NR	6	4	4

Table 4.8 Micro-interlocutor Analysis Results for Legibility Theme

FG = Focus Group.

SA = Provided significant statement or example suggesting agreement.

A = Indicated agreement (i.e., verbal or nonverbal).

SD = Provided significant statement or example suggesting dissent.

D = Indicated dissent (i.e., verbal or nonverbal).

NR = Did not indicate agreement or dissent (i.e., non-response).

Participants provided some statements in support of the legibility theme. For example, in multiple focus groups, participants articulated this particular conceptualization in the form of a question. For example, during the first focus group, FG1-S4 asked, “trustworthiness, can I read it?” and during the third focus group, FG3-S2 asked, “by the word trustworthy, do you mean that they’re readable? Or, you know, that they’re visually trustworthy?” These participants’ questions about what the moderator meant by the term trustworthiness suggest that they conceptualized trustworthiness in terms of a document’s legibility or readability.

Evidence that other participants besides FG1-S4 and FG3-S2 also conceptualized trustworthiness in terms of legibility or readability came from participants’ responses to one of the moderator’s prompts, which asked participants to describe a document from WADA that they thought was trustworthy. During the first focus group, FG1-S4 replied, “readable, as in filmed properly” (by filmed

properly, FG1-S4 was referring to digitization). In essence, FG1-S4 conceptualized trustworthiness as a well-digitized, and therefore readable, document.

The following passage illustrates this theme along with the consensus and dissent it caused in the context of the second focus group:

FG2-S1—I was going to say, I think an important thing is legibility.

Moderator—Okay. For the document as a document?

FG2-S1—Yes, because if it's, you know, a real fuzzy reproduction-

FG2-S3—(Overlapping) oh yeah. ((nods head in agreement)).

FG2-S1—Or the original was, what, the ink bled and things like that, and then the reproduction isn't really clear, what you read on, that could be-

FG2-S5—No, now you're talking about the information though.

It's still a document even though it got run over by a train.

In the preceding passage, FG2-S1 articulated her conceptualization of trustworthiness in terms of a document's legibility. The moderator asked for clarification as to whether FG2-S1's conceptualization was specific to the document or the document content. FG2-S1 confirmed that her conceptualization of trustworthiness referred specifically to document trustworthiness, and then gave an example of reduced legibility as a result of poor digitization, to which FG2-S3 responded by nodding his head in agreement (kinesic). FG2-S5 dissented with this

conceptualization of trustworthiness. According to FG2-S5, whether or not a document is more or less legible does not make it more or less trustworthy.

4.2.9 Coverage

The ninth theme for which participants expressed support was the theme of coverage. Kelton et al. (2008) define coverage as “the completeness of the information” (p. 370). Table 4.9 presents results of micro-interlocutor analysis regarding the coverage theme. As shown in Table 4.9, during the second focus group, two participants provided significant statements in support of the coverage theme; two participants provided expressions in support of the coverage theme, with the remaining three participants not providing any response. During the first and third focus groups, I did not identify the coverage theme. Across the three focus groups, 4 of 22 participants either made strong statements in support of this theme or expressed agreement.

	FG 1	FG 2	FG3
SA	0	2	0
A	0	2	0
SD	0	0	0
D	0	0	0
NR	0	3	0

Table 4.9 Micro-interlocutor Analysis Results for Coverage Theme.

FG = Focus Group.

SA = Provided significant statement or example suggesting agreement.

A = Indicated agreement (i.e., verbal or nonverbal).

SD = Provided significant statement or example suggesting dissent.

D = Indicated dissent (i.e., verbal or nonverbal).

NR = Did not indicate agreement or dissent (i.e., non-response).

Participants provided some statements in support of the coverage theme. For example, the following interaction among the focus group participants illustrates this theme:

Moderator—How would you describe a document you found using WADA that you think is trustworthy? Like, what adjectives would you use?

FG2-S6—Complete.

FG2-S2—((Nods head in agreement)) Yeah, yeah. Complete.

Moderator—So, if some of that information was not there, or if it was ((points at FG2-S6 and then turns back to FG2-S2)) incomplete-

FG2-S2—((Nods head in agreement)) Correct.

Moderator—Then you would think that that document is not trustworthy?

FG2-S2—Not perhaps as trustworthy were it complete.

In the preceding passage, FG2-S6 and FG2-S2 both state that they would use the adjective “complete” to describe a document that they think is trustworthy, suggesting that they conceptualize document trustworthiness in terms of completeness. As a result of the moderator’s probing question, it becomes clear that participants’ document trustworthiness perception in terms of completeness is not binary, but continuous. For example, it was FG2-S2 who remarked that the more complete a document is, the more trustworthy she perceives it. Evidence that other participants shared a similar conceptualization came from FG2-S3 who nodded his head in agreement (kinesic) with FG2-S2.

4.2.10 Objectivity

The theme for which participants expressed the least support was the theme of objectivity. Kelton et al. (2008) define objectivity as “balance of content” (p. 370). Table 4.10 presents results of micro-interlocutor analysis regarding the objectivity theme. As shown in Table 4.10, findings provide minimal support for conceptualizing trustworthiness in terms of objectivity. During the third focus group, only one participant provided a statement in support of this theme, two participants expressed agreement using nonverbal communication, and the remaining five participants did not provide a response. During the first and second focus groups, I did not identify the objectivity theme. Across the three focus groups, 3 of 22 participants either made strong statements in support of this theme or expressed agreement.

	FG 1	FG 2	FG3
SA	0	0	1
A	0	0	2
SD	0	0	0
D	0	0	0
NR	0	0	5

Table 4.10 Micro-interlocutor Analysis Results for Objectivity Theme

FG = Focus Group.

SA = Provided significant statement or example suggesting agreement.

A = Indicated agreement (i.e., verbal or nonverbal).

SD = Provided significant statement or example suggesting dissent.

D = Indicated dissent (i.e., verbal or nonverbal).

NR = Did not indicate agreement or dissent (i.e., non-response).

One participant provided statements in support of the coverage theme. FG3-S3 mentioned the importance of questioning the bias that might be inherent in WADA document content, just as she suggested one must do for news information encountered on a daily basis:

It's like looking at the news today . . . on the web. . . You have to question the bias or how the person, that person, got the information. I think that the document, like I keep saying, I trust that the document on the archive is as presented and is the true document. The facts in the document, I may not always trust.

FG3-S3 suggested that one must consider the bias in WADA document content even if she or he considers the document trustworthy in terms of it being the document that it claims to be. Thus, FG3-S3 drew a distinction between document trustworthiness and document content trustworthiness such that a document can be perceived as trustworthy while its content can be perceived as untrustworthy in terms of bias. Evidence that other participants shared a similar conceptualization came from FG3-S1 and FG3-S2, who both nodded their heads in agreement (kinesic) with FG3-S3.

4.3 Conclusion

Results of this study provide valuable insight into how members of a specific designated community conceptualize document trustworthiness. These findings indicate that a broad range of concepts and issues pertain to the participants'

concept of document trustworthiness. Specifically, the findings suggest that, to varying degrees, genealogists, WADA's largest designated community, conceptualize trustworthiness in terms of a document's perceived authenticity, accuracy, first-hand or primary nature, believability, validity, proper form, stability, legibility, coverage, and objectivity. At one level, these findings underscore the complexity of the concept of document trustworthiness.

At another level, document trustworthiness may not be as complex as all ten of the themes I identified in this study might suggest. Participants did not express equal support for all themes. If the number of statements and expressions of support for a theme serve as indications of a theme's importance, then some conceptualizations of document trustworthiness are clearly more important than others. Considering the results of micro-interlocutor analysis, the themes of authenticity and accuracy are most important and most salient with regard to user document trustworthiness perception, and the themes of coverage and objectivity are least important.

Regardless of the frequency with which participants discussed the themes I identified, the findings are important for scale development. During the early stages of a scale development project, the researcher develops an exhaustive list of all possible items that could potentially be useful for measuring the construct of interest (DeVellis, 2012). Even inclusion of conceptually redundant items is encouraged during the early stages of scale development, as some items usually perform better than others (DeVellis, 2012). DeVellis (2012) points out that focus groups can be useful for developing survey items for a scale development project.

Since no instruments exist for measurement of digitized archival document trustworthiness, the focus groups provided an opportunity for item development. I derived a total of fifty-one items for development of a scale to measure document trustworthiness from participants' responses during the focus groups. Table 4.11 lists those items.

Items
1. The document is altered with the intention to deceive.
2. Names are misspelled.
3. The document lists an incorrect marital status for a person.
4. The document lists an incorrect age for a person.
5. The document is hearsay.
6. The document is readable.
7. The document is a primary source.
8. The document lists fake names.
9. The document is from the time period it claims to be.
10. The document is a secondary source.
11. The document matches research I have found using other sources.
12. The document is censored or blacked out.
13. The document is properly digitized.
14. The document contains a seal.
15. The handwriting looks to be of the correct time period.
16. The document represents an account of the event afterwards.
17. The document lists a maiden name instead of a married name.
18. Parts of the document are crossed out.
19. The document is the actual scanned image.
20. I have seen the original physical document that was used to create the digitized document.
21. The document is real.
22. The document was written at the time of the event.
23. The document has documentation of where it came from.
24. The document looks like it comes from where I think it comes from.
25. The document has all the appropriate fields, even though all the fields are not completely filled out.
26. I know who provided the information.
27. The document looks like what an official document should look like.
28. The document lists the wrong place of residence for a person.
29. The digitized document is an actual picture of the original physical document.
30. The document is fake.

31. The digitized document looks like what the original document should look like.
32. The document is accurate, but the actual information is erroneous.
33. The document is the closest to the original "action" that happened.
34. I found information elsewhere that conflicts with the document.
35. The document is a fuzzy reproduction.
36. I can see that some of the spelling has been changed.
37. The document has the proper form.
38. The document matches information I know about.
39. The document was altered by the witnesses.
40. The document lists the wrong place of birth for a person.
41. The document contains false information.
42. The document was written by the minister who filled out the form for the marriage.
43. The document lists an adopted parent as a birth parent.
44. The document is second-hand.
45. The person the document is about was alive during the time the document was created.
46. The document is mistakenly identified.
47. I can see that at least one of the dates have been changed.
48. The document is handwritten.
49. The document is notarized.
50. A person is listed as being in two places at the same time.
51. The document is typed.

Table 4.11 Items Derived from the Focus Groups Study Findings

Overall, the findings from this study shed light on what document trustworthiness is from the perspective of a small yet representative and appropriate sample of members of a specific designated community. At the conclusion of the focus group study, it was still too early to know which conceptualizations of document trustworthiness I identified were most important to the designated community under investigation, as the focus groups involved a limited number of participants. Consequently, the findings of the focus groups were utilized in the primary project study, which deployed selected responses from participants during the focus groups as survey items to a much larger sample of

designated community members for their evaluation. I discuss those results in the following chapter.

4.4 References

- Boyatzis, R. E. (1998). Transforming qualitative information: Thematic analysis and code development. Thousand Oaks, CA: Sage Publications.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101.
- Case, D. O. (2008). Collection of family health histories: The link between genealogy and public health. *Journal of the American Society for Information Science and Technology*, 59(14), 2312-2319.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37-46.
- Crowe, E. P. (2011). *Genealogy online* (9th ed.). New York: McGraw Hill.
- Darby, P., & Clough, P. (2013). Investigating the information-seeking behaviour of genealogists and family historians. *Journal of Information Science*, 39(1), 73-84.
- DeVellis, R. F. (2012). *Scale development: Theory and applications*. Thousand Oaks, Calif.: SAGE.
- Duff, W. M., & Johnson, C. A. (2003). Where is the list with all the names? information-seeking behavior of genealogists. *American Archivist*, 66(1), 79-95.
- Fogg, B. J. (2003). *Persuasive technology: Using computers to change what we think and do*. Amsterdam ;Boston: Morgan Kaufmann Publishers.

- Kelton, K., Fleischmann, K. R., & Wallace, W. A. (2008). Trust in digital information. *Journal of the American Society for Information Science and Technology*, 59(3), 363-374.
- Krueger, R. A. (1998). *Moderating focus groups*. Thousand Oaks ; London: Sage.
- Krueger, R. A., & Casey, M. A. (2009). *Focus groups : A practical guide for applied research* (4th ed.). Los Angeles: SAGE.
- Liamputtong, P. (2011). *Focus group methodology : Principles and practices*. Los Angeles: SAGE.
- Lucas, S. A. (2008). *The information seeking processes of genealogists*. (Unpublished doctoral dissertation, Emporia State University).
- Maxwell, J. (1992). Understanding and validity in qualitative research. *Harvard Educational Review*, 62(3), 279-301.
- Morgan, D. L. (1997). *Focus groups as qualitative research*. Thousand Oaks, Calif.; London: SAGE.
- Multon, K. (2010). Interrater reliability. In N. Salkind (Ed.), *Encyclopedia of research design*. (pp. 627-629). Thousand Oaks, CA: SAGE Publications, Inc. doi: <http://dx.doi.org/10.4135/9781412961288.n194>
- Onwuegbuzie, A. J., Dickinson, W. B., Leech, N. L., & Zoran, A. G. (2009). A qualitative framework for collecting and analyzing data in focus group research. *International Journal of Qualitative Methods*, 8(3), 1-21.
- Stewart, D. W., & Shamdasani, P. N. (1990). *Focus groups: Theory and practice*. Newbury Park, Calif.: Sage.

- Stewart, D. W., Shamdasani, P. N., & Rook, D. W. (2007). Focus groups : Theory and practice. Thousand Oaks ; London ; New Delhi: Sage.
- Tseng, S., & Fogg, B. J. (1999). Credibility and computing technology. *Commun.ACM*, 42(5), 39-44.
- Veale, K. H. (2005). A doctoral study of the use of the internet for genealogy. *Historia Actual Online*, 7, 7-14.
- Yakel, E. (2004). Seeking information, seeking connections, seeking meaning: Genealogists and family historians. *Information Research*, 10(1)
- Yakel, E., & Torres, D. A. (2007). Genealogists as a "community of records". *American Archivist*, 70(1), 93-113.

CHAPTER 5

RESULTS OF STEPS 3 AND 4 OF SCALE DEVELOPMENT: MEASURING TRUSTWORTHINESS

This chapter presents results of the culmination of Steps 3 and 4 of scale development: Designing the Scale (Step 3) and Full Administration and Item Analysis (Step 4) (see Figure 5.1). Step 3 involved transforming the item pool resulting from Step 2 of scale development into a web-administered survey (see Appendix G) such that actual users of digitized archival documents could evaluate those items with their experiences in mind. Step 3 also involved pretesting that instrument for purposes of refinement and clarification. During Step 4, I administered the instrument resulting from Step 3 to a large sample of participants for their evaluation of the trustworthiness items. The results of analysis of the participants' data are the focus of this chapter. In this respect, this chapter addresses both Step 3 and Step 4 of scale development.

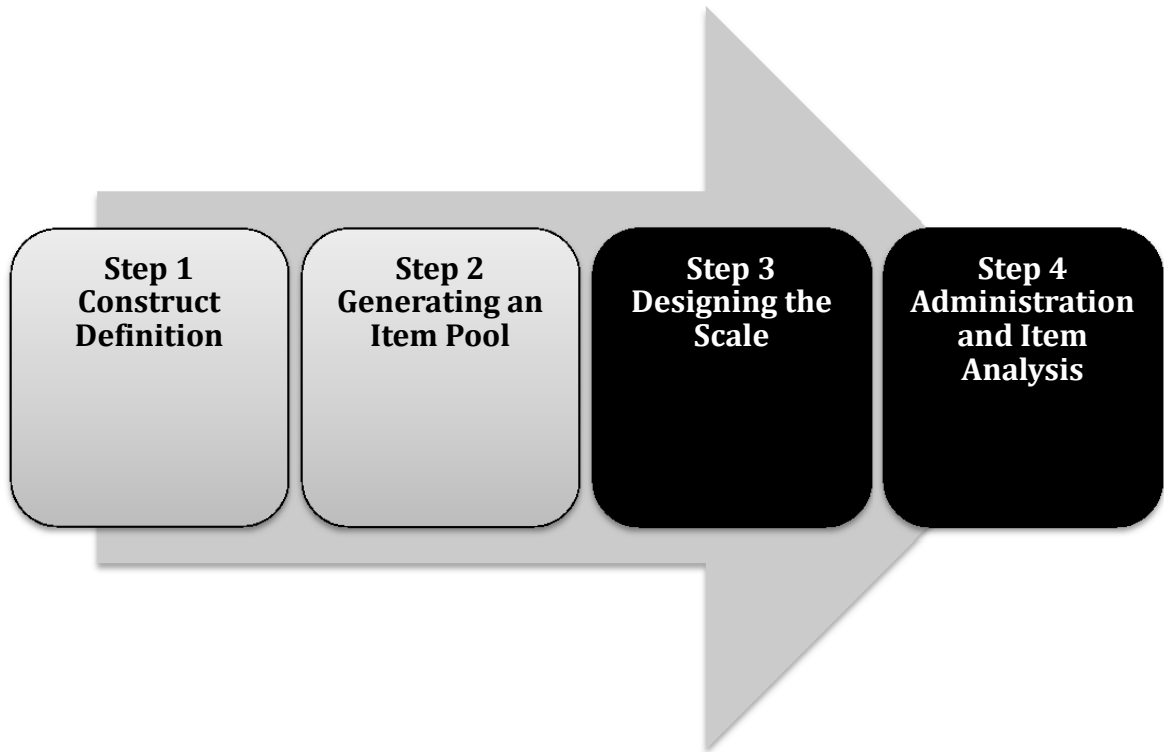


Figure 5.1 Contribution of the Chapter 5 Study Findings to Scale Development.

As denoted in black, the Chapter 5 study findings culminate Steps 3 and 4 of scale development. The study uses the instrument developed as a result of activities during Step 3 of scale development to gather data from a large sample of participants and analyze the responses to trustworthiness items during Step 4. These findings speak specifically to the measurability of the construct of trustworthiness perception specifically for archival documents delivered digitally within a digital repository context.

The item analysis aspect of Step 4 requires a series of tests and analyses aimed at providing valuable insight into whether items that are hypothesized to measure a construct are actually good indicators of that construct (DeVellis, 2012). In the social sciences, many constructs are latent variables; they are not directly observable. Regardless, researchers can introduce a construct into a mathematical model for measurement and treat that construct just like any other variable (Bartholomew, Steele, Moustaki, & Galbraith, 2008). Specifically, this chapter

presents results of an attempt to introduce the construct of trustworthiness into a mathematical model for purposes of empirical measurement.

To address the second research question, “To what extent are designated community members’ perceptions of document trustworthiness measurable?”, I conducted a quantitative survey research study. These data represent participants’ ratings of a broad range of items that I hypothesized would measure the construct of document trustworthiness.

Taken together, the results of item analysis suggest the development of a viable and statistically reliable scale—the Digitized Archival Document Trustworthiness Scale (DADTS). DADTS has propitious psychometric properties and includes twelve items that I derived from one very powerful factor; it offers a practical means of measuring user document trustworthiness perception.

The structure of the chapter is as follows. First, I describe participant recruitment and present survey data regarding the study participants whose responses were used to generate data for item analysis. Second, I discuss results of item analysis. Specifically, I present results of five types of item analysis that DeVellis (2012) recommends as part of scale development, including examination of item variances, item-total correlations, item means, factor analysis, and computation of reliability coefficients. Afterwards, I present findings from a validation study for DADTS.

5.1 Study Participants

According to DeVellis (2012), scale development depends upon administration of items to “an appropriately large and representative sample” (p. 104). Since the main goal of the study was to develop a measure for the trustworthiness perception of digital archival documents, an appropriate sample had to include participants with experience using digital archival documents. Prior contact with the Deputy State Archivist of the Washington State Archives confirmed that WADA’s largest designated community of users included genealogists (T.S. Badger, personal communication, March 8, 2013). Thus, an appropriate and representative sample for this study required participants with demographics similar to the larger population of genealogists as well as experience in using digital archival documents.

While there is no sampling frame available to assess genealogists as a population, there are multiple studies of genealogists in archival science and information science. Most of what is known about genealogists in archival science and information science is based upon samples of predominantly older females (e.g., Case, 2008; Duff & Johnson, 2003; Lucas, 2008; Yakel, 2004; Yakel & Torres, 2007). I derived a profile of demographic characteristics from studies of genealogists in archival science and information science and used that profile as a proxy for the demographics of the larger population. In this study, participants answered questions about their age and gender. Afterwards, I compared the participants’ responses to findings of previous studies in archival science and information science involving genealogists as a measure of the representativeness of my sample.

In addition to demographic characteristics, the study also required participants with experience using digital archival documents. Verifying the behavior of the study participants beforehand, e.g., observing them as they utilized digital archival documents over time, would have ensured that they had experience using digital archival documents. However, this approach was not practically feasible. Alternatively, asking a set of questions pertaining to participants' WADA document usage was possible. Although this study is limited by no attempts to verify the self-reports, they at least provide some insight into the extent to which the study participants thought they used digital archival documents.

In addition to answering questions pertaining to their age and gender, the study participants also answered demographic questions pertaining to their use of WADA documents. I included these questions in order to understand participants' demographic characteristics and participants' appropriateness for the study. After discussing participant recruitment, sampling strategies, and the sample size, this section examines participants' responses to the demographic questions and WADA document use questions, highlighting the trends that I identified as a result of the data analysis.

5.1.1 Participant Recruitment and Sampling Strategies

I recruited participants for this study via the intercept survey method. According to Couper (2000), "intercept surveys target visitors to a Web site [sic]" and "generally [use] systematic sampling to invite every *n*th visitor to a site to

participate in a survey” (p. 485). In this study, I targeted visitors to WADA’s homepage who would self-report that they were genealogists.

Between December 21, 2013 and February 2, 2014, there were 712,598 page views to WADA’s homepage (T.S. Badger, personal communication, December 9, 2014). WADA staff set their web site to provide a pop-up invitation to thirteen percent of the visitors; 54,815 of the 712,598 page viewers received an invitation to participate in WADA’s customer satisfaction survey (T.S. Badger, personal communication, December 9, 2014). To ensure that WADA users with pop-up blockers would also receive an invitation to participate, WADA staff had the invitation to participate appear as a small window overlaying WADA’s homepage rather than as a pop up requiring a second window. This was important because using a separate window to administer the invitation would have excluded participants with pop-up blockers (Fricker, Jr., 2008).

Of these 54,815 invitations, 969 visitors took WADA’s customer satisfaction survey (T.S. Badger, personal communication, December 9, 2014). The last question of WADA’s customer survey included an invitation and a link to participate in my survey. Of the 969 visitors who took WADA’s survey, 127 clicked the link to my survey and indicated either that “yes” they wanted to take my survey or “no” they did not want to take my survey. Specifically, 116 participants answered, “yes” that they wanted to take my survey. Of those 116, 51 self-reported that they were genealogists.

To increase participation, I modified the sampling plan in two important ways. First, I removed WADA’s customer satisfaction survey from the sampling plan.

The rationale for this choice was that participants might be willing to take one survey, but they were far less willing to take two surveys. Second, I increased the percentage of pop-up invitations to participate in my survey from thirteen percent to thirty percent.

From February 10, 2014 to March 31, 2014, there were 618,179 page views to WADA's home page (T.S. Badger, personal communication, December 9, 2014). WADA staff set their web site to provide a pop-up invitation to thirty percent of the visitors; 20,606 of the 618,179 page viewers received an invitation to participate in my study that included a link to my survey (T.S. Badger, personal communication, December 9, 2014). Of these 20,606 invitations, 610 clicked the link to my survey and indicated either that "yes" they wanted to take my survey or "no" they did not want to take my survey. Specifically, 476 participants answered, "yes" that they wanted to take my survey. Of those 476, 182 self-reported that they were genealogists.

5.1.2 Sample Size

From December 21, 2013 through March 31, 2014, 233 subjects self-reported that they were genealogists and participated in the study. After listwise case deletion for the demographics questions, WADA document use questions, and the trustworthiness items, 180 subjects remained. As a result of following procedures for outlier detection (see Appendix H), I removed another participant's data from further analysis, bringing the sample to 179 participants. I employed listwise case deletion; thus, there were no missing data in the dataset.

Nunnally and Bernstein (1994) recommend participation of four to ten subjects per item in a scale development project. Since I initially retained twenty-six items for analysis in this study, a sample of 179 participants falls within Nunnally and Bernstein’s guidelines for sample sizes.

5.1.3 Demographic Information

Questions pertaining to participants’ demographic information included questions about participants’ age and gender.

5.1.3.1 Participants by Age

Table 5.1 displays the study participants by age. As shown in Table 5.1, all but seven participants (96%) reported that they were 40 years of age or older. The rest were either between 30 and 39 (n=3) or 20 and 29 (n=4).

Age Ranges	Frequency	Percent
20 to 29	4	2%
30 to 39	3	1%
40 to 49	20	11%
50 to 59	48	27%
60 to 69	66	38%
70 or older	38	21%
Total	179	100%

Table 5.1 Participants by Age.

Overall, these results provide empirical support for the representativeness of the sample. Most of what is known about genealogists in archival science and

information science is based upon samples of older participants (e.g., Case, 2008; Duff & Johnson, 2003; Lucas, 2008; Yakel, 2004; Yakel & Torres, 2007). While demographic characteristics of the population of genealogists are unknown, what can be said with certainty is that results of this study's participants' self-reports are consistent with studies in archival science and information science regarding the average age of genealogists who agree to participate in studies.

5.1.3.2 Participants by Gender

Table 5.2 displays the study participants by gender. As shown in Table 5.2, almost three-fourths (74%) of the participants reported that they were female. The remaining 26% reported that they were male.

Sex	Frequency	Percent
Female	132	74%
Male	47	26%
Total	179	100%

Table 5.2 Participants by Gender.

Overall, these results provide additional empirical support for the representativeness of the sample. Most of what is known about genealogists in archival science and information science is based upon samples of predominantly female participants (e.g., Duff & Johnson, 2003; Lucas, 2008; Yakel, 2004; Yakel & Torres, 2007). In his review of recent studies on genealogists in information science, Case (2008) concluded that genealogy is a hobby that is perhaps predominantly

female. Thus, while demographic characteristics of the population of genealogists are unknown, what can be said with certainty is that results of this study's participants' self-reports are consistent with studies in archival science and information science regarding the breakdown of participants by gender.

5.1.4 Participants' WADA Document Usage

Participants answered questions related to their: primary reason for using WADA documents, frequency of using WADA documents, frequency of use by document type, and time spent using documents on a typical visit to WADA. Responses to these questions create a profile of the sample with regard to their actual experience using WADA documents and provide a basis for the appropriateness of the sample for the study, given its aims.

5.1.4.1 Primary Reason for Using WADA Documents

All study participants reported that their primary reason for using WADA documents was to conduct genealogical research. Thus, all of the participants in this study are considered self-reported genealogists.

5.1.4.2 Frequency of Using WADA Documents

Table 5.3 presents study participants' frequency of WADA document usage. As shown in Table 5.3, nearly three-fourths of the survey sample reported that they used WADA documents daily (4%), weekly (28%), or monthly (42%). Nearly a quarter (24%) of the participants indicated that they used WADA documents a few times a year. The remaining four participants indicated that they had not used WADA documents within the last year. Most of the sample of participants uses WADA documents regularly.

WADA Document Use	Frequency	Percent
Not at all (within the last year)	4	2%
A few times a year	42	24%
Monthly	75	42%
Weekly	50	28%
Daily	8	4%
Total	179	100%

Table 5.3 Participants' Frequency of Using WADA Documents

Overall, these results provide empirical support for the appropriateness of the sample for the study. The survey instructions asked participants to draw upon their recent experiences using WADA documents in their evaluation of the trustworthiness items. These findings demonstrate that a majority of the study participants actually had recent experiences with using WADA documents that they could draw upon in their evaluation of the trustworthiness items that the participants evaluated during the study.

At one level, I cannot rightfully compare these data with results from other studies in archival science and information science. Prior studies have not typically asked participants specifically about their frequency of digital archival document use. Instead, researchers have typically asked participants questions concerning how many years they have conducted genealogical research in general (e.g., Darby & Clough, 2013; Lucas, 2008; Yakel, 2004; Yakel & Torres, 2007), their use of online resources (Darby & Clough, 2013; Lucas 2008), and their use of county archives or record office websites (Darby & Clough, 2013). At another level, I can compare the data regarding participants' frequency of WADA document use to results of other studies in archival science and information science. It is plausible that in these previous studies at least some of the participants' genealogical research, use of online resources, and use of county archives websites involved use of digital archival documents. Considering this plausibility, the data collected in this study are congruous with the findings of previous studies regarding frequency of use of digital archival documents.

5.1.4.3 Frequency of Use by Document Type

Table 5.4 presents participants' frequency of use by document type. As shown in Table 5.4, the study participants reported using digitized marriage records most frequently (55%) followed by death records (31%), census records (3%), birth records (2%), and land records (1%). Fourteen participants (8%) used the "other" category to indicate that they use birth, death, marriage, census, and land records. These findings demonstrate that, although a majority of the sample reported that

they used marriage records most frequently, the sample had experience using a broad range of digital archival documents.

Document Types	Frequency	Percent
Birth records	5	2%
Death records	55	31%
Marriage records	99	55%
Census records	5	3%
Land records	1	1%
Other	14	8%
Total	179	100%

Table 5.4 Participants' Frequency of Use by Document Type

Overall, these results provide additional empirical support for the appropriateness of the sample for the study. To a certain extent, these findings cross-validate other findings pertaining to participants' frequency of use regarding WADA documents. Specifically, these findings provide insight into which type of documents participants reported using. Had the participants indicated that they used a document type not preserved by WADA, this would have suggested error on the part of the participants or underscored their inappropriateness for the study.

While findings from prior studies suggest that genealogists utilize birth, death, marriage, census, and land records, these studies focus less on whether or not those documents are digital (e.g., Case, 2008; Darby & Clough, 2013; Duff & Johnson, 2003; Lucas, 2008; Yakel, 2004; Yakel & Torres, 2007). This makes the findings from

this study difficult to compare with results from previous studies. While participants did not specific questions about participants’ use of digital archival documents in these studies, it is possible that the participants used digitized birth, death, marriage, census, and land records. If so, the findings from this study align with previous studies regarding the representativeness of the sample regarding digital archival document use.

5.1.4.4 Time Spent Using Documents on a Typical Visit to WADA

Table 5.5 presents participants’ time spent using WADA documents on a typical visit. As shown in Table 5.5, nearly half (47%) of the participants indicated that they spent either over 30 minutes to an hour (29%) or between 1 and 2 hours (18%) using WADA documents on a typical visit. Ten percent reported that they spend over 2 hours using WADA documents per visit. The remaining 43% indicated that they spend between 0 to 30 minutes using WADA documents on a typical visit. Overall, these results demonstrate that the study participants spend a substantial amount of time using WADA documents per visit.

Time Spent	Frequency	Percent
0-30 minutes	76	43%
31-59 minutes	52	29%
1-2 hours	33	18%
More than 2 hours	18	10%
Total	179	100%

Table 5.5 Participants’ Time Spent Using Documents on a Typical Visit to WADA.

Overall, these results provide additional empirical support for the appropriateness of the sample for the study. Specifically, these findings provide some sort of perspective on how long the participants think they spend interacting with WADA documents. For example, if participants indicated that they used WADA documents frequently, e.g., weekly, but only look at WADA documents for seconds at a time, they may not have enough experience in terms of time spent with documents to inform their perspective of WADA documents. Participants without experience using WADA documents in terms of quantity and quality would have been inappropriate for this study. Answers to this question suggest that the majority of participants were appropriate for this study because they had enough experience with WADA documents per visit to draw upon in their evaluation of the trustworthiness items in the survey.

Previous studies on genealogists in archival science and information science tend to focus more on participants' years of conducting genealogy rather than addressing how long participants use digital archival documents (e.g., Case, 2008; Darby & Clough, 2013; Duff & Johnson, 2003; Lucas, 2008; Yakel, 2004; Yakel & Torres, 2007). This makes the findings from this study difficult to compare with results from previous studies. Regardless, it is plausible that participants in previous studies spent at least some of their time using digital archival documents.

5.1.5 Study Participants Section Summary

Considering the data from Tables 5.1 through 5.5 creates a profile of the study participants, which suggests that they comprise an appropriate and representative sample. The study participants:

- Are mostly female and 40 years of age or older,
- Use WADA documents frequently, and spend ample time using WADA documents per visit (over 50% of the study participants reported spending over 30 minutes using WADA documents per visit), and
- Utilize specific types of documents, including digitized marriage, death, birth, census, and land records.

This composite profile suggests that it was possible for the study participants to evaluate the trustworthiness items in this study based on the quantity and quality of their reported experiences with WADA documents. This was critical for this study, which focuses on trustworthiness perception regarding digital archival documents.

5.2 Initial Examination of Items' Performance

Evaluating items is essential to ensuring that a scale actually measures the construct one seeks to measure (DeVellis, 2012). DeVellis (2012) emphasizes that “item evaluation is second perhaps only to item development in its importance” (p. 104).

In a scale development project, a high correlation with the true score of the latent variable is an ideal quality (DeVellis, 2012). In reality, researchers cannot

assess the true score on a latent variable because latent variables are not directly observable (DeVellis, 2012). Researchers can assume relationships between latent or unobservable variables (i.e., factors) and manifest or observable variables (i.e., items) based upon relationships that researchers can assess among items (Bartholomew et al., 2008). However, correlations among items cannot exist if items do not exhibit certain properties. Consequently, DeVellis (2012) recommends assessing items' performance after collecting data on those items. Applied to this study, I can make assumptions about trustworthiness—the latent or unobservable variable that is the focus of this study—based upon correlations among the items that I hypothesize to relate to the construct of trustworthiness—such as the items that I administered to the study participants during Step 4. As a prerequisite, I examined the trustworthiness items' performance. DeVellis (2012) recommends three tests for assessing items' performance, including examination of items' variances, item-total correlations, and item means.

5.2.1 Items' Variances

An important attribute of an item is relatively high variance (DeVellis, 2012). According to DeVellis (2012), if all individuals provide the same rating of an item, then that item is not capable of discriminating against people who possess different levels of the construct. DeVellis argues that the development sample of participants should be diverse with respect to the attribute of interest, and thus people's responses to items should also be diverse.

In this study, participants evaluated document trustworthiness items via a web-administered survey (see Appendix G). Each item represented a circumstance people might encounter regarding a digitized archival document that they have recently found and subsequently used. On a 7-point scale, participants rated how trustworthy they would perceive a digitized archival document, given the circumstance described in each item. The scale points ranged from -3, “very untrustworthy,” to +3, “very trustworthy.” The survey also included the option 99, “not applicable,” for participants to utilize if they felt the circumstance described in an item was not relevant to their trustworthiness perceptions. The participants evaluated seventy-four items. Of those items, I excluded forty-eight from further analysis because ten or more participants indicated that the circumstances described by those items were irrelevant (see Appendix I for a list of these items). Consequently, I retained the remaining twenty-six items for further analysis. I discuss the results of analysis for these items in the rest of this section.

Applying DeVellis’s position on item variances to this study, an item that participants provided a range of trustworthiness ratings for is more useful item than one that everyone provides the same rating for. The idea is that if everyone provides the same trustworthiness rating for an item, that item is not capable of discriminating against those who possess different levels of document trustworthiness perception.

To assess item variances, I inspected the range of responses for each item. In particular, I examined items’ minimums and maximums. Table 5.6 presents participants’ minimum and maximum ratings for each item. To highlight item

variances more clearly, I recoded participants' responses from -3 to +3 to 1 to 7. As shown in Table 5.6, answers ranged from 1 to 7 for five items, answers ranged from 2 to 7 for six items, answers ranged from 3 to 7 for three items, and answers ranged from 4 to 7 for the remaining twelve items.

Items	Minimums	Maximums
The information is correct.	1	7
The document appears complete.	1	7
The document resembles what an official document should look like.	1	7
The document includes verifiable data.	1	7
The document accurately reflects what happened.	1	7
The document is authoritative.	2	7
The document is from the time period it claims to be.	2	7
The document appears free from error.	2	7
The document is credible.	2	7
The document includes documentation of where it came from.	2	7
The document is official.	2	7
The document is factual.	3	7
The document is authentic.	3	7
The digitized document resembles what the original document should look like.	3	7
The document is readable.	4	7
The document matches research I have found using other sources.	4	7
The document is a primary source.	4	7
The document was written at the time of the event.	4	7
The document is what it claims to be.	4	7
The document is believable.	4	7
The document has all the appropriate fields, even though all the fields are not completely filled out.	4	7
The document was created using responsible and accepted practices.	4	7
The digitized document is an actual picture of the original physical document.	4	7
The document is legible.	4	7
The document matches other information I know about.	4	7
The document is the same every time I download or view it.	4	7

Table 5.6 Items' Recoded Minimums and Maximums

The discrepancy between the minimum and maximum response choices represent variation in responses among the study participants regarding the items. This variation is key because it suggests that the items are capable of discriminating

among individuals with different levels of trustworthiness perception. Thus, analysis of the items' minimum and maximum responses suggest that the items are good for measuring trustworthiness.

5.2.2 Item-Total Correlations

In scale development, a set of observable items is assumed to relate to the same underlying, unobservable construct (DeVellis, 2012). Specifically, the more strongly items correlate, the more strongly those items are assumed to reflect the same underlying construct (DeVellis, 2012). Thus, a set of highly intercorrelated items is highly desirable in scale development (DeVellis, 2012).

To obtain a set of highly intercorrelated items, DeVellis (2012) recommends that, "each individual item should correlate substantially with the collection of remaining items" (pp. 106-107). According to DeVellis (2012), this property of items is most appropriately assessed through examination of items' corrected item-total correlations. A corrected item-total correlation represents the correlation of the item being evaluated with all the scale items, excluding itself (DeVellis, 2012). Items with high corrected item-total correlations are desirable and should be retained for further analysis, and items with low item-total correlations should be discarded (DeVellis, 2012).

Table 5.7 presents the items' corrected item-total correlations. As shown in Table 5.7, the items' corrected item-total correlations ranged between .811 and .514, with an average corrected item-total correlation of .697. All of the item-total correlations fall above Nunnally and Bernstein's (1994) recommended cut off of .30.

Given such high corrected item-total correlations, I retained all items for further analysis during this stage of item evaluation.

Items	Corrected Item-Total Correlations
The document is from the time period it claims to be.	0.811
The document is credible.	0.790
The document has all the appropriate fields, even though all the fields are not completely filled out.	0.770
The document resembles what an official document should look like.	0.762
The digitized document resembles what the original document should look like.	0.753
The document is believable.	0.749
The document matches other information I know about.	0.741
The document is what it claims to be.	0.740
The document is authentic.	0.740
The document was created using responsible and accepted practices.	0.740
The document is authoritative.	0.734
The document is official.	0.733
The document appears complete.	0.731
The document includes documentation of where it came from.	0.730
The document is factual.	0.728
The document is the same every time I download or view it.	0.719
The document appears free from error.	0.694
The document was written at the time of the event.	0.691
The document includes verifiable data.	0.686
The document is legible.	0.641
The document is readable.	0.638
The information is correct.	0.589
The document accurately reflects what happened.	0.583
The document is a primary source.	0.578
The document matches research I have found using other sources.	0.549
The digitized document is an actual picture of the original physical document.	0.514

Table 5.7 Items' Corrected Item-Total Correlations

5.2.3 Items' Means and Standard Deviations

According to DeVellis (2012), another favorable attribute of an item is having “a mean close to the center of the range of possible scores” (p. 107). For example, if using a 7-point scale, a mean of 4 is ideal. According to DeVellis, a mean of 4 would

indicate that participants used a range of responses to arrive at that mean. In actuality, there are two ways to arrive at a mean which lies at the midpoint of possible responses to an item: 1) people use a range of responses (i.e., some low responses and some high responses), such that the average response is the midpoint of possible responses, or 2) everyone provides the same rating, given that rating is the midpoint of possible responses. DeVellis assumes positive implications from the former way of attaining a midpoint mean.

Table 5.8 presents the means for each item. As shown in Table 5.8, the means for the trustworthiness items ranged from 5.46 to 6.49, with an average of 5.99. For scale development, more important than the items' means was how the means were generated, which data pertaining to items' minimums and maximums (see Section 5.2.1) as well as items' standard deviations do more to explain.

Items	Means
The digitized document is an actual picture of the original physical document.	6.49
The document matches research I have found using other sources.	6.43
The information is correct.	6.22
The document is authentic.	6.22
The document is a primary source.	6.19
The document matches other information I know about.	6.19
The document includes verifiable data.	6.18
The document is what it claims to be.	6.15
The document accurately reflects what happened.	6.15
The document was written at the time of the event.	6.13
The document includes documentation of where it came from.	6.12
The document is official.	6.10
The document was created using responsible and accepted practices.	6.08
The document is factual.	6.06
The document is credible.	6.03
The document is from the time period it claims to be.	5.97
The document is the same every time I download or view it.	5.93
The digitized document resembles what the original document should look like.	5.84

The document is believable.	5.81
The document appears free from error.	5.79
The document appears complete.	5.74
The document is authoritative.	5.71
The document has all the appropriate fields, even though all the fields are not completely filled out.	5.66
The document resembles what an official document should look like.	5.61
The document is legible.	5.47
The document is readable.	5.46

Table 5.8 Items' Means

Table 5.9 presents the standard deviations for each item. As shown in Table 5.9, the standard deviations for the trustworthiness items ranged from .757 to 1.264. These standard deviations demonstrate the variation in responses among participants for each of the items. Results confirm more variation in responses for some items than others. For example, 12 (46%) of the items had standard deviations greater than 1, indicating greater variation in responses among participants for those specific items. For the remaining 14 (54%) items, the standard deviations were less than 1, indicating less variation in responses among the study participants for those items. Despite more variation in responses for some items than others, overall, the standard deviations for the items provide additional support for the idea that the items are capable of discriminating among individuals with different levels of document trustworthiness perception.

Items	Standard Deviations
The document resembles what an official document should look like.	1.264
The document is authoritative.	1.215
The document is readable.	1.214
The document is legible.	1.191
The document is the same every time I download or view it.	1.159
The document appears complete.	1.158
The digitized document resembles what the original document should look like.	1.136
The document has all the appropriate fields, even though all the fields are not completely filled out.	1.127

The document appears free from error.	1.112
The document is believable.	1.059
The document is official.	1.023
The document is from the time period it claims to be.	1.016
The document is a primary source.	0.982
The information is correct.	0.973
The document accurately reflects what happened.	0.968
The document is factual.	0.967
The document is credible.	0.962
The document is what it claims to be.	0.955
The document was written at the time of the event.	0.948
The document is authentic.	0.945
The document was created using responsible and accepted practices.	0.935
The document includes verifiable data.	0.933
The document includes documentation of where it came from.	0.910
The digitized document is an actual picture of the original physical document.	0.830
The document matches other information I know about.	0.820
The document matches research I have found using other sources.	0.757

Table 5.9 Items' Standard Deviations

5.2.4 Initial Evaluation of Items' Performance Section Summary

Variation within and across items is a key factor in building a scale (DeVellis, 2012). Together, the statistical results regarding items' minimums, maximums, means, and standard deviations suggest sufficient variation in responses, which serves as evidence of the items' suitability as scale items. Specifically, variation in responses for these items suggests that they can measure different levels of trustworthiness perception that individuals may have regarding digitized archival documents. Had there been little or no variation, this would have suggested that, no matter how the participants considered trustworthiness, the items would not detect differences in terms of trustworthiness perception. Since variation exists, these items presumably are capable of measuring different levels of trustworthiness perception. Also, the extent to which a set of items intercorrelate is critical in developing a scale (DeVellis, 2012). Results of item-total correlations indicate that

the items participants evaluated in this study are highly intercorrelated, thus reflective of the construct of trustworthiness. Given the favorable results of all the tests of items' initial performance, I used these items in subsequent analysis, in particular Exploratory Common Factor Analysis.

5.3 Exploratory Common Factor Analysis Results

According to DeVellis (2012), factor analysis is “an essential tool in scale development” (p. 158). He suggests that factor analysis can reveal important properties of a scale. In particular, factor analysis can help “determine *empirically* how many constructs, or latent variables, or factors underlie a set of items” (DeVellis, 2012, p. 116; Jacoby, 1991).

I performed Common factor analysis in this study. Common factor analysis, also known as Principal Axis Factoring (PAF) (Kline, 2013), focuses only on the variance that is common or shared among items, excluding variance that is unique to any specific item (DeVellis, 2012). Thus, common factor analysis involves extraction of common factors, those factors that represent only the variance shared among the items the factors contribute to. According to DeVellis (2012), common factors are idealized—“they are estimates of what an error-free variable determining a set of items might look like” (p. 149).

When determining the underlying structure of a set of items, DeVellis (2012) recommends use of exploratory common factor analysis. It was clear, based upon the sources of the items in this study (e.g., the literature, trustworthiness subject matter experts, and actual users of digital archival documents), that the items in this

study were related to trustworthiness. However, it was not clear exactly *how* the items related. Consequently, I selected exploratory common factor analysis as the appropriate choice of factor analysis.

I employed an oblique, promax rotation in this study. I rotated Items during factor analysis to identify the most interpretable solution (DeVellis, 2012).

According to DeVellis (2012), factor rotation involves “identifying clusters of [items] that can be characterized predominantly in terms of a single latent variable” or factor (p. 133). An oblique rotation allows factors to correlate (DeVellis, 2012). Since I hypothesized that the items were related to one another, I assumed that the factors, if any, underlying the items would correlate. Therefore, I performed an oblique rotation on the items. I selected promax rotation as the type of oblique rotation to perform as it is most commonly used (Kline, 2013).

As I had no a priori hypotheses regarding how many factors to extract in a study of document trustworthiness, I performed several different factor extractions. In particular, I performed exploratory common factor analysis specifying extraction of two through eight factors. After specifying extraction of nine factors, SPSS 22.0, the statistical software program that I used to perform all factor analyses, terminated. Specification of nine factors caused one of the items to have a communality of greater than one, which SPSS 22.0 would not permit. I closely examined the two-, three-, four-, five-, six-, seven-, and eight-factor solutions to identify the factor solution with the simplest structure, i.e., the largest communalities with the fewest cross-loadings and the largest amount of items

loading strongly on each factor (Costello & Osborne, 2005). Consequently, the three-factor solution maintained the simplest structure.

The following presents the results of the exploratory common factor analysis (EFA) specifying a three-factor extraction. Appendix J lists the SPSS 22.0 syntax that submits these results. I used an oblique (promax) rotation in the analysis, as the factors are correlated. The data matrix in the analysis is the correlation matrix, which includes the correlations of all the items considered in the analysis (Hair, Black, Babin, & Anderson, 2010). The run of SPSS terminated normally with a converged and admissible solution after 500 iterations. After presenting results of tests regarding the appropriateness of the data for EFA, this section discusses results of tests concerning the appropriate number of factors to retain, the amount of variance the factors explain, and afterwards focuses on factor interpretation results.

5.3.1 Results of Tests for Appropriateness of the Data for EFA

I performed two tests to assess the appropriateness of the data for EFA: the Kaiser-Meyer-Olkin Measure of Sampling Adequacy (Kaiser, 1970; Kaiser & Rice, 1974) and Bartlett's Test of Sphericity (Bartlett, 1954). I performed both tests using the KMO function in SPSS 22.0.

5.3.1.1 Kaiser-Meyer-Olkin Measure of Sampling Adequacy Results

The Kaiser-Meyer-Olkin Measure of Sampling Adequacy (MSA) is necessary for evaluating the appropriateness of applying factor analysis (Hair et al., 2010). It is calculated for the correlation matrix, which contains the correlations of all items considered during factor analysis. According to Hair et al. (2010), values greater than or equal to .50 are acceptable. Computation of MSA during this study obtained a value of .946. According to Kaiser and Rice (1974), values above .90 for MSA are “marvelous” (p. 112). Thus, the data that I collected for this study were highly acceptable for factor analysis.

5.3.1.2 Bartlett’s Test of Sphericity Results

Bartlett’s Test of Sphericity is also necessary for assessing the appropriateness of data for factor analysis (Hair et al., 2010). It measures “the overall significance of all correlations within a correlation matrix” (Hair et al., 2010, p. 92). According to Hair et al. (2010), when values for Bartlett’s Test of Sphericity are significant (e.g., $p < .05$), it means that there is an acceptable number of significant correlations among the items. As factor analysis considers the correlations among items, it cannot be performed with too few of these correlations. In this study, results of Bartlett’s Test of Sphericity were significant (i.e., .000), suggesting a high enough proportion of significant correlations among the items for factor analysis.

5.3.1.3 Results of Tests for Appropriateness of the Data for EFA Summary

Taken together, values of the Kaiser-Meyer-Olkin Measure of Sampling Adequacy and Bartlett's Test of Sphericity obtained during this study ensured that the data had inherent sufficient correlations and justified the use of EFA.

5.3.2 Results of Tests for Selecting the Number of Factors to Retain

I performed three separate tests and compared their results to reach a conclusion regarding the appropriate number of factors to retain as a result of EFA: Kaiser's criterion (Kaiser, 1960), parallel analysis (Horn, 1965), and scree test examination (Cattell, 1966). They all involve examination of factors' eigenvalues. An eigenvalue is a measure of the proportion of variance in the items that a particular factor explains (DeVellis, 2012; Kline, 2013). Regarding Kaiser's criterion, I examined the eigenvalues of each factor to identify all factors with eigenvalues greater than 1 (Kaiser, 1960). Regarding parallel analysis, I compared the observed eigenvalues against those expected from random data (Horn, 1965). Specifically, I compared the eigenvalues to identify which observed eigenvalues exceeded the eigenvalues computed for the corresponding factors based on randomized data (O'Connor, 2000). I used eigenvalues that correspond to the 95th percentile of the distribution of random data eigenvalues as the baseline for comparison with the eigenvalues generated from the actual data (DeVellis, 2012; Kline, 2013). Regarding scree plot examination, I visually inspected a line graph of the eigenvalues for each of the total possible number of factors to locate the point where the drop in eigenvalues over successive factors levels out and from which the slope of the line is

basically horizontal, known as “the elbow” (Cattell, 1966; DeVellis, 2012; Kline, 2013).

5.3.2.1 Kaiser’s Criterion Results

Table 5.10 presents the initial eigenvalues that the EFA generated. As shown in Table 5.10, only the first three factors had eigenvalues greater than 1, satisfying Kaiser’s criterion. These three initial eigenvalues are in bold in Table 5.10 (Factor 1 (F1), eigenvalue = 13.674; Factor 2 (F2), eigenvalue = 1.509; Factor 3 (F3), eigenvalue = 1.132). These results suggest retention of the first three common factors.

Factors	Initial Eigenvalues
1	13.764
2	1.509
3	1.132
4	0.939
5	0.848
6	0.780
7	0.714
8	0.598
9	0.591
10	0.504
11	0.443
12	0.431
13	0.428
14	0.399
15	0.362
16	0.344
17	0.341
18	0.317
19	0.258
20	0.240
21	0.228
22	0.196
23	0.191
24	0.172
25	0.150
26	0.122

Table 5.10 Factors’ Initial Eigenvalues

5.3.2.2 Parallel Analysis Results

Table 5.11 presents results of the parallel analysis. I performed Principal Axis Factoring (PAF) with raw data permutations with 1,000 simulated datasets. The eigenvalues for the first two common factors that are based upon the actual data exceed the eigenvalues based upon the simulated data for those same common factors, indicating statistically significant eigenvalues ($p=.05$). I made this determination by comparing the eigenvalues for the first two common factors in the “Raw Data Eigenvalues” column towards the left side of Table 5.11 with the eigenvalues for the same first two common factors in the “95th Percentile Eigenvalues” column on the far right side of Table 5.11. After the first two common factors, the remaining eigenvalues based upon the actual data are smaller than the eigenvalues for the same factors based upon the simulated data, suggesting nonsignificance. Comparison of the eigenvalues for the third through the twenty-sixth factors in the “Raw Data Eigenvalues” column on the left of Table 5.11 with the eigenvalues for the third through the twenty-sixth factors in the “95th Percentile Eigenvalues” column on the far right of Table 5.11 makes this apparent. Taken together, the parallel analysis results suggest retention of the first two common factors. Appendix K lists the SPSS syntax for the parallel analysis.

Factors	Raw Data Eigenvalues	50 th Percentile Eigenvalues	95 th Percentile Eigenvalues
1	13.422936	.934043	1.057112
2	1.157534	.804354	.907927
3	.741705	.707404	.790966
4	.584191	.623563	.698691
5	.439226	.552003	.621369
6	.373118	.486158	.548883
7	.307873	.421159	.482995
8	.247436	.363392	.419606
9	.227552	.307734	.361823
10	.148517	.254698	.307694
11	.114208	.205478	.253372
12	.096299	.156580	.200343
13	.066305	.110046	.153031
14	.031638	.065681	.106263
15	.010626	.021732	.062270
16	-.008058	-.019917	.018365
17	-.012417	-.061295	-.025837
18	-.039364	-.100490	-.063991
19	-.076929	-.139139	-.104795
20	-.093221	-.177390	-.145390
21	-.125392	-.214712	-.184338
22	-.133349	-.251430	-.220532
23	-.146562	-.289629	-.261748
24	-.168788	-.327103	-.298585
25	-.191713	-.366127	-.336380
26	-.205382	-.413491	-.380977

Table 5.11 Parallel Analysis Results.

Results based upon 1,000 parallel datasets, 50th percentile (mean), and 95th percentile of the distribution and random data eigenvalues using permutations of the raw dataset. The eigenvalues in bold denote statistical significance at the 95th percentile, suggesting retention of two factors.

5.3.2.3 Scree Test Results

Figure 5.2 displays the scree plot of the initial eigenvalues for each possible factor. As shown in Figure 5.2, only the initial eigenvalue for the first factor lies above “the elbow” of eigenvalue plots. These results suggest retention of only the first common factor.

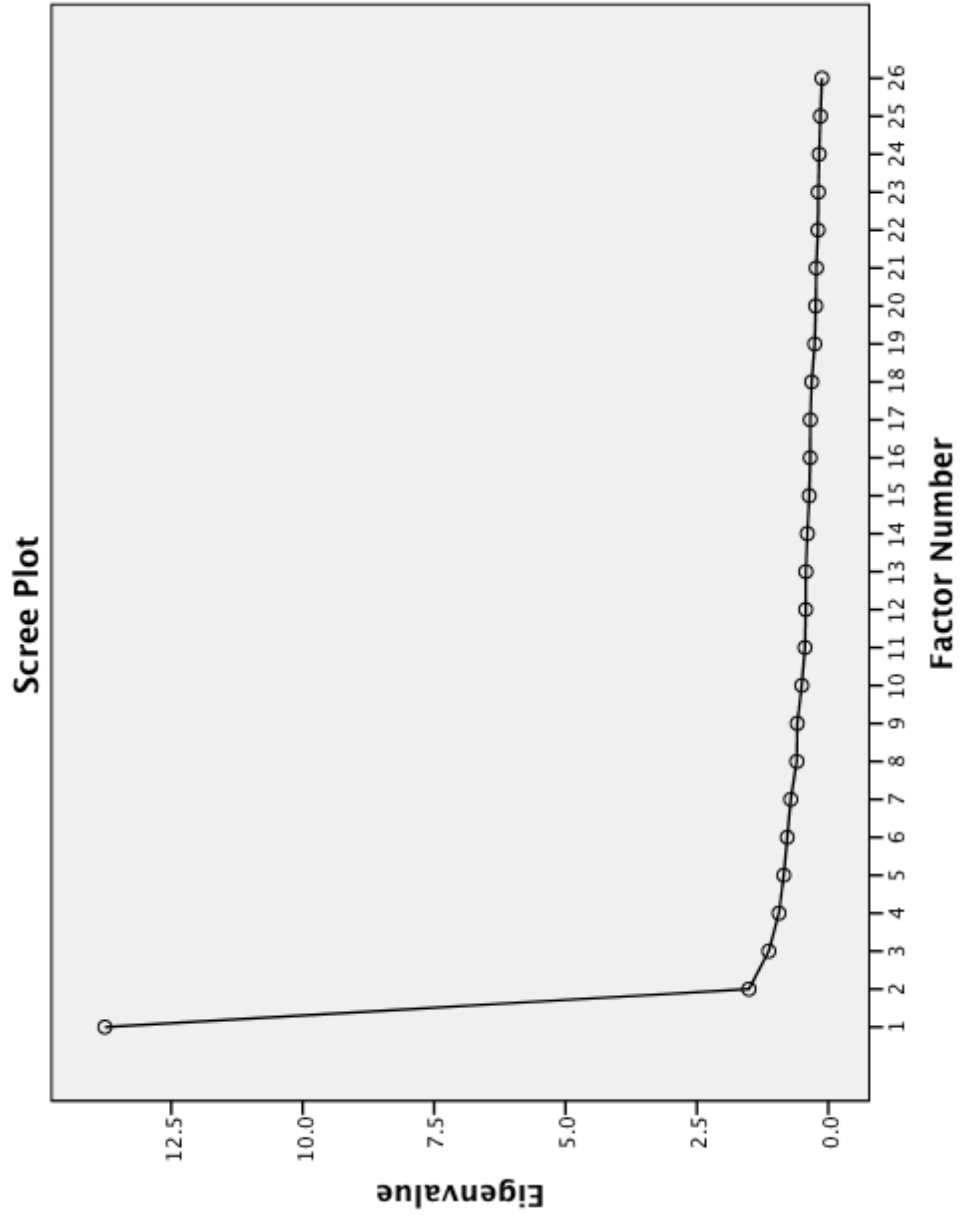


Figure 5.2 Scree Plot of Factors.

5.3.2.4 Results of Tests for Selecting the Number of Factors to Retain Summary

Taken together, results of the first two tests, Kaiser's criterion and parallel analysis, suggest retention of too many factors or overfactoring, while results of the third test, scree test examination, suggest retention of the appropriate amount of factors. Results of applying Kaiser's criterion suggest retention of the first three factors. Results of parallel analysis suggest retention of the same first two factors identified using Kaiser's criterion. In contrast, results of scree plot examination suggest retention of only the first factor.

In the case of Kaiser's criterion, the first three factors had eigenvalues larger than 1.0. While the first three factors satisfy Kaiser's criterion, the second and third factors just barely meet this criterion. The eigenvalue for the second factor is 1.509 and the eigenvalue for the third factor is 1.132. Such small eigenvalues indicate that the factors to which they correspond barely explain more than any individual item that those factors contribute to. This is because, in factor analysis, any individual item is often considered as having an eigenvalue of 1 (Kline, 2013). According to DeVellis (2012), when factors have eigenvalues just slightly above 1.0, they should not be retained, despite satisfying Kaiser's criterion.

Results of parallel analysis lead to the same type of overfactoring that the results of Kaiser's criterion suggest. While the eigenvalue of the second factor based upon the actual data is larger than the eigenvalue based upon randomized or simulated data for the same factor, the eigenvalue of the second factor is still too small to explain much more variance than any individual item that the second factor contributes to. According to DeVellis (2012):

What the scale developer often is after is a *parsimonious* account of the factors. That is, in the course of scale development, we often want to know about the few, most influential, sources of variation underlying a set of items, not every possible source we can ferret out. (p. 127)

Applying DeVellis's mantra to this study, retaining only the first factor explains the most variance while contributing to the fewest items.

Results of the scree test most effectively illustrate the importance of the first factor as well as the unimportance of the second and third factors. According to DeVellis (2012), "the vertical portion of the plot is where the substantial factors are located while the horizontal portion is the scree, or rubble, that should be discarded" (p. 129). Applying this analysis strategy to this study, I retained the first factor while discarding the second and third factors. As shown in Figure 5.2, the vertical portion of the plot, where the eigenvalue for the first factor lies, is high above the horizontal portion of the plot, where the second, third, and all remaining factors lie.

In summary, even though results of two tests suggest retention of multiple factors, I retained only the first factor for further analysis in this study. Retention of the first common factor explains the most variance using the fewest items, providing the most parsimonious account of the factors. Therefore, the remainder of this chapter reports on the first factor and the items it contributes to exclusively.

5.3.3 Percentage of Variance Explained by the Factor

Examination of the total variance explained by the factors helps to distinguish which factors are most important for understanding a particular construct (Kline, 2013). The items that the factors explaining the most variance contribute to are most worthy of consideration because those items most strongly reflect the construct that underlies them. Applied to this study, factors with high percentages of variance explained would do the most to explain user document trustworthiness perception. Further, the items those factors contribute to are most worthy of consideration, as they most pointedly reflect the construct of trustworthiness.

The left side of Table 5.12 lists the eigenvalues and percentages of total explained variance for each of the twenty-six factors in the initial solution extracted with Principal Axis Factoring (PAF) of the unreduced correlation matrix, the matrix containing the correlations of all the items before rotation (Kline, 2013). For example, the eigenvalue for the first initial factor is 13.764, so this factor accounts for $13.764/26 = .5294$ or 52.94% of the variance across all items. As expected, the next twenty-five initial factors explain successively decreasing percentages of variance, and all twenty-six factors together explain a total of 100% of the variance.

The right side of Table 5.12 reports the eigenvalues and percentages of explained variance for the three factors extracted with Principal Axis Factoring (PAF) of the reduced correlation matrix, the matrix containing the correlations of all the items after rotation (Kline, 2013). As expected, the eigenvalues and amount of variance explained by the three extracted factors are smaller because they are based

upon common variance, which excludes error and is always smaller than the total variance, which includes error (Kline, 2013). As shown on the right side of Table 5.12, the first PAF-extracted factor explains 51.405% of the total variance. This result suggests that the first common factor alone explains a majority of the common variance. Thus, focusing on this factor and the items it contributes to would reveal the most about the construct of trustworthiness using the smallest amount of items.

Factors	Initial			Reduced		
	Eigenvalues	% variance	Cumulative % variance	Eigenvalues	% variance	Cumulative % variance
1	13.764	52.940	52.940	13.365	51.405	51.405
2	1.509	5.804	58.743	1.108	4.260	55.665
3	1.132	4.353	63.097	.710	2.732	58.397
4	0.939	3.611	66.708			
5	0.848	3.262	69.970			
6	0.78	3.000	72.971			
7	0.714	2.746	75.716			
8	0.598	2.302	78.018			
9	0.591	2.271	80.289			
10	0.504	1.939	82.229			
11	0.443	1.703	83.932			
12	0.431	1.657	85.588			
13	0.428	1.645	87.233			
14	0.399	1.533	88.766			
15	0.362	1.391	90.157			
16	0.344	1.324	91.481			
17	0.341	1.313	92.794			
18	0.317	1.219	94.014			
19	0.258	.991	95.004			
20	0.24	.923	95.927			
21	0.228	.877	96.804			
22	0.196	.754	97.558			
23	0.191	.733	98.292			
24	0.172	.663	98.954			
25	0.15	.577	99.531			
26	0.122	.469	100.000			

Table 5.12 Eigenvalues and Percentages of Explained Variance for Analysis of an Exploratory 3-Factor Model of Document Trustworthiness Perception

5.3.4 Factor Interpretation

During factor interpretation, the researcher selects names or labels that synthesize the content of the items that load the highest on a given factor. Items with the highest factor loadings are assumed to reflect the concept underlying them (DeVellis, 2012). Factor interpretation involves four steps, including (Hair et al., 2010):

- (Step 1) identifying significant coefficients or loadings for each item,
- (Step 2) assessing the communalities of the items,
- (Step 3) respecifying the factor model if needed, and
- (Step 4) labeling the factors.

The following presents results from execution of these four steps.

5.3.4.1 Identification of Significant Loadings for Each Item Results

According to Hair et al. (2010), the process of identifying significant loadings for each factor involves:

start[ing] with the first [item] and moving horizontally from left to right, looking for the highest loading for that [item] on any factor. When the highest loading (largest absolute factor loading) is identified, it should be underlined if significant.... Attention then focuses on the second [item] and, again moving from left to right horizontally, looking for the highest loading for that [item] on any factor and underlining it. This procedure should continue for each [item] until all [items] have been reviewed for their highest loading on a factor. (p. 119)

In addition to the technique described above, the process of identifying significant loadings also involves detection of cross-loadings. A cross-loading occurs when an item loads significantly on more than one factor (Hair et al., 2010).

In this study, I applied the entire process of identifying significant loadings on a factor to analysis of both factor pattern and structure matrices. The factor pattern matrix includes the pattern coefficients or loadings for each item, which represent the unique contribution of each factor to each item (Hair et al., 2010; Jacoby, 1991). The factor structure matrix contains the structure coefficients or loadings for each item, which represent the correlations of each item with each factor (Hair et al., 2010). Both coefficients provide insight into which items specific factors contribute to the most, and, in this case, are necessary for understanding the items that most strongly reflect the construct of trustworthiness.

5.3.4.1.1 Items' Pattern Coefficients

Hair et al. (2010) point out that researchers typically examine the factor pattern matrix because it is easier to identify the unique contribution of each factor to each variable. Table 5.13 presents the factor pattern coefficients for the items of the factor retained in this study. The highest loadings of each item on any factor are underlined in the table. As shown in Table 5.13, all items' highest loadings are on the first factor. The items' highest factor loadings ranged between .928 and .374. According to Hair et al. (2010), there are two benchmarks regarding the evaluation of pattern coefficients. The first is the strictest, requiring items to have loadings greater than or equal to .50. The first nine items in Table 5.13 satisfy this criterion.

The second, less stringent benchmark states that loadings between .30 and .40 are minimally acceptable for retaining items. All of the items in Table 5.13 either meet or exceed this criterion. Consequently, I did not discard any of the items from further analysis on the basis of criteria pertaining to the magnitude of the pattern coefficients.

Only one item had a pattern coefficient cross-loading, which means that the item had a significant loading on more than one factor (Hair et al., 2010). As shown in Table 5.13, the item, “The document accurately reflects what happened” had a pattern coefficient of .451 on the first factor and a pattern coefficient of .310 on the third factor. Despite the cross-loading, I kept this item for two reasons:

- 1) I was interested in all items that loaded most strongly onto the first factor, because the first factor was the only one that explained a significant amount of variance.
- 2) I discarded the second and third factors, thus the fact that the item “The document accurately reflects what happened” had a cross-loading became irrelevant.

Items	Factors		
	1	2	3
The document is authentic.	<u>.928</u>	-.091	-.025
The document is factual.	<u>.926</u>	-.156	.033
The document includes documentation of where it came from.	<u>.772</u>	-.094	.137
The document was created using responsible and accepted practices.	<u>.678</u>	.122	.014
The digitized document is an actual picture of the original physical document.	<u>.677</u>	.062	-.193
The document is credible.	<u>.663</u>	.107	.107
The document appears free from error.	<u>.594</u>	.017	.166
The document is what it claims to be.	<u>.538</u>	.211	.069
The document is a primary source.	<u>.538</u>	.030	.070
The document accurately reflects what happened.	<u>.451</u>	-.088	.310
The document is official.	<u>.443</u>	.242	.128
The document was written at the time of the event.	<u>.374</u>	.295	.100

Table 5.13 Items' Pattern Coefficients

5.3.4.1.2 Items' Structure Coefficients

Table 5.14 presents the factor structure coefficients for the items of the factor retained in this study. The highest loadings of each item on any factor are underlined in the table. As shown in Table 5.14, all items' highest loadings are on the first factor. The items' highest factor loadings ranged between .842 and .589.

Although there are several structure coefficient cross-loadings (i.e., items with significant loadings on more than one factor), I expected this, since the factors themselves are highly correlated. Hair et al. (2010) point out, "as the correlation among factors becomes greater, it becomes more difficult to distinguish which [items] load uniquely on each factor in the structure matrix" (p. 119). The fact that there are several structure coefficient cross-loadings demonstrates that the factors are highly correlated. In the context of this study, this fact is insignificant because I

did not retain the second and third factors, as they explain very little variance as compared to the first factor (see Section 5.3.3). Therefore, the most important consideration regarding the structure coefficients is that they are the highest on the first factor, which explains more variance than any of the other factors.

Items	Factors		
	1	2	3
The document is authentic.	<u>.842</u>	.587	.555
The document is factual.	<u>.832</u>	.561	.565
The document is credible.	<u>.818</u>	.680	.644
The document includes documentation of where it came from.	<u>.797</u>	.582	.606
The document was created using responsible and accepted practices.	<u>.779</u>	.640	.572
The document is what it claims to be.	<u>.743</u>	.663	.592
The document appears free from error.	<u>.722</u>	.581	.591
The document is official.	<u>.714</u>	.666	.608
The document was written at the time of the event.	<u>.665</u>	.647	.570
The document is a primary source.	<u>.610</u>	.483	.465
The document accurately reflects what happened.	<u>.600</u>	.471	.560
The digitized document is an actual picture of the original physical document.	<u>.589</u>	.432	.321

Table 5.14 Items' Structure Coefficients

Factor loadings are important because they establish the degree to which a factor contributes to an item (Hair et al., 2010). Taken together, the results of Step 1 demonstrate the emergence of a set of items to which the first factor contributes the most. As comparison of Tables 5.13 and 5.14 demonstrates, the sequence of the

highest factor loadings on the first factor was the same, whether considering items' pattern or structure coefficients. Taken together, the results of this study confirm that the first factor contributes significantly to the items that loaded onto it the highest.

5.3.4.2 Assessment of Items' Communalities Results

Step 2 involves assessment of items' communalities. A communality "represent[s] the amount of variance accounted for by the factor solution for each [item]" (Hair et al., 2010, p. 119; Jacoby, 1991). I examined the communalities to assess whether I should retain the items for further analysis. Specifically, I compared the communalities to the benchmark of .40 or above for the retention of items in further analysis. According to Costello and Osborne (2005), if an item has a communality of less than .40, "it may either a) not be related to the other items, or b) suggest an additional factor that should be explored" (p. 4). Their comprehensive review of best practices regarding EFA includes recommendation for retaining items with final communalities of .40 or above, as these items are considered to be well enough explained by the factors.

Table 5.15 lists the items' final communalities. Final communalities represent items' communalities after factor rotation. As shown in Table 5.15, items' final communalities ranged between .364 and .714, with the average communality of .556. Most (8 or 66%) of the items' communalities were high (i.e., greater than .50). Two items' communalities ranged between .40 and .50. The communalities of the remaining two items, "The document is a primary source" and "The document is an

actual picture of the original physical document,” fell slightly below .40. These low communalities suggest that no factor contributed significantly to those items.

High communalities suggest that items are capable of loading significantly on a factor, which is critical to interpreting EFA results (Costello & Osborne, 2005). Results of computing items’ communalities suggest that most of the items have high communalities, but two items do not. Regardless, I retained the two items with low communalities because they only fell slightly below the recommended benchmark .40.

Items	Communalities
The document is authentic.	0.714
The document is factual.	0.702
The document is credible.	0.684
The document includes documentation of where it came from.	0.643
The document was created using responsible and accepted practices.	0.615
The document is what it claims to be.	0.580
The document is official.	0.555
The document appears free from error.	0.537
The document was written at the time of the event.	0.497
The document accurately reflects what happened.	0.403
The document is a primary source.	0.375
The document is an actual picture of the original physical document.	0.364

Table 5.15 Items’ Final Communalities.

5.3.4.3 Factor Model Respecification

According to Hair et al. (2010), respecification of the factor model may be necessary if a factor has too few items, an item has no significant loadings, an item’s communality is too low, or an item has a cross-loading. While having too few items for a factor was not an issue in this study, two items had low communalities (e.g.,

“The document is a primary source” and “The digitized document is an actual picture of the original physical document”) and one item had a cross-loading (e.g., “The document accurately reflects what happened”). Despite these results, I did not respecify the factor model for two reasons. First, the two items with low communalities fell only slightly below the recommended benchmark of .40. Second, it was irrelevant that one of the items had a cross-loading because I only retained the first factor.

5.3.4.4 Factor Labeling Results

During Step 4, factors are labeled. As Hair et al. point out (2010), factor labeling “is not derived or assigned by the factor analysis computer program; rather, [a factor] label is intuitively developed by the researcher based on its appropriateness for representing the underlying dimensions of a particular factor” (p. 120). Items with the highest factor loadings are considered demonstrative of the factors underlying them (DeVellis, 2012; Hair et al., 2010). Therefore, during Step 4, those items’ content are examined and afterwards form the basis of factor labeling (Hair et al., 2010).

In this study, I examined factor pattern coefficients. As previously stated, factor pattern coefficients represent the unique contribution of each factor to each item (Hair et al., 2010). In particular, I used items with coefficients of .65 or greater to interpret the factor that I retained for further analysis, as this benchmark constitutes a substantial loading on a factor (DeVellis, 2012). According to DeVellis (2012), “the items with the highest loadings are the ones that are the most similar to

the latent variable (and thus correlate most strongly). Therefore, they can provide a window into the nature of the factor in question” (p. 147).

Table 5.16 presents the items with their highest pattern coefficients that are greater than .65 on the factor that I retained. As shown in Table 5.16, the items are ordered by the magnitude of their pattern coefficients from greatest to least. I considered the item with the highest pattern coefficient, “The document is authentic,” as most influential in labeling the factor. It directly pertains to the concept of authenticity (Smith, 2000). I considered the item with the second highest pattern coefficient, “The document is factual,” as influential in labeling the factor. It pertains to the concepts of reliability and accuracy (Duranti, 1995). I considered the item with the third highest pattern coefficient, “The document includes documentation of where it came from,” as influential in labeling the factor. It pertains to the concept of provenance that relates to the concept of authenticity (Smith, 2000). I considered the item with the fourth highest pattern coefficient, “The document was created using responsible and accepted practices,” as influential in labeling the factor. It pertains to the concept of validity as described in Kelton et al. (2008) but is more commonly associated with the concept of reliability in the digital curation literature (e.g., Duff, 1998; Duranti, 1995; Duranti, Suderman, & Todd, 2008). I considered the item with the fifth highest pattern coefficient, “The digitized document is an actual picture of the original physical document,” as influential in labeling the factor. It pertains to the concepts of reliability and authenticity (Duranti, 1995). I also considered the item with the sixth highest pattern coefficient, “The document is credible,” as influential in labeling the factor. It

pertains to the concepts of trustworthiness and credibility (Rieh, 2010). Given these six best performing items, I chose to label the factor contributing to these items Trustworthiness.

Items	Factors		
	1	2	3
The document is authentic.	<u>.928</u>	-.091	-.025
The document is factual.	<u>.926</u>	-.156	.033
The document includes documentation of where it came from.	<u>.772</u>	-.094	.137
The document was created using responsible and accepted practices.	<u>.678</u>	.122	.014
The digitized document is an actual picture of the original physical document.	<u>.677</u>	.062	-.193
The document is credible.	<u>.663</u>	.107	.107

Table 5.16 Items' Pattern Coefficients with Highest Loadings on the First Factor that are Also Greater than .65

Although I focused on the six items with the highest pattern coefficients to label the factor, six other items also loaded onto the same factor strongly. Table 5.17 lists those other items and their pattern coefficients. These items largely pertain to the concepts of authenticity, reliability, and accuracy, which lend additional empirical support to naming the factor in terms of perceived trustworthiness. For example, the item with the eighth highest pattern coefficient, "The document is what is claims to be," corresponds to Duranti's (1995) classic definition of authenticity. The item with the tenth highest pattern coefficient, "The document accurately reflects what happened," corresponds to Duff et al.'s (2004) definition of reliability. The item with the lowest pattern coefficient, "The document was written at the time

of the event,” corresponds to notions of accuracy (Association for Information and Image Management, 1992).

Items	Factors		
	1	2	3
The document appears free from error.	<u>.594</u>	.017	.166
The document is what it claims to be.	<u>.538</u>	.211	.069
The document is a primary source.	<u>.538</u>	.030	.070
The document accurately reflects what happened.	<u>.451</u>	-.088	.310
The document is official.	<u>.443</u>	.242	.128
The document was written at the time of the event.	<u>.374</u>	.295	.100

Table 5.17 Items’ Pattern Coefficients that also Loaded Strongly on the First Factor

As a result of examining the content of items with the largest loadings on the retained factor, I applied the label, “Trustworthiness.” The label appropriately represents the underlying dimension of the factor. Although the factor contributes to a diverse set of items, at a higher level of abstraction, all the items address some aspect of perceived trustworthiness, hence the factor label.

5.3.4.5 Factor Interpretation Summary

Factor interpretation is a critical component of a scale development project (DeVellis, 2012). During factor interpretation, the researcher examines the items to determine the name for the retained factors. In this study, execution of the four steps for factor interpretation outlined in Hair et al. (2010) led to the labeling the retained factor. This process involved examining items’ factor pattern coefficients,

structure coefficients, and communalities. As a result, I labeled a very powerful factor containing twelve items “Trustworthiness.”

5.4 Reliability

The next step in the scale development project was to assess whether the final factor that I identified and labeled as a result of EFA and factor interpretation could serve as a scale for measurement. One way to assess the suitability of a set of items as a scale is to examine its reliability. A commonly used measure of reliability is internal consistency, which refers to the strength of the correlation among the items in a scale (Hair et al., 2010). Assessing a scale’s internal consistency involves computing a reliability coefficient. Also known as Cronbach’s alpha, the reliability coefficient measures how strongly a set of items intercorrelate (Cronbach, 1951). According to DeVellis (2012), “one of the most important indicators of a scale’s quality is the reliability coefficient, alpha” (p. 108).

To assess reliability, I followed two procedures. First, I calculated Cronbach’s alpha. Second, I performed item-deleted alpha analysis, whereby I analyzed each item separately to assess its impact on Cronbach’s alpha if I deleted it (DeVellis, 2012).

5.4.1 Cronbach’s Alpha Results

The value of Cronbach’s alpha for the scale that I derived from the factor retained during factor analysis was .931. This value for Cronbach’s alpha is well

above the benchmark of .70, which is common in the social sciences (Hair et al., 2010). According to DeVellis (2012), a Cronbach's alpha value greater than or equal to .90 is excellent. The results of calculating Cronbach's alpha suggest that the items the Trustworthiness factor contributed to could together serve as a highly reliable scale, as its items exhibit a high degree of internal consistency.

5.4.2 Item-Deleted Alpha Analysis Results

I compared results of item-deleted alpha analysis to assess whether or not exclusion of any particular item from further analysis would increase the internal consistency reliability of the scale. Table 5.18 presents results of the item-deleted alpha analysis. The left column lists the scale items and the right column demonstrates what value Cronbach's alpha would either increase or decrease to, given each item's deletion from the scale. As shown in Table 5.18, none of the items' deletion would cause Cronbach's alpha to increase; thus, the results suggest that the scale would not become more reliable as a result of deleting any of its items. These results offer strong empirical support for the scale derived from the Trustworthiness factor as is.

Items	Cronbach's Alpha if Item Deleted
The document is an actual picture of the original physical document.	.931
The document is a primary source.	.930
The document was written at the time of the event.	.928
The document appears free from error.	.926
The document is official.	.925
The document is what it claims to be.	.924
The document was created using responsible and accepted practices.	.924
The document is credible.	.923
The document includes documentation of where it came from.	.923
The document is factual.	.922
The document is authentic.	.922

Table 5.18 Item-Deleted Alpha Analysis Results Regarding the Scale Derived from the Trustworthiness Factor.

Table 5.19 lists the final set of items comprising the Digitized Archival Document Trustworthiness Scale (DADTS). I developed DADTS as a result of all the processes described in this chapter, including two tests regarding reliability and analysis of the resulting data. DADTS consists of twelve items. Table 5.19 lists items in order from the item that performed the best, i.e., the item that loaded most strongly onto the Trustworthiness factor, to the item that performed the weakest, i.e., the item that loaded onto the Trustworthiness factor least strongly.

Items	Pattern Coefficients
The document is authentic.	<u>.928</u>
The document is factual.	<u>.926</u>
The document includes documentation of where it came from.	<u>.772</u>
The document was created using responsible and accepted practices.	<u>.678</u>
The digitized document is an actual picture of the original physical document.	<u>.677</u>
The document is credible.	<u>.663</u>
The document appears free from error.	<u>.594</u>
The document is what it claims to be.	<u>.538</u>
The document is a primary source.	<u>.538</u>
The document accurately reflects what happened.	<u>.451</u>
The document is official.	<u>.443</u>
The document was written at the time of the event.	<u>.374</u>

Table 5.19 Items in the Digitized Archival Document Trustworthiness Scale (DADTS)

5.5 Validation Study

After determining the reliability of the Digitized Archival Document Trustworthiness Scale (DADTS), I assessed its validity. According to DeVellis (2012), validity refers to the extent to which an “item set continues to perform as the assigned name implies” (p.147). To assess validity, DeVellis (2012) recommends using a different sample of participants to ensure that the scale resulting from the EFA is not a mere quirk of the initial development sample (DeVellis, 2012).

5.5.1 Research Questions

The validation study had two aims. The first aim was to assess whether a different subgroup of genealogists who regularly utilize digitized archival documents would consider the items comprising DADTS as important with respect to their document trustworthiness perceptions. The second aim was to examine the

utility of the scale in measuring different levels of trustworthiness perception for specific digitized archival documents. The aims restated as research questions are:

1. To what extent would a different subgroup of genealogists who regularly utilize digitized archival documents consider the items comprising the scale as important with respect to their document trustworthiness perceptions?
2. To what extent can the scale measure trustworthiness perception regarding specific digitized archival documents?

5.5.2 Study Participants

I drew the study sample from the Genealogical Society of Washtenaw County (GSWC). Founded in 1974, the purpose of the GSWC is to “aid and assist members in genealogical studies, to encourage the collection and preservation of family and public records, and to promote the exchange of genealogical information” (Genealogical Society of Washtenaw County, 2014). The GSWC runs on a volunteer basis, has an active membership of 374, and conducts monthly meetings that are open to the public (M. McCrary, personal communication, July 9, 2014). I selected the GSWC for this study for three reasons. First, the GSWC is well established, having existed for well over three decades. Second, the GSWC has active membership. Third, the GSWC has a targeted mission of aiding their members in genealogical research.

I recruited the validation study participants with the help of the GSWC president. She provided access to the study participants during one of their monthly research meetings. In particular, the study participants were members of the

Research Committee for the GSWC. The research committee members handle all research requests that the GSWC receives. Comprised of eight members, the Research Committee was a suitable sample for this validation study, because, according to the President of the GSWC, each Research Committee member has gained experience utilizing digitized archival documents in the course of addressing the GSWC's research requests (M. McCrary, personal communication, July 9, 2014).

For the most part, the validation study participants were similar to the study participants who were involved in the development of the scale. All participants reported that they were over 60 years of age. Five were females and two were males. Most participants indicated that they use digital documents weekly (n=5), monthly (n=1), or a few times a year (n=1). They reported using death records (n=4), census records (n=2), or manuscripts (n=1) most frequently. In contrast to the participants involved in the development of DADTS, the validation study participants reported using seekingmichigan.org, a digital archive hosted by the Michigan History Collection and Archives of Michigan, rather than WADA. This was the most desirable use characteristic of the participants, as one of the primary aims of the validation study was to understand the importance of the items comprising DADTS to genealogists other than those who frequently utilize archival documents delivered digitally via WADA.

5.5.3 Study Details

To address the first research question, the validation study participants provided ratings for the scale. Their ratings served as evidence of the importance of those items to their document trustworthiness perceptions. To address the second research question, the validation study participants provided ratings for the scale items in relation to two digitized archival documents.

5.5.3.1 Instrumentation

To facilitate participants' evaluation of the scale items, I created a web survey. The survey included two digitized archival documents that participants were supposed to evaluate in relation to the scale items. Appendix L includes the survey. The remainder of this subsection describes the two digitized archival documents that the participants evaluated using the scale during the validation study: a digitized birth certificate (see Appendix M) and a digitized marriage certificate (see Appendix N).

5.5.3.1.1 The Digitized Birth Certificate

I used the digitized birth certificate in this study for two reasons. First, the digitized birth certificate was not a WADA document. It was necessary during the validation stage of the project to assess the appropriateness of the items comprising

the scale in measuring trustworthiness perception for documents beyond those found specifically in the context of WADA.

Second, there has been a lot of controversy regarding the perceived trustworthiness of the particular digitized birth certificate that I used in this study—the birth certificate of Barack Obama, the forty-fourth and current President of the United States as of the time of this study. Specifically, “birthers” claim that President Obama is not qualified for the presidency because they believe that he was not actually born in the United States (Hughey, 2012). Since being a natural-born citizen of the United States is a requirement for holding presidential office (U.S. Const. art. II, § 1), lack of natural-born citizenship is grounds for removal from office.

The controversy regarding President Obama’s natural-born citizenship began during his first presidential campaign (Smith & Tau, 2011). According to Smith and Tau (2011), a group of Hillary Clinton supporters sent out anonymous emails stating that Barack Obama was actually born in Kenya, not the United States. Another theory claimed that, since Barack Obama’s father was Kenyan and his mother was only 18 when he was born, he was not eligible for citizenship under existing immigration law (Smith & Tau, 2011). Yet another theory claimed that Obama lost his U.S. citizenship either when he lived in Indonesia or visited Pakistan (Smith & Tau, 2011). In order to settle the issue of whether or not Obama was born in the United States, Jim Geraghty asked members of the Obama campaign to release his birth certificate to the public (Geraghty, 2008). As a result, the Obama campaign released his birth certificate on their Fight the Smears website and also gave a copy

of his birth certificate to staff of the Daily Kos, which they posted on their website (Moulitsas, 2008; Smith & Tau, 2011).

Instead of settling the issue, release of Obama's birth certificate caused even more controversy. Some considered the birth certificate a forgery (Smith & Tau, 2011). Others criticized the released version of the birth certificate because it was the Certificate of Live Birth form, not Obama's original birth certificate (Smith & Tau, 2011). In 2001, the State of Hawaii changed its policy on vital records so that only computer-generated abstracts of birth and marriage records would be issued, as opposed to photocopies of original records, for the purpose of consistency in reporting information contained in those records (Watanabe, 2009). Although the U.S. Supreme Court recognizes Hawaii's Certification of Live Birth forms as official birth certificates (Watanabe, 2009), several politicians (e.g., Roy Blunt and Nathan Deal), celebrities (e.g., Donald Trump), and conspiracy theorists (e.g., Jerome Corsi) rejected Obama's Certification of Live Birth form (Associated Press, 2011; Corsi, 2008; Galloway, 2009; Justice v. Fuddy, 2011; Martin v. Lingle, 2008; Shear, 2011; Weigel, 2009). They requested release of his "original" birth certificate, also referred to as the long form certificate, which contains more details concerning the circumstances of his birth. After over two years of controversy and media attention devoted to the issue of President Obama's birth certificate, he released a digitized version of his "original," long form birth certificate on the White House website on April 27, 2011. During a press conference on that day, President Obama remarked:

We've had every official in Hawaii, Democrat and Republican, every news outlet that has investigated this, confirm that, yes, in fact, I was born in Hawaii, August 4, 1961, in Kapiolani Hospital. We've posted the

certification that is given by the state of Hawaii on the Internet for everybody to see. People have provided affidavits that they, in fact, have seen this birth certificate. And yet this thing just keeps on going. (The White House, Office of the Press Secretary, 2011)

Despite public release of President Obama's long form birth certificate, questions regarding the trustworthiness of his birth certificate persisted. For example, Maricopa County, Arizona Sheriff Joe Arpaio and his staff of volunteer investigators claimed that they completed a forensic analysis of the digitized birth certificate that is posted on the White House website and were able to confirm that it is a forgery (Associated Press, 2012; Vuoto, 2013). Arizona State officials have since rejected Arpaio's claims (Davenport & Billeaud, 2012; Fischer, 2012).

At any rate, the controversy regarding President Obama's birth certificate is a prime example of the importance of digitized archival document trustworthiness perception. President Barack Obama's birth certificate is just as trustworthy as any other vital record archived by the Department of Health of the State of Hawaii (State of Hawaii, Department of Health, 2008; 2009). Dr. Chiyome Fukino, then Director of the State of Hawaii's Department of Health, released the following statement:

I ... along with the Registrar of Vital Statistics who has statutory authority to oversee and maintain these type of vital records, have personally seen and verified that the Hawai'i State Department of Health has Sen. Obama's original birth certificate on record in accordance with state policies and procedures. No state official, including Governor Linda Lingle, has ever instructed that this vital record be handled in a manner different from any other vital record in the possession of the State of Hawai'i. (State of Hawaii, Department of Health, 2008)

Despite Hawaiian officials' efforts to verify Obama's birth certificate, for whatever reason, some people choose to perceive it as untrustworthy. Given the well-documented and highly publicized controversy regarding President Barack Obama's birth certificate, I included a digitized version of his original birth certificate in this study.

5.5.3.1.2 The Digitized Marriage Certificate

The other digitized archival document that was used in the study was a digitized marriage certificate. A digitized marriage certificate was chosen for two reasons. First, the particular digitized marriage certificate that was used in the study was a WADA document. Since the development sample was asked to provide their ratings of the trustworthiness items in relation to WADA documents, it was necessary to have the validation study participants provide ratings of the trustworthiness items for a WADA document. Inclusion of a WADA document addresses the issue of whether or not different subgroups of the population of genealogists perceive document trustworthiness differently for WADA documents. Second, a digitized marriage certificate was chosen for the study because it was the type of document that the development sample reported using most frequently (see Section 5.1.4.3). It was important to assess whether or not a different sample of users would consider the scale items similarly for the same type of document. Including a digitized marriage certificate allowed for this type of comparison.

Beyond these two reasons, there was no particular reason for including the specific marriage certificate that was used in the study. To my knowledge, there were no controversies surrounding the trustworthiness of the particular digitized

marriage certificate, as was the case with the digitized birth certificate. This contrast provided balance in the study.

5.5.4 Study Procedures

On Monday, July 21, 2014 at 11AM, the validation study took place in the Family History Library at the Church of Jesus Christ of Latter-Day Saints (Saline Ward) (525 Woodland Drive, Saline, MI). Using computers in the Family History Library, each Research Committee member took the web-based Digital Archive Document Trustworthiness Perception Survey (see Appendix L). The survey consisted of four main parts: an introduction and instructions; presentation of two digitized archival documents for participants' examination and comment (a digitized birth certificate and a digitized marriage certificate); the scale items for participants' evaluation; and digitized archival document usage as well as demographics questions.

While taking the web-administered survey, half of the participants viewed the birth certificate first, and then they supplied ratings of the items comprising the scale based on their evaluation of the birth certificate. Afterwards, they viewed the marriage certificate, and then they supplied separate ratings of the items comprising the scale based on their evaluation of the marriage certificate. The other half of the participants viewed the marriage certificate first, and then they supplied ratings of the items comprising the scale based on their evaluation of the marriage certificate. Afterwards, they viewed the birth certificate, and then they supplied separate ratings of the items comprising the scale based on their evaluation of the

birth certificate. Participants provided their ratings for each item based on a 7-point scale, which ranged from -3, “very untrustworthy,” to +3, “very trustworthy.” The 7-point scale was used in the study to provide insight into the relative importance of the items comprising the scale in terms of participants’ trustworthiness perceptions specifically for the two digitized archival documents included in the study.

5.5.5 Results

Results of the validation study provide some indication of the importance of DADTS to genealogists beyond those who participated in its development. All of the means of the ratings that the validation study participants provided for the items comprising DADTS were above 4, the recoded midpoint of response choices. This was true for participants’ evaluation of both the birth certificate (see Table 5.20) and the marriage certificate (see Table 5.21). Recoded values of response choices above the midpoint, e.g., 5, 6, and 7, refer to positive document trustworthiness perception (i.e., 5 = “slightly trustworthy,” 6 = “trustworthy,” and 7 = “very trustworthy”). Thus, means above 4 suggest that, on average, the participants associated the items with positive document trustworthiness perceptions. These findings address the first aim of the study, which was to assess the extent to which the items comprising DADTS were relevant to a different subgroup of genealogists. In particular, these findings suggest that the items are relevant to another subgroup’s positive document trustworthiness perceptions.

Items	Means
The document includes documentation of where it came from.	6.57
The digitized document is an actual picture of the original physical document.	6.49
The document is factual.	6.29
The document was created using responsible and accepted practices.	6.29
The document was written at the time of the event.	6.29
The document is authentic.	6.22
The document is a primary source.	6.19
The document is what it claims to be.	6.00
The document accurately reflects what happened.	6.00
The document is credible.	5.71
The document is official.	5.71
The document appears free from error.	5.57

Table 5.20 Recoded Means of Validation Study Participants' Ratings for DADTS items (Birth Certificate)

Items	Means
The document is what it claims to be.	6.57
The digitized document is an actual picture of the original physical document.	6.49
The document accurately reflects what happened.	6.43
The document was written at the time of the event.	6.29
The document is authentic.	6.22
The document is a primary source.	6.19
The document is factual.	6.14
The document is credible.	6.14
The document is official.	6.14
The document includes documentation of where it came from.	6.14
The document was created using responsible and accepted practices.	6.00
The document appears free from error.	4.43

Table 5.21 Recoded Means of Validation Study Participants' Ratings for DADTS items (Marriage Certificate).

To address the second aim of the study, examination of the utility of DADTS in measuring different levels of document trustworthiness perception, two types of data were analyzed: items' minimums and maximums and the sum of items' means for both documents that participants viewed during the study.

Table 5.22 presents the minimums and maximums of the items based upon participants' evaluation of the birth certificate, and Table 5.23 presents the

minimums and maximums of the items based upon participants' evaluation of the marriage certificate. As both Tables 5.22 and 5.23 demonstrate, participants provided a broad range of responses per item. These findings suggest that the items are capable of discriminating against those who have differing trustworthiness perceptions. The ability of items to discriminate against those with different levels of the construct of interest is key for any scale (DeVellis, 2012).

Items	Minimums	Maximums
The document accurately reflects what happened.	3	7
The document is authentic.	3	7
The document is what it claims to be.	3	7
The digitized document is an actual picture of the original physical document.	4	7
The document is a primary source.	4	7
The document is official.	4	7
The document was written at the time of the event.	4	7
The document is credible.	4	7
The document appears free from error.	4	7
The document was created using responsible and accepted practices.	4	7
The document is factual.	5	7
The document includes documentation of where it came from.	6	7

Table 5.22 Recoded Minimums and Maximums of Validation Study Participants' Ratings for DADTS items (Birth Certificate)

Items	Minimums	Maximums
The document appears free from error.	1	7
The document is authentic.	3	7
The digitized document is an actual picture of the original physical document.	4	7
The document is official.	4	7
The document was written at the time of the event.	4	7
The document is a primary source.	4	7
The document accurately reflects what happened.	4	7
The document includes documentation of where it came from.	4	7
The document was created using responsible and accepted practices.	4	7
The document is factual.	5	7
The document is credible.	5	7

Table 5.23 Recoded Minimums and Maximums of Validation Study Participants' Ratings for DADTS items (Marriage Certificate)

Analysis of the average of the means for DADTS items in relation to the birth certificate and the marriage certificate suggest that DADTS is capable of measuring different levels of document trustworthiness perception in relation to different documents. The average of the means for each item for the birth certificate was 6.11. The average of the means for each item for the marriage certificate was 6.09. The difference in the average of the means for the DADTS items based upon the birth and marriage certificates suggest that document trustworthiness perception was slightly more positive for the birth certificate than the marriage certificate. Thus, these findings suggest that when applying DADTS to different digitized archival documents, it is able to detect differences in document trustworthiness perception. These results provide some indication of the utility of DADTS as a scale capable of measuring different levels of document trustworthiness perception for different documents.

In summary, results from this validation study suggest that:

- DADTS items are relevant to genealogists besides those whose responses were used to develop it,
- DADTS items are capable of discriminating against those with different trustworthiness perceptions, and

- DADTS can be applied to specific digital archival documents and detect differences in trustworthiness perceptions for those documents.

All of these findings underscore the utility of DADTS as a bona fide scale, not simply a mere quirk of the development sample that I used to create it.

5.6 Conclusion

The goal of this study was to assess the measurability of user document trustworthiness perception. Toward this end, I undertook a scale development project. This approach involved creating a comprehensive list of items that relate to the construct of trustworthiness, administering those items to a large and representative sample of users of digitized archival documents, and collecting those users' responses to the items. The value of item analysis, the results of which are the focus of this chapter, is that the processes associated with item analysis provided a means of culling a list of twenty-six items that a large sample of participants were able to evaluate in terms of their document trustworthiness perceptions down to twelve items that are most reflexive of the construct.

The Digitized Archival Document Trustworthiness Scale (DADTS) includes twelve items. Together these items represent what was most salient to a specific designated community of users with experience utilizing particular types of archival documents made available by one digital repository.

The results discussed in this chapter confirm the existence of a viable and reliable scale of trustworthiness—the Digitized Archival Document Trustworthiness Scale (DADTS). I built DADTS upon a sample of users with relevant, substantive

experiences using digitized archival documents. I also validated the relevance of the DADTS items with another sample of users different from the participants whose responses I used to develop DADTS. Findings from the validation study demonstrate the viability to measure the trustworthiness perceptions of others besides those who frequently utilize WADA documents.

5.7 References

- Associated Press. (2012, July 17). Obama birth record 'definitely fraudulent,' Sheriff Joe Arpaio says. *Yahoo! News*. Retrieved from <http://news.yahoo.com/arpaio-obama-birth-record-definitely-fraudulent-010211250.html>
- Associated Press. (2011, April 7). Trump goes after Obama on US citizenship, says citizenship questions remain unanswered. *Star Tribune*. Retrieved from <http://www.startribune.com/nation/119391219.html>
- Bartholomew, D. J., Steele, F., Moustaki, I., & Galbraith, J. I. (2008). *Analysis of multivariate social science data*. Boca Raton: CRC Press.
- Bartlett, M. S. (1954). A note on the multiplying factors for various χ^2 approximations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 16(2), 296-298.
- Baudoin, P. (2008). The principle of digital preservation. *The Serials Librarian*, 55(4), 556-559.
- Case, D. O. (2008). Collection of family health histories: The link between genealogy and public health. *Journal of the American Society for Information Science and Technology*, 59(14), 2312-2319.
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1(2), 245-276.
- Corsi, J. R. (2008). *The Obama nation: Leftist politics and the cult of personality*. New York: Threshold Editions.

- Costello, A. B., & Osborne, J. W. (2005). □Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research and Evaluation, 10*(7), 1-9.
- Couper, M. P. (2000). Review: Web surveys: A review of issues and approaches. *Public Opinion Quarterly, 64*(4), 464-494.
- Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334.
- Darby, P., & Clough, P. (2013). Investigating the information-seeking behaviour of genealogists and family historians. *Journal of Information Science, 39*(1), 73-84.
- Davenport, P., & Billeaud, J. (2012, May 23). Arizona officially accepts Obama's birth record. *Fox News Latino*. Retrieved from <http://latino.foxnews.com/latino/news/2012/05/23/arizona-officially-accepts-obama-birth-record/>
- DeVellis, R. F. (2012). *Scale development: Theory and applications*. Thousand Oaks, Calif.: SAGE.
- Duff, W. M., & Johnson, C. A. (2003). Where is the list with all the names? information-seeking behavior of genealogists. *American Archivist, 66*(1), 79-95.

- Duranti, L., Suderman, J., & Todd, M. (2008). Part seven-structuring the relationship between records creators and preservers: Policy cross-domain task force report. In L. Duranti, & R. Preston (Eds.), *International research on permanent authentic records in electronic systems (InterPARES) 2: Experiential, interactive and dynamic records* (pp. 1-42). Padova, Italy: Associazione Nazionale Archivistica Italiana. Retrieved from http://www.interpares.org/display_file.cfm?doc=ip2_book_part_7_policy_task_force.pdf
- Fischer, H. (2012, March 6). Brewer disagrees with Arpaio findings, believes Obama birth record is real. *East Valley Tribune*. Retrieved from http://www.eastvalleytribune.com/arizona/politics/article_7ceda460-67e1-11e1-977c-0019bb2963f4.html
- Fricker Jr., R. D. (2008). Sampling methods for web and E-mail surveys. In N. Fielding, R. M. Lee & G. Blank (Eds.), *The SAGE handbook of online research methods* (pp. 195-217). Los Angeles: SAGE.
- Galloway, J. (2009, November 5). Nathan Deal says he'll ask for Barack Obama's birth certificate. *Atlanta Journal-Constitution*. Retrieved from <http://blogs.ajc.com/political-insider-jim-galloway/2009/11/05/nathan-deal-says-hell-ask-for-barack-obamas-birth-certificate/>
- Genealogical Society of Washtenaw County. (2014). The Genealogical Society of Washtenaw County. Retrieved from <http://washtenawgenealogy.org/>

- Geraghty, J. (2008, June 9). Obama could debunk some rumors by releasing his birth certificate. *National Review Online*. Retrieved from <http://www.nationalreview.com/campaign-spot/9490/obama-could-debunk-some-rumors-releasing-his-birth-certificate>
- Hair, J. F., Black, W.C., Babin, B.J., & Anderson, R.E. (2010). *Multivariate data analysis* (7th ed.). Upper Saddle River, NJ: Prentice Hall.
- Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179-185.
- Hughey, M.W. (2012). Show me your papers! Obama's birth and the whiteness of belonging. *Qualitative Sociology*, *35*, 163-181.
- Jacoby, W. G. (1991). *Data theory and dimensional analysis*. Newbury Park, Calif.: Sage Publications.
- Justice v. Fuddy, H.I. 1ST Cir. (2011).
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, *20*(1), 141-151.
- Kaiser, H. F., & Rice, J. (1974). Little jiffy, mark iv. *Educational and Psychological Measurement*, *34*(1), 111-117.
- Kelton, K., Fleischmann, K. R., & Wallace, W. A. (2008). Trust in digital information. *Journal of the American Society for Information Science and Technology*, *59*(3), 363-374.
- Kline, R. B. (2013). Exploratory and confirmatory factor analysis. In Y. Petscher, C. Schatschneider & D. L. Compton (Eds.), *Applied quantitative analysis in education and the social sciences* (pp. 171-207). New York: Routledge.

- Lucas, S. A. (2008). The information seeking processes of genealogists. (Unpublished doctoral dissertation, Emporia State University).
- Martin v. Lingle, H.I. 1ST Cir. (2008).
- Moulitsas, M. (2008, June 12). Obama's birth certificate. *Daily Kos*. Retrieved from <http://www.dailykos.com/story/2008/06/12/534616/-Obama-s-birth-certificate>
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. New York: McGraw-Hill.
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, 32(3), 396-402.
- Rieh, S. Y. (2010). Credibility and cognitive authority of information. *Encyclopedia of Library and Information Sciences, Third Edition*, 1337-1344.
- Shear, M. D. (2011, March 1). Huckabee Questions Obama Birth Certificate. *The News York Times*. Retrieved from http://thecaucus.blogs.nytimes.com/2011/03/01/huckabee-questions-obama-birth-certificate-claims-he-was-raised-in-kenya/?_php=true&_type=blogs&r=0
- Smith, B., & Tau, B. (2011, April 22). Birtherism: Where it all began. *Politico*. Retrieved from <http://www.politico.com/news/stories/0411/53563.html>

State of Hawaii, Department of Health. (2008). Statement by Dr. Chiyome Fukino [News release]. Retrieved from

<http://health.hawaii.gov/vitalrecords/files/2013/05/08-93.pdf>

State of Hawaii, Department of Health. (2009). Statement by Health Director Dr. Chiyome Fukino, M.D. [News release]. Retrieved from

<http://health.hawaii.gov/vitalrecords/files/2013/05/09-063.pdf>

The White House, Office of the Press Secretary. (2011, April 27). Remarks by the President.

Retrieved from <http://www.whitehouse.gov/the-press-office/2011/04/27/remarks-president>

U.S. Const. art. II, § 1.

Vuoto, G. (2013, July 8). Forensic findings on Obama's birth certificate: 'A 100 percent forgery, no doubt about it.' *World Tribune*. Retrieved from <http://www.worldtribune.com/2013/07/08/forensic-findings-on-obamas-birth-certificate-a-100-percent-forgery-no-doubt-about-it/>

Watanabe, J. (2009, June 6). Born identity. *Honolulu Star-Bulletin*. Retrieved from http://archives.starbulletin.com/content/20090606_kokua_line

Weigel, D. (2009, July 29). And now, Roy Blunt. *The Washington Independent*.

Retrieved from <https://washingtonindependent.com/53127/and-now-roy-blunt>

Yakel, E. (2004). Seeking information, seeking connections, seeking meaning:

Genealogists and family historians. *Information Research*, 10(1)

Yakel, E., & Torres, D. A. (2007). Genealogists as a "community of records". *American Archivist*, 70(1), 93-113.

CHAPTER 6

DISCUSSION & CONCLUSION

According to the RLG-OCLC Working Group on Digital Archive Attributes (WGDAAs) (2002), one level of trust that applies to the establishment of Trusted Digital Repositories (TDRs) involves “how users trust documents provided to them by a repository” (p. 9). Further, the WGDAAs (2002) asserts that “a user must be able to trust digital documents provided by digital repositories” (p. 10). These quotations describe a mandate for digital repositories; digital repositories are responsible for the trust relationship between a user and a document within the context of a digital repository. Interestingly, what this mandate misses is what it means to a user for a digital document to be considered as trustworthy.

The purpose of this study was to measure trustworthiness perception, motivated by recognition of the extent to which trustworthiness is asserted as a property of a Trusted Digital Repository (TDR) instead of proposed as a measurement of user perceptions. This chapter summarizes the relationship between the two research questions at the crux of this dissertation: (1) How do members of a designated community conceptualize trustworthiness for documents they find in a digital repository? (2) To what extent are designated community

members' perceptions of document trustworthiness measurable? It also addresses the implications of the findings for understanding:

- trustworthiness as a concept,
- trustworthiness in a TDR context, and
- the relationship between document trustworthiness and repository trustworthiness.

To address both research questions, I implemented the methodology of scale development, which consists of four steps: 1) Construct Definition, 2) Generating an Item Pool, 3) Designing the Scale, and 4) Full Administration and Item Analysis.

To enhance execution of Steps 1 and 2 of scale development—Construct Definition (Step 1) and Generating an Item Pool (Step 2)—I conducted a focus groups study. The findings from the focus groups study addressed the first research question, How do members of a designated community conceptualize trustworthiness for documents they find in a digital repository? The findings from the focus groups study enhanced Step 1 of scale development by providing clarity by contextualizing trustworthiness from the perspective of actual users of digitized archival documents. The findings from the focus groups study enhanced Step 2 of scale development by providing key insights into actual users' terminology regarding trustworthiness; this allowed me to develop trustworthiness items using actual digitized archival document users' verbiage, thereby increasing the face validity of the items.

I conducted three focus groups with twenty-two members of the largest designated community of users of the Washington State Digital Archives (WADA),

i.e., genealogists, to uncover how they conceptualized digitized archival document trustworthiness. The study participants conceptualized trustworthiness in terms of a digitized archival document's perceived authenticity, accuracy, primary or first-hand nature, believability, validity, proper form, stability, legibility, coverage (i.e., completeness), and objectivity. Overall, the value of the focus groups study findings is that they identify the range of concepts that underlie these designated community members' perceptions of trustworthiness. More importantly, the focus groups study findings pinpointed those perceptions of trustworthiness that were most important to real-world users of digitized archival documents in TDRs, as participants most frequently mentioned trustworthiness in terms of perceived authenticity, accuracy, primary or first-hand nature, believability, validity, and proper form.

At a broader level, addressing Research Question 1 advances the field of digital curation regarding trustworthiness. The focus groups study findings change the terms of discussion regarding trustworthiness from a property of a document to a judgment that designated community members perceive.

After addressing Steps 1 and 2 of scale development, with the aid of the focus groups study, I culminated execution of Steps 3 and 4 of scale development— Designing the Scale (Step 3) and Full Administration and Item Analysis (Step 4)— with a quantitative survey research study involving 233 genealogists. The findings from this study addressed the second research question, To what extent are designated community members' perceptions of document trustworthiness measurable? After performing item evaluation, including exploratory factor analysis (EFA), on twenty-six items, I found that only one dimension or factor, the

Trustworthiness factor, explained a significant amount of variance, indicating that the twelve items which loaded onto *the Trustworthiness factor* were most critical for measurement of trustworthiness. Consequently, I derived the Digitized Archival Document Trustworthiness Scale (DADTS) from the *Trustworthiness factor* for measurement of trustworthiness.

At a broader level, results from addressing Research Question 2 offer a breakthrough in the ways TDR managers can engage the concept of trustworthiness. The Digitized Archival Document Trustworthiness Scale (DADTS) provides an example of a relatively simple measure of trustworthiness perception with items that were most critical to a sample of designated community members. Thus, DADTS is an example of a measurable operationalization of trustworthiness that can allow TDR managers the opportunity to actually monitor the trust relationship between a designated community member and a document in the context of a digital repository.

This chapter is as follows. First, I discuss the implications of the research findings for measurement of trustworthiness. Second, I examine insights the research findings provide regarding trustworthiness in a TDR context. Third, I situate the Digitized Archival Document Trustworthiness Scale (DADTS) within existing relevant literatures regarding trustworthiness. Fourth, I consider what the research findings contribute to understanding the relationship between document and repository trustworthiness. Fifth, I discuss limitations and delimitations of the research. Sixth, I provide recommendations for future research. Finally, I discuss

how measurement of trustworthiness advances the understanding and practice of TDRs.

6.1 Trustworthiness is Unidimensional with Multiple Nuances

The results of this study confirm that document trustworthiness perception is a unidimensional concept. At the same time, results of this study confirm that the single dimension of document trustworthiness perception includes multiple nuances, which relate to the concepts of authenticity, accuracy, reliability, and credibility, but are not separate dimensions of trustworthiness. Even though it may seem as though there is a contradiction between unidimensionality, which implies singularity, and multiple nuances, which implies multiplicity, there actually is not a tension at all. The remainder of this section explains how both unidimensionality and multiple nuances with respect to the concept of user document trustworthiness perception can coexist.

Results of this study confirm that document trustworthiness perception is a unidimensional concept. This is an important finding given some digital curation researchers' conceptualizations of trustworthiness. For example, MacNeil (2000) asserts that trustworthiness consists of two dimensions: reliability and authenticity. While results of this study confirm the importance of MacNeil's (2000) definitions of reliability and authenticity as nuances of a single trustworthiness dimension, results of this study do not provide empirical support for considering reliability and authenticity as dimensions themselves.

The fact that results of this study confirm that document trustworthiness perception is unidimensional has important implications for measurement. Here it is necessary to distinguish among dimensions, factors, scales, and subscales and to describe the goals of item analysis and factor analysis. Item analysis is necessary to identify items that correlate with each other (DeVellis, 2012). The assumption when items correlate highly with one another is that they correlate because they share a common dimension or factor (Bartholomew, Steele, Moustaki, & Galbraith, 2008; Jacoby, 1991). Factor analysis examines the correlations among items and groups of items to determine how many dimensions or factors underlie a set of items (Jacoby, 1991; Kline, 2013). Sometimes, only one factor or dimension underlies a set of items (i.e., unidimensionality); at other times, multiple dimensions underlie a set of items (i.e., multidimensionality) (Jacoby, 1991).

Comparing MacNeil's (2000) conceptualization of document trustworthiness with my study, one might expect that two dimensions, reliability and authenticity, would underlie the trustworthiness items that the participants evaluated. If this were the case, then a separate subscale comprised of items pertaining to the concept of reliability (to reflect the dimension of reliability) and a separate subscale comprised of items pertaining to the concept of authenticity (to reflect the dimension of authenticity) would both be used to measure document trustworthiness perception. This is not what I found.

After performing item analysis and factor analysis, I found that items pertaining to the concepts of reliability and authenticity were so highly correlated with each other that those items needed to join together in a single scale to measure

trustworthiness. In other words, I found only one dimension—trustworthiness. As a result, only one scale was necessary for measuring user document trustworthiness perception—the Digitized Archival Document Trustworthiness Scale (DADTS).

My findings do not weaken current definitions of document trustworthiness or distinctions among concepts that some digital curation researchers have associated with trustworthiness. My findings do, however, shed light on what researchers who endeavor to measure document trustworthiness within a digital repository context might encounter after running statistical analysis on their data. In particular, items pertaining to different concepts may come together in a single scale because they represent one dimension of document trustworthiness perception from an empirical, statistical standpoint. Albeit, that one empirical, statistical dimension may include multiple nuances.

The idea that distinctions among concepts do not always apply when measuring concepts is well documented in the literature on scale development. For example, DeVellis (2012) used the example of combining items pertaining to “Conscientiousness” and “Dependability” together in a single scale to underscore this point: “it might very well be, for example, that real-world data would support a single factor combining items about conscientiousness and dependability rather than separating them” (p. 146). In this study, real-world data collected from actual users of digitized archival documents supported the combination of items pertaining to reliability and items pertaining to authenticity in a single scale for measurement of trustworthiness.

Empirical findings from this study suggest that authenticity and reliability are not the only concepts that pertain to user document trustworthiness perception. I also found empirical support for measuring trustworthiness in terms of items that relate to the concept of accuracy. This is actually not surprising given the close association between the concepts of reliability and accuracy as they relate to some digital curation researchers' conceptualizations of trustworthiness. For example, according to Duranti, Suderman, and Todd (2008), "because, by definition, the content of a reliable record is trustworthy, and trustworthy content is, in turn, predicated on accurate data, it follows that a reliable record is also an accurate record" (p. 667). Thus, it would seem critical to include items that relate to the concept of accuracy in a scale for measurement of trustworthiness.

I found that items pertaining to factual accuracy (e.g., "The document is factual") and items pertaining to technical accuracy or errors pertaining to the technical environment of the digital realm (e.g., "The document appears free from error") both highly correlated with the trustworthiness factor. Similar to items pertaining to the concepts of reliability and authenticity, there was no empirical support for inclusion of items pertaining to the concept of accuracy in a separate scale (to reflect a dimension of accuracy) as part of measurement of trustworthiness. In this respect, accuracy is but a nuance of the larger trustworthiness dimension.

I also found empirical support for measuring the dimension of trustworthiness using an item that included the term "credible": "The document is credible." This empirical finding brings into question some of the traditions in the

Information Quality (IQ) and Web Credibility (WC) research literatures, where the converse is the case. Despite the departure of my empirical findings regarding the item “The document is credible” from IQ and WC studies, I retain this item in the Digitized Archival Document Trustworthiness Scale (DADTS) because of its statistical relevance to the other items in the scale.²

In sum, multiple items are necessary to measure all of the relevant nuances of user document trustworthiness perception. Use of single items to measure the concept of trustworthiness are not appropriate because they do not comprehensively sample the domain of the most important nuances of trustworthiness for users of digitized archival documents found within a digital repository context (Nunnally, 1978).

This study’s findings have implications for digital curation as well as information quality and web credibility researchers. Mainly, in each of these research domains, researchers should first find out what nuances of trustworthiness

² The fact that “The document is credible” is an item in a scale for measuring trustworthiness represents a conundrum in the information quality and web credibility research literatures. Some information quality and web credibility researchers use the terms credibility and believability synonymously (e.g., Fogg et al., 2001; Knight, 2008). In turn, some information quality researchers use items that include the term “trustworthy” to measure the concept of credibility or believability, which they consider a dimension of information quality (e.g., Lee, Strong, Kahn, & Wang, 2002). Other information quality researchers use the term “credibility” in items to measure believability (e.g., Knight, 2008). In the web credibility literature, some researchers define credibility in terms of trustworthiness (e.g., Rieh, 2010). Some web credibility researchers use the terms “trustworthy” and “believable” in items to measure credibility (e.g., Flanagin & Metzger, 2000). My dissertation does not resolve the issue of whether items that include the term “trustworthy” should be used to measure credibility, or items that include the term “credible” should be used to measure trustworthiness. In this study, “The document is credible” appears as an item in the Digitized Archival Document Trustworthiness Scale (DADTS) based upon overwhelming empirical evidence that I gathered as well as recommendations in the scale development literature for the inclusion of items in a scale (DeVellis, 2012).

matter most to their study participants, and then use items that can address those nuances in a scale for measurement of trustworthiness. Information quality researchers and web credibility researchers should refrain from using single items to measure trustworthiness perception at the document or information level. For example, information quality researchers and web credibility researchers should stop using single items such as “the document is trustworthy” or “the information is trustworthy” to measure trustworthiness perception. Further, digital curation researchers should not pick up this “bad habit” just because it occurs in other research disciplines, because measurement of trustworthiness using a single item does not actually provide much information about the concept.

6.2 Trustworthiness in a TDR Context

As stated in Chapter 2, Kelton, Fleischmann, and Wallace’s (2008) Integrated Model of Trust in Information provides the most explicit model of trustworthiness perception. According to Kelton et al. (2008), trustworthiness consists of four elements: accuracy (including currency, coverage, i.e., completeness, and believability), objectivity, validity, and stability. In their model, trustworthiness is a judgment that users assess based upon their perception of digital information as accurate, current, complete, believable, objective, valid, and stable. In Kelton et al.’s (2008) model, trustworthiness is part of a larger conceptual model of trust in digital information. Prior to my dissertation, Kelton et al.’s (2008) model had not been empirically tested.

Findings from this study represent two specific implications for Kelton et al.'s (2008) specification of trustworthiness: 1) the model is a viable conceptual model for trustworthiness perception of digitized documents in a TDR, and 2) the model needs to be adapted to better suit a TDR context. Evidence for the viability of Kelton et al.'s (2008) specification of trustworthiness but also for the need to adapt the model comes from three specific empirical observations during this dissertation:

- 1) Strong empirical support for trustworthiness perception in terms of accuracy, believability, and validity during the study. All three of these perceptions correspond to Kelton et al.'s (2008) specification of trustworthiness and should be maintained.
- 2) Strong empirical support for trustworthiness perception in terms of authenticity, primary or first-hand nature, and proper form. These three perceptions of trustworthiness are not part of Kelton et al.'s (2008) specification of trustworthiness but should be added.
- 3) Weak empirical support for trustworthiness perceptions in terms of coverage (i.e., completeness) and objectivity, and no empirical support for trustworthiness perceptions in terms of currency. Both of these perceptions of trustworthiness are part of Kelton et al.'s (2008) specification of trustworthiness but should be discarded.

The following subsection describes these three empirical observations in greater detail.

6.2.1 Motivation for Revisions to Kelton et al.'s Specification of Trustworthiness

In this subsection, I take into account how much empirical support I gathered regarding each trustworthiness perception that I identified in this study to clarify which trustworthiness perceptions are most important or more important than others. As a result, findings from this dissertation underscore the importance of authenticity, accuracy, believability, primary or first-hand nature, validity, and proper form perceptions of trustworthiness. The findings also demonstrate that stability, legibility, coverage (i.e., completeness), objectivity, and currency perceptions of trustworthiness are relatively less important.

During analysis of the focus groups study data, I operationalized importance as the frequency with which participants mentioned specific trustworthiness perceptions. I found that participants articulated their perceptions of trustworthiness in terms of perceived authenticity, accuracy, primary or first-hand nature, believability, validity, and proper form more frequently than they articulated their perceptions of trustworthiness in terms of perceived stability, legibility, coverage (i.e., completeness), and objectivity.

During item analysis, I operationalized important items as those items that loaded onto the Trustworthiness factor, the factor explaining the most amount of variance and therefore containing the most critical items for measurement of trustworthiness. Those items pertained to perceived accuracy, authenticity, believability, primary or first-hand nature, validity, and proper form.

In contrast, I found minimal empirical support for trustworthiness in terms of perceived objectivity. For example, during the focus groups study, participants mentioned trustworthiness in terms of perceived objectivity least frequently. During item analysis, more than ten participants rated items pertaining to objectivity as not applicable to their experience of evaluating the trustworthiness of digitized archival documents. These findings make sense given the fact that the participants were considering digitized archival documents. Archival documents are examples of historical documents; the presence of bias can underscore documents' historical accuracy (Trouillot, 1995). The participants understood that they were dealing with digitized archival documents, and thus perceived objectivity was not much of a concern.

I found no empirical support, either in the focus groups study or the item analysis study, for trustworthiness in terms of perceived currency as Kelton et al. (2008) apply the term in their specification of trustworthiness. These findings make sense because participants were considering their experience utilizing digitized archival documents as they articulated their trustworthiness perceptions. Digitized archival documents by definition are non-current.

Taken together, the focus groups study findings and item analysis study findings triangulate regarding which trustworthiness perceptions are more important than others thereby underscoring the validity of the dissertation findings. These findings advance understanding of the concept of trustworthiness with empirical evidence that not all trustworthiness perceptions that Kelton et al. (2008)

specify in their model have equal weight from the perspective of real world users of preserved digital documents.

6.2.2 Revisions to Kelton et al.'s (2008) Specification of Trustworthiness

Based on this dissertation's findings, I revise Kelton et al.'s (2008) trustworthiness specification to incorporate only those trustworthiness perceptions that most strongly correspond to the real world users of preserved digital documents who participated in this study. Figure 6.1 displays the revised conceptual trustworthiness model. As shown in Figure 6.1, I retain Kelton et al.'s (2008) specification of trustworthiness in terms of accuracy, believability, and validity and discard Kelton et al.'s (2008) specification of trustworthiness in terms of stability, coverage (i.e., completeness), objectivity, and currency. I also add authenticity, primary or first-hand nature, and proper form to their model, as I gathered substantial empirical support for these trustworthiness perceptions, and yet they were not included in Kelton et al.'s (2008) original model.

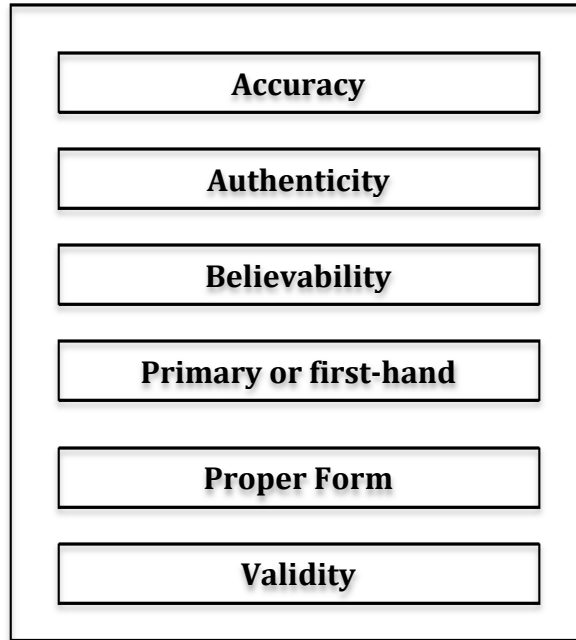


Figure 6.1 Conceptual Trustworthiness Model

Adapted from Kelton et al.'s (2008) Specification of Trustworthiness in their Integrated Model of Trust in Information.

6.2.2.1 Advantages and Disadvantages of Conceptual Trustworthiness Model

The revised conceptual trustworthiness model offers one primary advantage as well as one primary disadvantage. The obvious advantage of the revised model is its parsimony. It excludes perceptions of trustworthiness that participants provided little empirical support for during the focus groups study and the item analysis study. The model reflects the most important aspects of trustworthiness for actual designated community members who utilize digitized archival documents in a TDR context. The disadvantage of the revised model is that it may leave out aspects of trustworthiness that are actually important, despite a lack of empirical support for those aspects of trustworthiness during this study.

6.2.3 The Digitized Archival Document Trustworthiness Scale (DADTS)

Chief among Kelton et al.'s (2008) recommendations for research on trust was development of “an instrument for measuring trust in information” (p. 371). This dissertation addresses Kelton et al.'s call for research by operationalizing the revised version of their specification of trustworthiness (Figure 6.1) with a psychometric instrument—the Digitized Archival Document Trustworthiness Scale (DADTS) (see Table 6.1).

As shown in Table 6.1, DADTS includes twelve items. DADTS combines the twelve items pertaining to concepts, such as authenticity, accuracy, reliability, and credibility, in a single scale based upon overwhelming empirical support that only one dimension or factor, the Trustworthiness factor, underlies these items. The Trustworthiness factor alone explained 51.41% of the total variance with an eigenvalue of 13.764. The high eigenvalue means that the Trustworthiness factor is worth the value of nearly fourteen trustworthiness items; thus, the Trustworthiness factor is very powerful and influential (Kline, 2013). The items are rank-ordered by the pattern coefficients. Pattern coefficients numerically represent the influence of the factor on each item, i.e., how strongly each item reflects the concept of interest, which, in this case, is trustworthiness (Kline, 2013). The remainder of this section discusses DADTS in relation to Kelton et al.'s (2008) specification of trustworthiness as well as the advantages and disadvantages of using DADTS as a trustworthiness measurement model.

Items	Pattern Coefficients	Concepts the Items Represent
1. The document is authentic.	.928	Authenticity
2. The document is factual.	.926	Accuracy
3. The document includes documentation of where it came from.	.772	Provenance (Authenticity)
4. The document was created using responsible and accepted practices.	.678	Reliability
5. The digitized document is an actual picture of the original physical document.	.677	Reliability and Authenticity
6. The document is credible.	.663	Credibility
7. The document appears free from error.	.594	Accuracy
8. The document is what it claims to be.	.538	Authenticity
9. The document is a primary source.	.538	Reliability and Accuracy
10. The document accurately reflects what happened.	.451	Reliability
11. The document is official.	.443	Reliability
12. The document was written at the time of the event.	.374	Accuracy and Reliability

Table 6.1 The Digitized Archival Document Trustworthiness Scale (DADTS)

The Digitized Archival Document Trustworthiness Scale (DADTS) items pertain to the accuracy, believability, and validity aspects of Kelton et al.'s (2008) specification of trustworthiness. For example, the items, "The document appears free from error," and "The document is factual," pertain to perceived accuracy. The item, "The document is credible," pertains to perceived believability. The item, "The document was created using responsible and accepted practices" pertains to perceived validity.

DADTS items also pertain to authenticity, primary or first-hand, and proper form trustworthiness perceptions I identified during the focus groups study. For example, the item, "The document is what it claims to be" pertains to perceived

authenticity. The item, “The document is a primary source” pertains to perceived primary or first-hand nature. The item, “The document is official” pertains to perceived proper form. Taken together, the items comprising DADTS correspond to the six most important trustworthiness perceptions that I identified during this study: accuracy, authenticity, believability, primary or first-hand nature, validity, and proper form.

6.2.3.1 Advantages and Disadvantages of Trustworthiness Measurement Model

As a trustworthiness measurement model, the Digitized Archival Document Trustworthiness Scale (DADTS) offers two primary advantages as well as one primary disadvantage. In terms of advantages, first, DADTS operationalizes a conceptual model of trustworthiness perception. Prior to this dissertation, no scale for measurement of trustworthiness perception for any type of preserved digital information existed. DADTS demonstrates the power and utility of carrying forward a conceptual understanding of trustworthiness with measurement of the concept in concrete, empirical terms. The second advantage of DADTS is its parsimony. Its items exclude perceptions of trustworthiness that participants provided little empirical support for during the item analysis study therefore including items that reflect the most important aspects of trustworthiness for a group of designated community members who utilize specific types of digitized archival documents within a TDR context. Thus, the Digitized Archival Document Trustworthiness Scale (DADTS) measures the most important aspects of trustworthiness that I identified during this study. The disadvantage of using DADTS as a trustworthiness

measurement model is that it may leave out aspects of trustworthiness that are actually important, despite a lack of empirical support for those aspects of trustworthiness during this study.

6.2.3.2 Recommendations

Based on the most compelling empirical findings from this dissertation, I recommend measurement of a revised version of Kelton et al.'s (2008) specification of trustworthiness using the Digitized Archival Document Trustworthiness Scale (DADTS) for two reasons. First, DADTS takes into account only the most important trustworthiness perceptions. Second, DADTS is parsimonious.

6.3 The Digitized Archival Document Trustworthiness Scale in Relation to the Literature

Beyond Kelton et al.'s (2008) specification of trustworthiness, the Digitized Archival Document Trustworthiness Scale (DADTS) relates to the digital curation and web credibility literatures more broadly. This section maps the Digitized Archival Document Trustworthiness Scale (DADTS) as an empirical finding of this study to the digital curation and web credibility literatures. This section concludes with implications of this study's findings for understanding the concept of trustworthiness including its measurement.

6.3.1 The Best Performing Item: “The document is authentic”

The best performing item in DADTS is “The document is authentic.” The item had the largest and strongest loading onto the Trustworthiness factor, i.e., a factor loading (i.e., pattern coefficient) of .928. This finding demonstrates that the item is very highly correlated with the Trustworthiness factor. The importance of authenticity has been well documented in the digital curation literature (Smith, 2000b). What this study contributes to understanding authenticity is that the concept is chief among the study participants’ concept of trustworthiness; authenticity matters to actual users and affects their perception of documents as trustworthy.

It is difficult to ascertain what perceptions of authenticity the study participants had in mind while providing ratings for this particular item. The item, “The document is authentic” includes no specific definitions of the concept. Regardless, it is clear from the empirical results that, whatever it means to a user to be authentic, users have a high regard for the concept and it is primary among their concerns regarding the trustworthiness of digitized archival documents.

Kim (2009), Nunnally (1978), and Nunnally and Bernstein (1994) warn against using a single item to measure a concept because single items fail to comprehensively sample the totality of what a concept can mean. Despite this warning against using a single item to measure a construct, I used the item “The document is authentic” to measure document trustworthiness perception to assess its performance. Results of the study confirm that the item was the best performing item out of all of the items the participants evaluated during the study.

In sum, even though it is unclear exactly what authenticity meant to the participants who evaluated the item “The document is authentic” during the study, it was the best performing item. Thus, it is clear that authenticity means something very important with respect to the participants’ concept of trustworthiness.

6.3.2 The Second Best Performing Item: “The document is factual”

The second best performing item is “The document is factual.” I derived this item from participants during the focus groups. This item refers to the concept of accuracy, in particular the concept of factual accuracy. Some definitions of accuracy in digital curation pertain to the truthfulness of information, which is synonymous with the extent to which something, in this case, a digitized archival document, is factual (Duranti & Preston, 2008). Web credibility researchers would consider the concept of “factual” as related to trustworthiness, but distinct from it (Rieh & Danielson, 2007). In contrast, my findings suggest a strong and very close relationship between factual accuracy and trustworthiness, such that an item pertaining to perceived factual accuracy perception be used to measure document trustworthiness perception for archival documents delivered digitally.

6.3.3 The Third Best Performing Item: “The document includes documentation of where it came from”

The third best performing item was “The document includes documentation of where it came from.” I derived this item from participants during the focus

groups. This item pertains to the concept of provenance, which is used to support claims of authenticity. Applied to digital documents, provenance refers to the origins of the documents and change history (Baudoin, 2008). As Smith (2000a) points out, “the role of provenance is as important in the digital world as in the analog world, if not more so” (p. 71). Results of this study confirm that provenance information is important in the digital realm because it is critical to actual users’ concept of trustworthiness. Thus, items pertaining to provenance should be included in a scale to measure user document trustworthiness perception.

6.3.4 The Fourth Best Performing Item: “The document was created using responsible and accepted practices”

The fourth best performing item, “The document was created using responsible and accepted practices” is part of an explicit definition of validity in Kelton et al. (2008). Comparison of this item with the digital curation literature suggests that this item is more accurately described as a definition for reliability in terms of trustworthiness. For example, Duranti (1995) defines trustworthiness at the document level in part in terms of reliability: “reliability is provided to a record by its form and procedure of creation” (p. 6). As another example, Gilliland-Swetland and Eppard (2000) contend that, “the degree to which a record can be considered reliable is dependent upon the level of procedural and technical control exercised during its creation and management in its active life” (n.p.).

In addition, the item also corresponds to notions of authenticity. As Bearman and Trant (1998) point out, “authenticity (or content assurance) addresses what the

resource purports to be and how it was created and is, therefore, grounded in methods of identification and verification of the resource” (n.p.). The key aspect of their definition with respect to this particular item is emphasis on how the document was created, which empirical results of this study confirm matters to users in terms of their perceptions of digitized archival documents as trustworthy.

6.3.5 The Fifth Best Performing Item: “The digitized document is an actual picture of the original physical document”

The fifth best performing item is “The digitized document is an actual picture of the original physical document.” I derived this item from participants during the focus groups. The item conforms to the concepts of reliability and authenticity in the digital curation literature. Duranti (1995) points out that a document, such as the one described in the item above, is reliable because it well enough maintains “the form of original” (p. 7). When a digitized archival document is a picture of the original physical document, participants felt they understood the relationship between the original and what they were seeing, and that the contrast between the two was not only acceptable, but also worthy of positive document trustworthiness perception.

6.3.6 The Sixth Best Performing Item: “The document is credible”

The sixth best performing item was “The document is credible.” In the digital curation literature, for example, Duff (1998) links credibility to reliability, with

reliability serving as a component of credibility: “The mere existence of a record does not ensure that it will faithfully represent a transition or an event; its credibility must be ensured through the establishment of reliable methods and procedures for its creation, maintenance, and use over time” (p. 88).

Some web credibility researchers would argue that considering trustworthiness in terms of credibility is inverted and incorrect. For example, Fogg (2003) defines credibility in part in terms of believability, not believability in terms of credibility. Rieh (2010) defines credibility in part in terms of trustworthiness: “credibility is defined as people’s assessment of whether information is trustworthy based on their own expertise and knowledge” (p. 1338). In contrast, results of this study provide strong empirical support for a credibility conceptualization of trustworthiness. I found that an item pertaining to credibility can be used to measure trustworthiness.

6.3.7 The Seventh Best Performing Item: “The document appears free from error”

The seventh best performing item was “The document appears free from error.” Beyond Kelton et al.’s (2008) definition of accuracy, Duranti and Preston’s (2008) definition of accuracy includes the statement “free of error or distortion” (p. 769).

6.3.8 The Eighth Best Performing Item: “The document is what it claims to be”

The eighth best performing item is “The document is what it claims to be.” The item conforms to Duranti’s (1995) classic definition of authenticity, which she defines as part of document or record trustworthiness.

6.3.9 The Ninth Best Performing Item: “The document is a primary source”

The ninth best performing item is “The document is a primary source.” I derived this item from participants during the focus groups. The item pertains to the concepts of reliability and accuracy. Linking both accuracy and reliability together, Duranti and Preston (2008) contend that accuracy “is usually inferred on the basis of the degree of the records’ reliability and is only verified when such degree is very low” (p. 452). In addition, according to the Association for Information and Image Management (1992), records produced within a short period after the event or activity tend to be more readily acceptable as accurate than records produced long after the event or activity. Applied to this study, participants considered primary sources as closer to the events that they were about, and thus they perceived those documents as more accurate, and in turn, they perceived those documents as trustworthy. In this respect, my study confirms associations among the concepts of accuracy, reliability, and trustworthiness vis-à-vis primary sources specifically for actual users of digitized archival documents.

6.3.10 The Tenth Best Performing Item: “The document accurately reflects what happened”

The tenth best performing item is “The document accurately reflects what happened.” This item explicitly refers to Duff et al.’s (2004) definition of reliability, which they assert is a component of trustworthiness of archival sources broadly defined. The item also relates to Duranti’s (1995) definition of reliability. According to her, “reliability refers to the authority and trustworthiness of the records as evidence, the ability to stand for the facts they are about” (p. 6). Considering both Duff et al. (2004) and Duranti’s (1995) definitions of reliability, the item, which focuses on a document accurately reflecting what happened, is perhaps most able to stand for the facts it is about.

6.3.11 The Eleventh Best Performing Item: “The document is official”

The eleventh best performing item is “The document is official.” I derived this item from participants’ statements during the focus groups. Additional support for this item comes from the digital curation literature. For example, Duranti (1995) states that:

reliability is provided to a record by its form ... The form of a record is the whole of its characteristics that can be separated from the determination of the subjects, persons, or places the record is about. A record is regarded as reliable when it possesses all the elements that are required by the socio-juridical system in which the record is created for it to be able to generate consequences recognized by the system itself. (p. 6)

Less stringent than Duranti’s (1995) definition, the participants in this study perceived documents as trustworthy when they appeared to be official, which

implies a certain look or form for the participants. The participants believed that official documents ought to look a certain way and they believed that they were capable of perceiving what official documents should look like.

6.3.12 The Twelfth Best Performing Item: “The document was written at the time of the event”

The twelfth best performing item is “The document was written at the time of the event.” I derived this item from participants’ statements during the focus groups study. This item corresponds to the concept of factual accuracy as described in the archival science and digital curation literatures. For example, according to the Association for Information and Image Management (1992), records produced within a short period after the event or activity tend to be more readily acceptable as accurate than records produced long after the event or activity. Applied to this study, participants considered documents that were written at the time of the event as likely being more accurate than they would be if they were written at a later date. Thus, they perceived those documents as trustworthy.

6.4 Implications: The Relationship Between Document and Repository Trustworthiness

My study contributes to the literature on trustworthiness within the field of digital curation. Some prior researchers suggest that repository trustworthiness and document trustworthiness are separate and distinct phenomena (Yakel et al., 2013).

Other researchers recommend more research to better understand the relationship between trustworthiness at the repository level and trustworthiness at the document or data level (Yoon, 2014). My contribution to this literature is that repository and document trustworthiness are related. Comparison of DADTS items to criteria in standards for trustworthy digital repositories demonstrates the relationship between repository and document trustworthiness.

Essentially, requirements in international standards, such as *Criteria for Trustworthy Digital Archives* (DIN 31644) and *Audit and Certification of Trustworthy Digital Repositories* (ISO 16363), directly pertain to actions regarding archival management of digital documents that affect designated community members' perceptions of those documents as trustworthy.

For example, Criterion C19 in DIN 31644 requires digital repositories to, “[allow] the users and the administration of the digital archive[s] to check and maintain the authenticity of the representations” (NESTOR Certification Working Group, 2013, p. 25). In the context of DIN 31644, representations refer to the perceptible manifestations of digital information that repositories make accessible to their designated communities. Applied to this study, digitized archival documents are representations. Criterion C19 directly relates to the DADTS item, “The document is what it claims to be,” which serves as a definition of authenticity. This study’s findings underscore the importance of Criterion C19, which requires repositories to allow users to check the authenticity of digital documents, because designated community members’ perception of digitized archival documents as authentic directly affects their perception of those documents as trustworthy.

In another example, Criterion 4.6.2 in ISO 16363 requires repositories to, “follow policies and procedures that enable the dissemination of digital objects that are traceable to the originals, with evidence supporting their authenticity” (International Organization for Standardization, 2012, p. 62). Criterion 4.6.2 relates to the DADTS item, “The document includes documentation of where it came from.” Documentation regarding where a document comes from provides the information necessary to facilitate assessment of a document’s traceability. Thus, following policies and procedures that enable the dissemination of digital objects that are traceable to the originals, which is a requirement for repository trustworthiness, directly relates to documentation regarding where documents come from, which this study’s findings demonstrate affect designated community members’ perceptions of digitized archival documents as trustworthy. Criterion 4.6.2 also relates to the DADTS item, “The document is what it claims to be,” because the item serves as a definition of authenticity, and Criterion 4.6.2 is designed to provide evidence in support of authenticity.

In another example, Criterion 4.2.6.3 in ISO 16363 states that repositories must ensure that Preservation Description Information, including provenance information, be persistently associated with the relevant Content Information (International Organization for Standardization, 2012). Preservation Description Information and Content Information are two terms established in the Open Archival Information System (OAIS) model that is used in ISO 16363. Specifically, Preservation Description Information refers to, “information which is necessary for adequate preservation of the Content Information and which can be categorized as

Provenance, Reference, Fixity, and Context Information” (Consultative Committee on Space Data Systems, 2002, p. 1-12). The requirement that TDRs collect and maintain provenance information also relates to TDRs’ responsibility to uphold digital documents’ archival quality, one definition of which “could be seen as synonymous with the concept of record or directly associated with the maintenance of provenance” (Conway, 2011, p. 296). Content Information refers to, “the set of information that is the original target of preservation. It is an Information object comprised of its Content Data Object and its Representation Information” (Consultative Committee on Space Data Systems, 2002, p. 1-8). Applied to this study, digitized archival documents are Content Information and Preservation Description Information, in particular provenance, is the type of information described in the DADTS item, “The document includes documentation of where it came from.” Thus, it is important for repositories to persistently associate Preservation Description Information, in particular, provenance, with digitized archival documents, because findings from this study demonstrate that that is exactly the type of information that designated community members rely upon to inform their judgments of digitized archival documents as trustworthy.

In another example, Criterion 4.2.9 specifies that, “it is the responsibility of the repository to choose the appropriate mechanism for checking the completeness and correctness of its collections” (International Organization for Standardization, 2012, p. 51). By “correctness,” Criterion 4.2.9 implies that repositories must ensure that their collections are as accurate as they were when the repositories first ingested them. Applied to this study, Criterion 4.2.9 relates to the DADTS item, “The

document appears free from error,” in the sense that a TDR is required to ensure that the digitized archival documents it preserves are just as free from error as they were when the TDR first ingested them; TDRs need to ensure that no corruption has taken place which could jeopardize the accuracy of the digitized archival documents. Similarly, Criterion C16 in DIN 31644 requires digital repositories to provide an interface, “which allows users and the digital archive administration to check and maintain the integrity of the representations” (NESTOR Certification Working Group, 2013, p. 22). By “integrity,” Criterion C16 necessitates the ability of designated community members to check whether or not digitized archival documents have become corrupt while in the context of a TDR. To the extent that corruption affects digitized archival documents’ appearance, repository trustworthiness Criterion 4.2.9 and Criterion C16 relate to designated community members’ perception of digitized archival documents as trustworthy as evidenced by the importance of the DADTS item, “The document appears free from error” in this study’s findings.

In summary, findings from this study demonstrate that the connection between repository and document trustworthiness lies in associating digital object management activities with the effect of those activities on the perception of trustworthiness by the members of a designated community. In this respect, measurement of trustworthiness could grant TDRs an opportunity to assess the impact of their efforts to abide by international standards for repository trustworthiness on designated community members’ perceptions of digitized archival documents as trustworthy.

6.5 Limitations and Delimitations of the Research

This dissertation has two primary limitations and two primary delimitations. The primary limitations pertain to the participants, most of whom were older, Caucasian females, in the studies comprising the dissertation as well as issues regarding sample size. The two primary delimitations regard my intentional focus on perceived trustworthiness during the focus groups study, the item analysis study, and the validation study as well as my choice of conducting three focus groups.

6.5.1 Limitations

Limitations “refer to conditions outside the investigator’s control that affect data collection” (Mertler & Charles, 2011, p. 58). I had no control over who actually participated in my dissertation. All I could do was invite as many people as possible to participate using the most conventional and appropriate recruitment methods.

6.5.1.1 Homogeneity of Study Participants

Despite efforts to create a diverse and random sample, this dissertation had one primary limitation regarding the participants in the focus groups study, item analysis study, and the validation study; each study involved samples of participants who were mostly older, Caucasian females. Consequently, the focus groups study data may only reflect older Caucasian women’s trustworthiness perceptions and not be relevant to the trustworthiness perceptions of people with different genders,

ages, and/or racial and ethnic backgrounds. Also, the Digitized Archival Document Trustworthiness Scale (DADTS) may only be useful for measuring older, Caucasian women's document trustworthiness perceptions, as I used the responses of predominantly older, Caucasian women to develop and validate the scale.

6.5.1.2 Sample Size

During the item analysis study, sample size was an area of concern and a potential limitation of the research. While there is no consensus on the appropriate sample size for scale development projects (DeVellis, 2012; Spector, 1992), there are rules of thumb in the literature. For example, Nunnally (1978) recommends that 300 participants constitute a sufficient sample for the development of a scale. DeVellis (2012) argues that, in practical experience, scales with smaller sample sizes have been developed. DeVellis (2012) states that, "if only a single scale is to be extracted from a pool of about 20 items, fewer than 300 subjects might suffice" (p. 102). The item analysis study findings support this proposition. DADTS was successfully developed using a sample of 179 participants (after listwise case deletion on all items, including the demographic variables). However, I acknowledge that developing DADTS using a larger sample would have been preferable.

6.5.2 Delimitations

Delimitations "are the boundaries purposely put on the study, usually to narrow it for researchability" (Mertler & Charles, 2011, p. 58). I selected two

primary delimitations. The first involves focusing only on perceived trustworthiness, not actual trustworthiness. The second involves my choice of conducting three focus groups.

6.5.2.1 Actual versus Perceived Trustworthiness

While it was feasible to assess perceived trustworthiness, it was not practically feasible to assess actual trustworthiness in this dissertation. In order to generate data regarding participants' trustworthiness perceptions for digitized archival documents, I used the critical incident technique (Flanagan, 1954). Specifically, I asked participants to consider their recent experiences with digitized archival documents as they discussed trustworthiness during the focus groups study and evaluated survey items during the item analysis study. The advantage of using the critical incident technique during both studies was that it encouraged participants to focus on their real-world experiences in utilizing digitized archival documents. The disadvantage of using the critical incident technique was that, unless participants, in this case genealogists, were conducting research on the same families, each participant was probably considering different documents.

In order to verify the actual trustworthiness of the documents the participants were considering, participants would have had to identify each document, tell me which document they were thinking of, and provide me with enough information so that I could find those documents and thereafter assess their trustworthiness. Or, they would have had to find those documents and send them to me so that I could examine them and determine their trustworthiness.

During the focus groups study, the activity of identifying which specific digitized archival documents participants were considering would have taken valuable time away from the focus groups discussion. It would have distracted participants from the main objective of the study, which was to discuss their perceptions of trustworthiness.

It would have been even less feasible to compare perceived versus actual trustworthiness during the item analysis study. Since 233 genealogists participated in the item analysis study, each participant could have been considering different documents and multiple documents. Instructions during the item analysis study asked participants to consider documents they had recently viewed or downloaded and subsequently used during their evaluation of items that corresponded to their perceptions of trustworthiness. At a minimum, each participant could have been considering the same document. In which case, verifying the actual trustworthiness of one document would have been feasible. However, it seems rather unlikely that all of the participants would have been considering the same document while employing the critical incident technique, given what is known in the archival science and information science research literature regarding genealogists' information needs; they seek information regarding their own or others' families (e.g., Yakel, 2004). As such, it is much more likely the case that each participant was considering a different document or different documents, unless some participants were researching the same families. Therefore, it was not practically feasible for me to verify the actual trustworthiness of 233 or more digitized archival documents.

In sum, I selected the research delimitation of focusing on perceived trustworthiness rather than actual trustworthiness because, based on the goals of the research, it was more important to understand designated community members' trustworthiness perceptions for digitized archival documents that they actually had experience with and which actually related to their information needs. This focus pushed the scope of digitized archival documents that the study participants were considering beyond a scope that I could practically assess in terms of actual trustworthiness. Without attention toward actual trustworthiness, the study participants may have been considering the trustworthiness of documents that are not actually trustworthy.

6.5.2.2 Data Saturation

Regarding the issue of data saturation, some researchers refer to saturation as the point at which themes reoccur and no additional information regarding previously identified themes. Applying the concept to analysis of focus group data, Krueger and Casey (2009) state:

The accepted rule of thumb is to plan three or four focus groups with each type or category of individual. Once you have conducted these first three or four groups, determine if you have reached saturation. *Saturation* is a term used to describe the point where you have heard the range of ideas and aren't getting new information. If after three or four groups you were still getting new information, you would conduct more focus groups. The reason you plan three to four groups is that focus groups are analyzed across groups. The analyst looks for patterns and themes across groups. (p. 21)

Thus, if I had collected the same type of data regarding the themes during all three focus groups, I could have claimed saturation, as Krueger and Casey (2009) define the term. For example, for the identified themes of perceived authenticity, accuracy, first-hand/primary or first-hand nature, believability, and legibility, participants gave similar statements, such that the themes were reoccurring and I identified no new information. Data regarding those themes that I collected during subsequent focus groups confirmed data collected during previous focus groups. Had this been the case for all of the themes under investigation in the study, I could have claimed reaching a point of saturation for the focus groups study.

In actuality, for some of the themes under investigation, I did not collect data regarding those themes during all three focus groups, and for other themes, I did not collect data until the third focus group, suggesting collection of new information, which would suggest that my data had not reached a saturation point. For example, data collected regarding stability and validity did not consistently emerge across all three focus groups. Data pertaining to stability emerged during the first and the third focus groups, but not the second focus group. Similarly, data pertaining to validity emerged during the second and the third focus groups, but not the first focus group. In order to substantiate a claim of saturation, as Krueger and Casey (2009) define the concept, data regarding validity and stability should have emerged during all three focus groups and would have needed to be the same type of data—data confirming previously collected data. However, data saturation in the focus groups was not an essential requirement of the scale development process,

given the large number of seventy-four items that participants evaluated during later phases of the project.

In sum, there is a possibility that additional items could be used in a scale for measurement of digitized archival document trustworthiness that I did not find because I did not reach a point of data saturation during the focus groups. Future research could attempt to mine user perceptions of trustworthiness for digitized archival documents more deeply to examine whether additional items for measurement of digitized archival document trustworthiness perception could be derived from study participants' statements.

6.6 Recommendations for Future Research

This study proposes the Digitized Archival Document Trustworthiness Scale (DADTS) as a meaningful and valid measure of user document trustworthiness perception. In addition, results of the validation study provide some indication that the DADTS items resonate with other genealogists besides those who frequently utilize WADA documents. Future research could examine the propensity of the DADTS to measure the trustworthiness perceptions of different samples of designated community members beyond those who participated in the development and validation of DADTS.

If future research confirms the ability of DADTS to generalize to different designated communities, then additional research should address new questions going forward. These questions could be aimed at understanding what the correspondence between document trustworthiness perception and actual

document trustworthiness is as well as what it should be. Researchers need to empirically ask whether or not the digitized archival documents that users perceive as trustworthy actually deserve to be considered trustworthy. It is insufficient to settle for users perceiving digitized archival documents as trustworthy simply because they are housed in a Trustworthy Digital Repository (TDR), even if users are quite willing to do so. Instead, future research should seek to address whether what TDRs do (or do not do) causes users to perceive digitized archival documents as more trustworthy than those documents deserve. To start, future research could focus on whether it is possible to measure the correspondence between perceived digitized archival document trustworthiness and actual digitized archival document trustworthiness within the context of a TDR. If this measurement is possible, future research could then examine what the appropriate correspondence between perceived and actual document trustworthiness should be.

TDRs are not innocent bystanders in the trust relationship between a user and a document in the context of a digital repository. In fact, TDRs are responsible for this particular trust relationship (RLG-OCLC WGDAA, 2002). Further, findings from this study link concepts that comprise users' document trustworthiness perception to TDRs' curatorial functions and responsibilities as codified in international standards for digital repositories. Going forward, research could examine what role TDRs can play in the correspondence between perceived and actual document trustworthiness. For example, TDRs may have the potential to throw into relief the distinction between actual trustworthiness and perceived trustworthiness. Specifically, TDRs may be able to provide additional information or

take actions that can clarify for their designated community members whether digitized archival documents are actually trustworthy. Future research should be aimed at identifying what information could be used as evidence in support of actual trustworthiness of digitized archival documents, and thereafter examine the impact of providing that information to designated community members on their trustworthiness perceptions. Ultimately, future research should investigate what TDRs can do to ensure that perceived trustworthiness aligns with the intrinsic properties of the documents that TDRs preserve and make accessible.

6.7 Conclusion

Since 1996 digital curation has become established as a distinctive domain of professional practice, bounded by a suite of international standards, determined by an international network of best practices, and founded on the principle that long-term preservation depends on the development and persistence of trust. The emergence of trustworthy digital repositories as viable storehouses of data, information, and knowledge is the most compelling evidence to date that digital curation practices are capable of affecting the long-term preservation of digital information. Digital repositories are special domains of managed information. By intention and design, repositories protect the authenticity and accuracy of digital documents, and in doing so, establish and maintain their trustworthiness. The results of this study demonstrate that it is possible and valuable to measure with statistical soundness and conceptual nuance how a designated community of users perceives the trustworthiness of digital information. In doing so, the Digitized

Archival Document Trustworthiness Scale (DADTS) lays the groundwork for future investigations of how trust, beyond serving as a symbolic brand, truly functions as an operational component of digital repositories.

6.8 References

- Association for Information and Image Management. (1992). *Performance guideline for the legal acceptance of records produced by information technology systems*. Silver Spring, MD: The Association for Information and Image Management.
- Bartholomew, D. J., Steele, F., Moustaki, I., & Galbraith, J. I. (2008). *Analysis of multivariate social science data*. Boca Raton: CRC Press.
- Baudoin, P. (2008). The principle of digital preservation. *The Serials Librarian*, 55(4), 556-559.
- Bearman, D., & Trant, J. (1998). Authenticity of digital resources: Towards a statement of requirements in the research process. *D-Lib Magazine*, 4(6)
Retrieved from <http://dlib.org/dlib/june98/06bearman.html>
- Consultative Committee for Space Data Systems. (2002). *Reference model for an open archival information system (OAIS) recommendation for space data system standards ; blue book* (Blue Book. Washington, DC: CCSDS Secretariat. doi: <http://public.ccsds.org/publications/archive/650x0b1.pdf>
- Conway, P. (2011). Archival quality and long-term preservation: A research framework for validating the usefulness of digital surrogates. *Archival Science*, 11(3-4), 293-309.
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches*. 4th ed. Thousand Oaks, Calif.: Sage Publications.
- DeVellis, R. F. (2012). *Scale development: Theory and applications*. Thousand Oaks, Calif.: SAGE.

- Duff, W., Craig, B., & Cherry, J. (2004). Finding and using archival resources: A cross-canada survey of historians studying canadian history. *Archivaria*, 58, 51-80.
- Duff, W. M. (1998). Harnessing the power of warrant. *American Archivist*, 61(1), 88-105.
- Duranti, L., & Preston, R. (Eds.). (2008). *International research on permanent authentic records in electronic systems (InterPARES) 2: Experiential, interactive and dynamic records*. Padova, Italy: Associazione Nazionale Archivistica Italiana.
- Duranti, L., Suderman, J., & Todd, M. (2008) Part seven-structuring the relationship between records creators and preservers: Policy cross-domain task force report. In L. Duranti, & R. Preston (Eds.), *International research on permanent authentic records in electronic systems (InterPARES) 2: Experiential, interactive and dynamic records* (). Padova, Italy: Associazione Nazionale Archivistica Italiana. Retrieved from http://www.interpares.org/display_file.cfm?doc=ip2_book_part_7_policy_task_force.pdf
- Duranti, L. (1995). Reliability and authenticity: The concepts and their implications. *Archivaria*, 39, 5-10.
- Fogg, B. J., Marshall, J., Laraki, O., Osipovich, A., Varma, C., Fang, N., et al. (2001). What makes web sites credible?: A report on a large quantitative study. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Seattle, Washington, United States. pp. 61-68.

- Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin*, 51(4), 327-358.
- Flanagin, A. J., & Metzger, M. J. (2000). Perceptions of internet information credibility. *Journalism and Mass Communication Quarterly*, 77(3), 515-540.
- (2007). The role of site features, user attributes, and information verification behaviors on the perceived credibility of web-based information. *New Media & Society*, 9(2), 319-342.
- Fogg, B. J. (2003). *Persuasive technology: Using computers to change what we think and do*. Amsterdam; Boston: Morgan Kaufmann Publishers.
- Gilliland-Swetland, A., & Eppard, P. B. (2000). Preserving the authenticity of contingent digital objects: The InterPARES project. *D-Lib Magazine*, 6(7/8).
- International Organization for Standardization. (2012). *Space data and information transfer systems : Audit and certification of trustworthy digital repositories* [Systèmes de transfert des informations et données spatiales : audit et certification des référentiels numériques de confiance] (1st ed.). Geneva, Switzerland: International Organization for Standardization.
- Jacoby, W. G. (1991). *Data theory and dimensional analysis*. Newbury Park, Calif.: Sage Publications.
- Kelton, K., Fleischmann, K. R., & Wallace, W. A. (2008). Trust in digital information. *Journal of the American Society for Information Science and Technology*, 59(3), 363-374.

- Kim, Y. (2009). Validation of psychometric research instruments: The case of information science. *Journal of the American Society for Information Science and Technology*, 60(6), 1178-1191.
- Kline, R. B. (2013). Exploratory and confirmatory factor analysis. In Y. Petscher, C. and the social sciences (pp. 171-207). New York: Routledge.
- Knight, S. (2008). User perceptions of information quality in world wide web information retrieval behaviour. (Doctoral Dissertation, Edith Cowan University), 1-362.
- Krueger, R. A., & Casey, M. A. (2009). *Focus groups : A practical guide for applied research* (4th ed.). Los Angeles: SAGE.
- Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). AIMQ: A methodology for information quality assessment. *Information & Management*, 40(2), 133-146.
- Liu, Z. (2004). Perceptions of credibility of scholarly information on the web. *Information Processing & Management*, 40(6), 1027-1038.
- MacNeil, H. (2000). *Trusting records: Legal, historical and diplomatic perspectives*. Dordrecht; London: Kluwer Academic.
- Mertler, C. A., & Charles, C. M. (2011). *Introduction to educational research* (7th ed.). Boston: Pearson/Allyn & Bacon.
- NESTOR Working Group on Trusted Repositories Certification. (2013). *Explanatory notes on the nestor seal for trustworthy digital archives*. Germany: Deutsche Nationalbibliothek.
- Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.

- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory*. New York: McGraw-Hill.
- OCLC/RLG Working Group on Digital Archive Attributes. (2002). *Trusted digital repositories: Attributes and responsibilities*. Mountain View, CA: RLG.
- Retrieved from <http://www.oclc.org/research/activities/past/rlg/trustedrep/repositories.pdf>
- Rieh, S. Y., & Danielson, D. R. (2007). Credibility: A multidisciplinary framework. *Annual Review of Information Science and Technology*, 41(1), 307-364.
- Rieh, S. Y. (2010). Credibility and cognitive authority of information. *Encyclopedia of Library and Information Sciences, Third Edition*, 1337-1344.
- Smith, A. (2000a). Authenticity in Perspective. *Authenticity in a digital environment* (pp. vi-viii). Washington D.C.: Council on Library and Information Resources.
- Retrieved from <http://clir.org/pubs/reports/pub92/pub92.pdf>
- . (2000b). Introduction. *Authenticity in a digital environment* (pp. 69-75). Washington D.C.: Council on Library and Information Resources. Retrieved from <http://clir.org/pubs/reports/pub92/pub92.pdf>
- Spector, P. E. (1992). *Summated rating scale construction: An introduction*. Newbury Park, Calif.: SAGE Publications.
- Trouillot, M. (1995). *Silencing the past: Power and the production of history*. Boston, Mass.: Beacon Press.
- Yakel, E. (2004). Seeking information, seeking connections, seeking meaning: Genealogists and family historians. *Information Research*, 10(1).

Yakel, E., Faniel, I., Kriesberg, A., & Yoon, A. (2013). Trust in digital repositories.

International Journal of Digital Curation, 8(1), 143-156.

Yoon, A. (2014). End users' trust in data repositories: Definition and influences on

trust development. *Archival Science, 14(1), 17-34.*

APPENDICES

Appendix A

Catalogue of Changes to Item Pool as a Result of Cognitive Interviews

- Table A.1 summarizes the revisions to the items as a result of the cognitive interview findings.
- Table A.2 provides a list of the nine items that I deleted based on the cognitive interviewees' feedback.
- Table A.3 provides a list of the eleven items that I added based on the cognitive interviewees' feedback.
- Table A.4 lists the twenty items that I revised as a result of the cognitive interviewees' feedback.
- Table A.5 lists the forty-two items that I retained as is.

Original Items Cognitive Interviewees Evaluated	71
Items Deleted	9
Items Added	11
Items Edited	20
Items Kept	42
Final Total	74

Appendix A Table 1. Summary of Revisions to Items as a Result of Cognitive interviewees' Feedback

Items

1. The document is dependable.
 2. The document is real.
 3. The document is current.
 4. The document looks like it comes from where I think it comes from.
 5. The document is the closest to the original “action” that happened.
 6. The document is accurate, but the actual information is erroneous.
 7. The document is valid.
 8. The document represents an account of the event afterwards.
 9. The document is reliable.
-

Appendix A Table 2. Items Deleted Based on Cognitive Interviewees’ Feedback.

Items

1. The document provides an impartial view of the topic.
 2. The document is susceptible to alteration.
 3. The document has been amended.
 4. The document was created using responsible and accepted practices.
 5. The document appears blurry.
 6. The document includes appropriate citation of sources.
 7. The document is persistent in its contents.
 8. Sound methods were used to create the document.
 9. The document accurately reflects what happened.
 10. The document is consistently available.
 - 11.** The document lists who provided the information.
-

Appendix A Table 3. Items Added Based on Cognitive Interviewees’ Feedback.

Original Items	Revised items
1. The document is altered.	1. The document has been altered.
2. The document is altered with the intention to deceive.	2. The document has been altered with the intention to deceive.
3. The document is properly digitized.	3. The document is clearly digitized.
4. The document lists a maiden name instead of a married name.	4. The document lists a maiden name for a person who is married.
5. The document is the actual scanned image.	5. The digitized document accurately presents the original document.
6. The document is what it says it is.	6. The document is what it claims to be.
7. The document is the same every time I download or click on it.	7. The document is the same every time I download or view it.
8. The document contains bias.	8. The document is objective.
9. The document has been verified by someone else.	9. The document's authenticity has been verified by someone else.
10. The document is complete.	10. The document appears complete.
11. The document has documentation of where it came from.	11. The document includes documentation of where it came from.
12. I know who provided the information.	12. I personally know who provided the information in the document.
13. I found information elsewhere that conflicts with the document.	13. I found information elsewhere that contradicts the document.
14. The digitized document looks like what the original document should look like.	14. The digitized document resembles what the original document should look like.
15. The document looks like what an official document should look like.	15. The document resembles what an official document should look like.
16. The document matches information I know about.	16. The document matches other information I know about.
17. The document was written by the minister who filled out the form for the marriage.	17. The document was completed by the minister who filled out the form for the marriage.
18. The document is second-hand.	18. The document presents a second-hand account of what happened.
19. The document is first-hand.	19. The document presents a first-hand account of what happened.
20. The document is accurate.	20. The document appears free from error.

Appendix A Table 4. Items Revised Based on Cognitive Interviewees' Feedback.

Items
1. The document is authoritative.
2. The document is hearsay.
3. Names are misspelled.
4. The document is from the time period it claims to be.
5. The information is correct.
6. The document lists an incorrect marital status for a person.
7. The document lists an incorrect age for a person.
8. The document is readable.
9. The document contains a seal.
10. The document lists fake names.
11. The document matches research I have found using other sources.
12. The handwriting looks to be of the correct time period.
13. The document is a primary source.
14. The document is censored or blacked out.
15. The document was written at the time of the event.
16. I have seen the original physical document that was used to create the digitized document.
17. Parts of the document are crossed out.
18. The document is typed.
19. The document is believable.
20. The document is factual.
21. The document is credible.
22. The document is authentic.
23. The document has all the appropriate fields, even though all the fields are not completely filled out.
24. The document is a secondary source.
25. The document lists the wrong place of residence for a person.
26. The document is notarized.
27. The document is fake.
28. The digitized document is an actual picture of the original physical document.
29. The document is a fuzzy reproduction.
30. The document contains false information.
31. The document is legible.
32. The document lists an adopted parent as a birth parent.
33. The document lists the wrong place of birth for a person.
34. I can see that some of the spelling has been changed.
35. The document is official.
36. The person the document is about was alive during the time the document was created.
37. The document is handwritten.
38. The document was altered by the witnesses.
39. A person is listed as being in two places at the same time.
40. The document is mistakenly identified.
41. The document has the proper form.
42. I can see that at least one of the dates have been changed.

Appendix A Table 5. Items Kept As Is After Cognitive Interviews.

Appendix B

Sample Recruitment Email for Focus Groups

Subject line: Free Lunch - Tuesday June 11 - Olympia

Greetings, All!

For those of you who live near Olympia and might be interested, we are hosting a free luncheon at the Washington State Archives, located at 1129 Washington Street SE, in Olympia, WA 98501 on Tuesday, June 11 at noon.

All we ask from you is your input regarding the Digital Archives. We sincerely hope that you will consider attending, as this event will directly help the Archives ensure that we continue to meet the needs of our customers.

If you are interested in attending, please contact Devan Ray Donaldson, at devand@umich.edu.

Thanks so much for your consideration!

Best regards,
Amber

Amber Raney

Historical Records Project Coordinator

Washington State Archives
1129 Washington St. SE
Mailstop 40238
Olympia, WA 98504-0238

Phone: (360) 586-2665

www.sos.wa.gov/archives

Appendix C

Sample Web Survey Form for Recruiting Focus Group Participants

Washington State Archives - Digital Archives Focus Group Recruitment Form

To help me place you in the right focus group, please answer the following questions:

Which location are you closest to: (Please select one choice)

- (Olympia) Washington State Archives – 1129 Washington Street SE, Olympia, WA, 98501
- (Seattle) Puget Sound Regional Archives - Pritchard-Fleming Building, 3000 Landerholm Circle SE, Bellevue, WA 98007
- (Cheney) Washington State Archives – Digital Archives, 960 Washington Street, Cheney WA 99004

What is your primary reason for visiting the Washington State Digital Archives web site?

- Business
- Genealogy
- Historical research
- Local government
- Other: please specify _____
- Title searches

Please provide your contact information below so that I can contact you for scheduling the focus group:

First name
Last name
Email address
Phone number

You will receive specific details from me about where and when the focus group will take place within 48 hours of your completion of this form. Thanks again!

Appendix D

Focus Groups Pre-survey

Study Participants' Internet Usage

Question 1: On average, how frequently do you use the Internet?

- Daily☑
- Weekly
- ☑• Monthly
- 1 or 2 times

☑Question 2: In general, I trust information I find on the Internet.

- Strongly Agree
- ☑• Agree
- ☑• Disagree
- ☑• Strongly Disagree

Study Participants' WADA Usage

Question 3: In the last year, how frequently have you used the Washington State Digital Archives?

- ☑• Daily
- ☑• Weekly
- Monthly
- ☑• 1 or 2 times☑

Question 4: What is your primary reason for visiting the Washington State Digital Archives?

- Business
- Genealogy
- Historical Research
- Local Government
- Other

Question 5: I usually find the documents I'm looking for when using the Washington State Digital Archives?

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

Question 6: How strongly do you trust the documents you find when using the Washington State Digital Archives?

- Trust very strongly
- Somewhat strongly
- Not very strongly
- Not at all

Question 7: How satisfied are you with the way the Washington State Digital Archives displays documents?

- Very satisfied
- Satisfied
- Dissatisfied
- Very Dissatisfied

☒ *Background Information* ☒

Question 8: What is your age range?

• 19 or under

• 20 to 29

☒ • 30 to 39

• 40 to 49

☒ • 50 to 59

• 60 to 69 ☒

• 70 or older ☒

Question 9: What is your gender?

• Female

• Male

Appendix E

Focus Groups Guide

1. Discuss the nature of the documents you use when using WADA and your purpose(s) for using them.
2. Discuss your perceptions of trustworthiness for the documents you find using WADA. Can you give an example of a document you found that you thought was trustworthy, or one that looks more trustworthy than typical documents you find in WADA?
3. How would you describe a document you found in WADA that you think is trustworthy? What adjectives would you use?
4. Under what circumstances would you question the trustworthiness of a document you encountered while using WADA?

Appendix F

Codebook for Focus Groups Data

- I. Identified themes
 - a. Authenticity - Authenticity refers to whether or not a document is what it claims to be (Duranti, 1995; Duranti, 2005; Duranti et al., 2008; Levy, 2000; MacNeil, 2000).
 - b. Primary or first-hand evidence - Primary or first-hand evidence refers to primary documents as documents that were written during the time period of the events that they were about as opposed to documents that serve as accounts of what happened that were written at a later time. Primary or first-hand also refers to having been written by the actual person(s) the documents were meant to represent.
 - c. Legibility - Refers to the readability of digitized documents.
 - d. Form - Refers to the perceived structure and presentation of digitized documents.
- II. Kelton-based themes
 - a. Accuracy - Kelton et al. (2008) define accuracy as “the extent to which information is free from error” (p. 370).

- b. Believability - Kelton et al. (2008) define believability as “the extent to which the information appears to be plausible” (p. 370). Kelton also treats believability as synonymous with credibility, which is consistent with other researchers who have conducted extensive research on credibility (Tseng & Fogg, 1999; Fogg, 2003).
- c. Coverage (i.e., completeness) - Kelton et al. (2008) define coverage as “the completeness of the information” (p. 370).
- d. Currency - Kelton et al. (2008) define currency as “the degree to which the information is up-to-date rather than obsolete” (p. 370).
- e. Objectivity - Kelton et al. (2008) define objectivity as “balance of content” (p. 370).
- f. Stability - Kelton et al. (2008) define stability as implying that “the information is persistent, in both its presence and its contents” (p. 370).
- g. Validity - Kelton et al. (2008) define validity as “the use of responsible and accepted practices” such as “the soundness of the methods used, the inclusion of verifiable data, and the appropriate citation of sources” (p. 370).

Appendix G

Final Survey Used During Step 4 of Scale Development

Revised Washington State Archives - Digital Archives Item Pool Survey - Option #1

Thank you for your interest in taking this survey! This survey is part of a research project at the University of Michigan. Your responses will help to greatly improve the experience for users of the Washington State Digital Archives. The following questions concern your perceptions of digitized documents (JPEGs) that you have recently viewed or downloaded from the Washington State Digital Archives. Because these questions ask about your perceptions, there are no right or wrong answers. Please try to provide an answer for every question. The survey will take approximately 15 minutes to complete. If you decide to provide your contact information at the end of this survey, it will only be used by the researcher responsible for this survey, University of Michigan Ph.D. student Devan Ray Donaldson, so that he can contact you to discuss some of your responses. Otherwise, your responses will be kept anonymous and your contact information will be kept confidential. Would you like to continue with this survey?

- YES
- NO THANKS

If NO THANKS! Is Selected, Then Skip To End of Survey

In the following screens, each statement represents a particular circumstance you might encounter while using a document. Please indicate whether the circumstance described in each statement would cause you to perceive a document as untrustworthy or trustworthy. Remember, these statements refer to digitized documents (JPEGs) that you have recently viewed or downloaded from the Washington State Digital Archives and subsequently used.

Instructions: Trustworthiness is a user judgment, based primarily on subjective factors, regarding a document found in a digital archive. Please indicate whether the circumstance described in each statement below would cause you to perceive a document as untrustworthy or trustworthy. Select the button below the number that most closely represents your perception for each statement. Remember, these statements refer to digitized documents (JPEGs) that you have recently viewed or downloaded from the Washington State Digital Archives and subsequently used.

	Very untrustworthy -3	Untrustworthy -2	Slightly untrustworthy -1	Neither untrustworthy or trustworthy 0	Slightly trustworthy 1	Trustworthy 2	Very trustworthy 3	Not applicable 99
The document is authoritative.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is hearsay.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Names are misspelled.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is from the time period it claims to be.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document has been amended.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document has	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<p>been altered with the intention to deceive.</p> <p>The information is correct.</p> <p>The document lists an incorrect marital status for a person.</p> <p>The document lists an incorrect age for a person.</p> <p>The document is readable.</p>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
---	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

Instructions: Trustworthiness is a user judgment, based primarily on subjective factors, regarding a document found in a digital archive. Please indicate whether the circumstance described in each statement below would cause you to perceive a document as untrustworthy or trustworthy. Select the button below the number that most closely represents your perception for each statement. Remember, these statements refer to digitized documents (JPEGs) that you have recently viewed or downloaded from the Washington State Digital Archives and subsequently used.

	Very untrustworthy - 3	Untrustworthy - 2	Slightly untrustworthy - 1	Neither untrustworthy or trustworthy 0	Slightly trustworthy 1	Trustworthy 2	Very trustworthy 3	Not applicable 99
The document contains a seal.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document lists fake names.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document matches research I have found using other sources.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is clearly digitized.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The handw	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<p>riting looks to be of the correct time period.</p> <p>The document has been altered .</p> <p>The document is a primary source.</p> <p>The document appears complete.</p> <p>The document is censored or blacked out.</p> <p>The document's authenticity has been verified by someone else.</p>	○	○	○	○	○	○	○	○
---	---	---	---	---	---	---	---	---

Instructions: Trustworthiness is a user judgment, based primarily on subjective factors, regarding a document found in a digital archive. Please indicate whether the circumstance described in each statement below would cause you to perceive a document as untrustworthy or trustworthy. Select the button below the number that most closely represents your perception for each statement. Remember, these statements refer to digitized documents (JPEGs) that you have recently viewed or downloaded from the Washington State Digital Archives and subsequently used.

	Very untrustworthy - 3	Untrustworthy - 2	Slightly untrustworthy - 1	Neither untrustworthy or trustworthy 0	Slightly trustworthy 1	Trustworthy 2	Very trustworthy 3	Not applicable 99
The document was written at the time of the event.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I have seen the original physical document that was used to create the digitized document.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Parts of the document are	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

crossed out.								
The document lists a maiden name for a person who is married.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is typed.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The digitized document accurately presents the original document.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is what it claims to be.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is believable.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

provides an impartial view of the topic. The document is susceptible to alteration.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
--	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

Instructions: Trustworthiness is a user judgment, based primarily on subjective factors, regarding a document found in a digital archive. Please indicate whether the circumstance described in each statement below would cause you to perceive a document as untrustworthy or trustworthy. Select the button below the number that most closely represents your perception for each statement. Remember, these statements refer to digitized documents (JPEGs) that you have recently viewed or downloaded from the Washington State Digital Archives and subsequently used.

	Very untrustworthy - 3	Untrustworthy - 2	Slightly untrustworthy - 1	Neither untrustworthy or trustworthy 0	Slightly trustworthy 1	Trustworthy 2	Very trustworthy 3	Not applicable 99
The document is objective.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document appears free from error.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is factual.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is credible.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is authentic.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document has all the appropriate fields, even	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<p>though all the fields are not completely filled out.</p> <p>The document is a secondary source.</p> <p>The document includes documentation of where it came from.</p> <p>The document was created using responsible and accepted practices.</p> <p>The document appears blurry.</p>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
--	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

Instructions: Trustworthiness is a user judgment, based primarily on subjective factors, regarding a document found in a digital archive. Please indicate whether the circumstance described in each statement below would cause you to perceive a document as untrustworthy or trustworthy. Select the button below the number that most closely represents your perception for each statement. Remember, these statements refer to digitized documents (JPEGs) that you have recently viewed or downloaded from the Washington State Digital Archives and subsequently used.

	Very untrustworthy - 3	Untrustworthy - 2	Slightly untrustworthy - 1	Neither untrustworthy or trustworthy 0	Slightly trustworthy 1	Trustworthy 2	Very trustworthy 3	Not applicable 99
The document presents a first-hand account of what happened.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document lists the wrong place of residence for a person.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is notarized.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is fake.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I found information elsewhere that	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<p>contradicts the document.</p> <p>The digitized document is an actual picture of the original physical document.</p> <p>The document is a fuzzy reproduction.</p> <p>The document resembles what an official document should look like.</p> <p>The document includes appropriate citation of sources.</p> <p>The document is persiste</p>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
---	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

nt in its content s.								
----------------------------	--	--	--	--	--	--	--	--

Instructions: Trustworthiness is a user judgment, based primarily on subjective factors, regarding a document found in a digital archive. Please indicate whether the circumstance described in each statement below would cause you to perceive a document as untrustworthy or trustworthy. Select the button below the number that most closely represents your perception for each statement. Remember, these statements refer to digitized documents (JPEGs) that you have recently viewed or downloaded from the Washington State Digital Archives and subsequently used.

	Very untrustworthy - 3	Untrustworthy - 2	Slightly untrustworthy - 1	Neither untrustworthy or trustworthy 0	Slightly trustworthy 1	Trustworthy 2	Very trustworthy 3	Not applicable 99
The document contains false information.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is legible.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The digitized document resembles what the original document should look like.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document matches other inform	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<p>ation I know about.</p> <p>The document lists an adopted parent as a birth parent.</p> <p>The document lists the wrong place of birth for a person.</p> <p>The document was completed by the minister who filled out the form for the marriage.</p> <p>The document lists who provided the information.</p> <p>I can</p>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
---	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

see that some of the spelling has been changed. The document includes verifiable data.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
--	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

Instructions: Trustworthiness is a user judgment, based primarily on subjective factors, regarding a document found in a digital archive. Please indicate whether the circumstance described in each statement below would cause you to perceive a document as untrustworthy or trustworthy. Select the button below the number that most closely represents your perception for each statement. Remember, these statements refer to digitized documents (JPEGs) that you have recently viewed or downloaded from the Washington State Digital Archives and subsequently used.

	Very untrustworthy - 3	Untrustworthy - 2	Slightly untrustworthy - 1	Neither untrustworthy or trustworthy 0	Slightly trustworthy 1	Trustworthy 2	Very trustworthy 3	Not applicable 99
The document presents a second-hand account of what happened.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Sound methods were used to create the document.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is official.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The person the document is about was alive during the	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

time the document was created								
The document is handwritten.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document was altered by the witnesses.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
A person is listed as being in two places at the same time.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is mistakenly identified.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document accurately reflects what happened.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

consist ently availabl e.								
------------------------------------	--	--	--	--	--	--	--	--

Instructions: Trustworthiness is a user judgment, based primarily on subjective factors, regarding a document found in a digital archive. Please indicate whether the circumstance described in each statement below would cause you to perceive a document as untrustworthy or trustworthy. Select the button below the number that most closely represents your perception for each statement. Remember, these statements refer to digitized documents (JPEGs) that you have recently viewed or downloaded from the Washington State Digital Archives and subsequently used.

	Very untrustworthy - 3	Untrustworthy - 2	Slightly untrustworthy - 1	Neither untrustworthy or trustworthy 0	Slightly trustworthy 1	Trustworthy 2	Very trustworthy 3	Not applicable 99
The document is the same every time I download or view it.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I personally know who provided the information in the document.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document has the proper form.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
I can see that at least one of	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

the dates have been change d.								
--	--	--	--	--	--	--	--	--

What type of document do you access most frequently when using the Washington State Digital Archives?

- Birth records
- Death records
- Marriage records
- Military records
- Census records
- Land records
- Other: please specify _____

What is your primary reason for using documents that you typically find in the Washington State Digital Archives?

- Business
- Genealogy
- Historical research
- Local government
- Other: please specify _____
- Title searches

In the last year, how frequently have you used documents in the Washington State Digital Archives?

- Not at all
- A few times a year
- Monthly
- Weekly
- Daily

How much time do you spend using documents on your typical visit to the Washington State Digital Archives?

- 0-30 minutes
- 31-59 minutes
- 1-2 hours
- More than 2 hours

What is your age range?

- 19 or under
- 20 to 29
- 30 to 39
- 40 to 49
- 50 to 59
- 60 to 69
- 70 or older

What is your gender?

- Male
- Female

How satisfied are you with the way the Washington State Digital Archives displays documents?

- Very Dissatisfied
- Dissatisfied
- Satisfied
- Very Satisfied

Are you willing to be contacted to discuss some of the answers you provided in this survey?

- Yes
- No

If YES Is Selected, Then Skip To Please provide your contact informati...If NO Is Selected, Then Skip To End of Survey

Please provide your contact information below:

First name

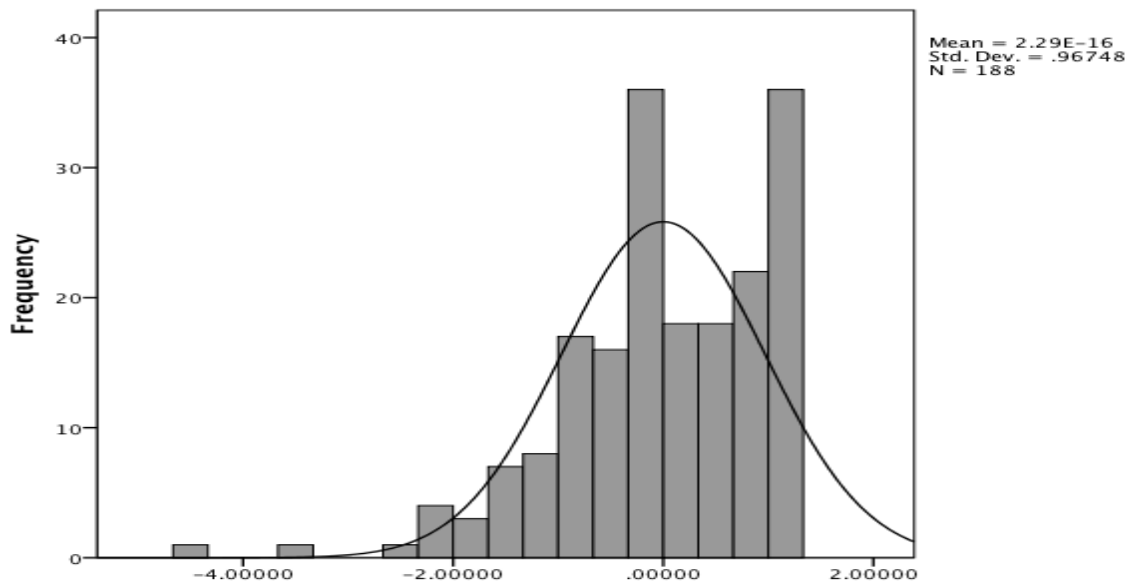
Last name

Email address

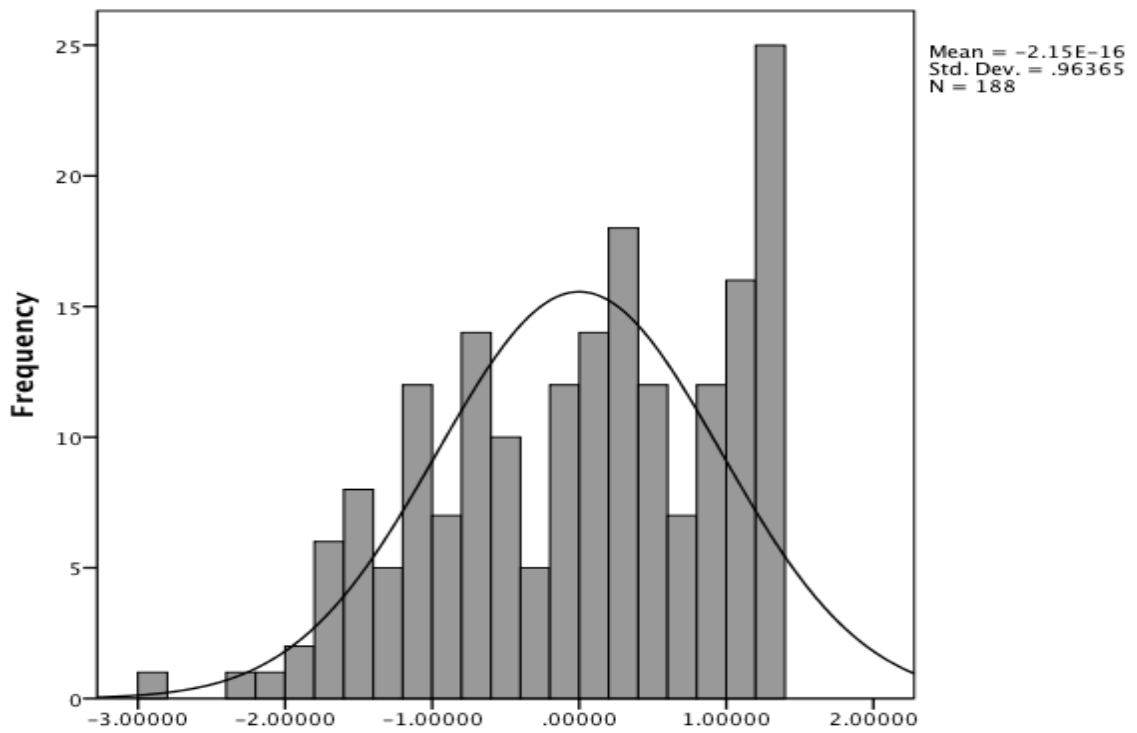
Appendix H

Outlier Detection Process

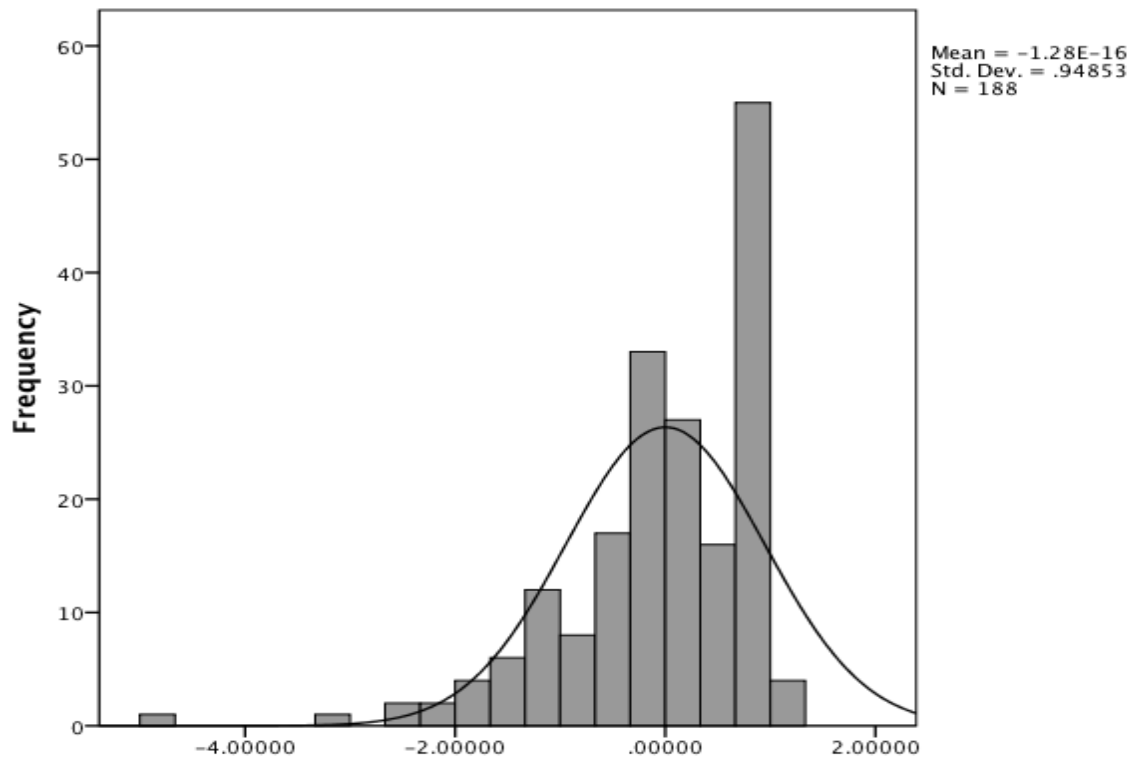
To identify the existence of outliers in the dataset, the following procedures were followed. First, computed scores for each participant on each factor were calculated using the REG(ALL) command in SPSS 22.0. Second, histograms were produced to display the distribution of participants' computed scores on each factor (see Figures H.1, H.2, and H.3). The histograms showed that there was a participant falling below the normal distribution of computed scores for each factor: below -4 for the first factor, below -2 for the second factor, and below -4 for the third factor. Afterwards, the dataset was sorted (ascending) by each of the factor scores to identify which participants fell below the normal distribution for each factor. Results from sorting the dataset by the computed scores for each factor revealed that one participant fell below the normal distribution of computed scores on each factor as shown in Figures 1, 2, and 3: -4.38262 on the first factor, -2.85236 on the second factor, and -4.98307 on the third factor. Consequently, this participant was removed from all analyses, and statistical analyses were rerun without using this outlier's data.



Appendix H Figure 1. Histogram of Computed Scores for Factor 1



Appendix H Figure 2. Histogram of Computed Scores for Factor 2.



Appendix H Figure 3. Histogram of Computed Scores for Factor 3.

Appendix I

List of Items Not Included in Exploratory Factor Analysis

Items

1. The document is hearsay.
2. Names are misspelled.
3. The document has been amended.
4. The document has been altered with the intention to deceive.
5. The document lists an incorrect marital status for a person.
6. The document lists an incorrect age for a person.
7. The document contains a seal.
8. The document lists fake names.
9. The document is clearly digitized.
10. The handwriting looks to be of the correct time period.
11. The document has been altered.
12. The document is censored or blacked out.
13. The document's authenticity has been verified by someone else.
14. I have seen the original physical document that was used to create the digitized document.
15. Parts of the document are crossed out.
16. The document lists a maiden name for a person who is married.
17. The document is typed.
18. The digitized document accurately presents the original document.
19. The document provides an impartial view of the topic.
20. The document is susceptible to alteration.
21. The document is objective.
22. The document is a secondary source.
23. The document appears blurry.
24. The document presents a first-hand account of what happened.
25. The document lists the wrong place of residence for a person.
26. The document is notarized.
27. The document is fake.
28. I found information elsewhere that contradicts the document.
29. The document is a fuzzy reproduction.
30. The document includes appropriate citation of sources.
31. The document is persistent in its contents.
32. The document contains false information.
33. The document lists an adopted parent as a birth parent.
34. The document lists the wrong place of birth for a person.
35. The document was completed by the minister who filled out the form for the marriage.
36. The document lists who provided the information.
37. I can see that some of the spelling has been changed.
38. The document presents a second-hand account of what happened.

39. Sound methods were used to create the document.
 40. The person the document is about was alive during the time the document was created.
 41. The document is handwritten.
 42. The document was altered by the witnesses.
 43. A person is listed as being in two places at the same time.
 44. The document is mistakenly identified.
 45. The document is consistently available.
 46. I personally know who provided the information in the document.
 47. The document has the proper form.
 - 48.** I can see that at least one of the dates have been changed.
-

Appendix I Table 1. List of Items Not Included in Exploratory Factor Analysis

Appendix J

SPSS Syntax for Item Analysis and Factor Analysis

* Syntax for the EFA of 179 valid cases, specifying 3 factors. LISTWISE. Rerunning EFA after getting rid of the outlier.

```
FACTOR
/VARIABLES Q2_1 Q2_4 Q2_7 Q2_10 Q4_3 Q4_7 Q4_8 Q6_1 Q6_7 Q6_8 Q7_3 Q7_4
Q7_5 Q7_6 Q7_7 Q7_9 Q7_11 Q8_7 Q8_9 Q9_3 Q9_4 Q9_5 Q9_11 Q10_4 Q10_10
Q21_2
/MISSING LISTWISE
/ANALYSIS Q2_1 Q2_4 Q2_7 Q2_10 Q4_3 Q4_7 Q4_8 Q6_1 Q6_7 Q6_8 Q7_3 Q7_4
Q7_5 Q7_6 Q7_7 Q7_9 Q7_11 Q8_7 Q8_9 Q9_3 Q9_4 Q9_5 Q9_11 Q10_4 Q10_10
Q21_2
/PRINT INITIAL DET KMO REPR EXTRACTION ROTATION FSCORE UNIVARIATE
/PLOT EIGEN ROTATION
/CRITERIA FACTORS(3) ITERATE(500)
/EXTRACTION PAF
/ROTATION PROMAX
/METHOD=CORRELATION
```

Appendix K

SPSS Syntax for Parallel Analysis

* Descriptive statistics on all variables used for Exploratory Factor Analysis (EFA).

```
DESCRIPTIVES VARIABLES=Q2_1 Q2_4 Q2_7 Q2_10 Q4_3 Q4_7 Q4_8 Q6_1 Q6_7  
Q6_8 Q7_3 Q7_4 Q7_5 Q7_6 Q7_7 Q7_9 Q7_11 Q8_7 Q8_9 Q9_3 Q9_4 Q9_5 Q9_11  
Q10_4 Q10_10 Q21_2  
/STATISTICS=MEAN STDDEV MIN MAX.
```

* Parallel Analysis Program For Raw Data and Data Permutations.

* To run this program you need to first specify the data for analysis and then RUN, all at once, the commands from the MATRIX statement to the END MATRIX statement.

* This program conducts parallel analyses on data files in which the rows of the data matrix are cases/individuals and the columns are variables; Data are read/entered into the program using the GET command (see the GET command below); The GET command reads an SPSS data file, which can be either the current, active SPSS data file or a previously saved data file; A valid filename/location must be specified on the GET command; A subset of variables for the analyses can be specified by using the "/ VAR =" subcommand with the GET statement; There can be no missing values.

* You must also specify:

- the # of parallel data sets for the analyses;
- the desired percentile of the distribution and random data eigenvalues;
- whether principal components analyses or principal axis/common factor analysis are to be conducted, and
- whether normally distributed random data generation or permutations of the raw data set are to be used in the parallel analyses.

* Permutations of the raw data set can be time consuming; Each parallel data set is based on column-wise random shufflings of the values in the raw data matrix using Castellán's (1992, BRMIC, 24, 72-77) algorithm; The distributions of the original raw variables are exactly preserved in the shuffled versions used in the parallel analyses; Permutations of the raw data set are thus highly accurate and most relevant, especially in cases where the raw data are not normally distributed or when they do not meet the assumption of multivariate normality (see Longman & Holden, 1992, BRMIC, 24, 493, for a Fortran version); If you would like to go this route, it is perhaps best to (1) first run a normally distributed random data generation parallel analysis to familiarize yourself with the program and to get a ballpark reference point for the number of factors/components; (2) then run a permutations of the raw data parallel analysis using a small number of datasets (e.g., 100), just to see how long the program takes to run; then (3) run a permutations of the raw data parallel analysis using the number of parallel data sets that you would like use for your final analyses; 1000 datasets are usually sufficient, although more datasets should be used if there are close calls.

```
set mxloops=9000 printback=off width=80 seed = 1953125.
matrix.
```

* Enter the name/location of the data file for analyses after "FILE ="; If you specify "FILE = *", then the program will read the current, active SPSS data file; Alternatively, enter the name/location of a previously saved SPSS data file instead of "*"; you can use the "/ VAR =" subcommand after "/ missing=omit" subcommand to select variables for the analyses.

```
GET raw / FILE = * / missing=omit / VAR = Q2_1 Q2_4 Q2_7 Q2_10 Q4_3 Q4_7 Q4_8
Q6_1 Q6_7 Q6_8 Q7_3 Q7_4 Q7_5 Q7_6 Q7_7 Q7_9 Q7_11 Q8_7 Q8_9 Q9_3 Q9_4 Q9_5
Q9_11 Q10_4 Q10_10 Q21_2.
```

* Enter the desired number of parallel data sets here. This is akin to a sample size of 1000.
compute ndatasets = 1000.

* Enter the desired percentile here. This is the equivalent of statistical significance testing, where you're specifying your alpha level. We will get percentiles associated with the eigenvalues that have been estimated on the 1000 datasets that are going to be generated from the parallel analysis. This program by default provides the 50th percentile, the means,

or medians in a normally distributed dataset, the means and the medians will correspond exactly. So you're going to get the mean eigenvalue from the parallel analysis (i.e., the 50th percentile), and you're going to get the 95th percentile.
compute percent = 95.

* Enter either
1 for principal components analysis, or
2 for principal axis/common factor analysis.
compute kind = 2 .

* Enter either
1 for normally distributed random data generation parallel analysis, or
2 for permutations of the raw data set.
compute randtype = 2.

***** End of user specifications. *****

compute ncases = nrow(raw).
compute nvars = ncol(raw).

* principal components analysis & random normal data generation.
do if (kind = 1 and randtype = 1).
compute nm1 = 1 / (ncases-1).
compute vcv = nm1 * (sscp(raw) - ((t(csum(raw))*csum(raw))/ncases)).
compute d = inv(mdiag(sqrt(diag(vcv)))).
compute realeval = eval(d * vcv * d).
compute evals = make(nvars,ndatsets,-9999).
loop #nds = 1 to ndatsets.
compute x = sqrt(2 * (ln(uniform(ncases,nvars)) * -1)) &*
cos(6.283185 * uniform(ncases,nvars)).
compute vcv = nm1 * (sscp(x) - ((t(csum(x))*csum(x))/ncases)).
compute d = inv(mdiag(sqrt(diag(vcv)))).
compute evals(:,#nds) = eval(d * vcv * d).
end loop.
end if.

* principal components analysis & raw data permutation.
do if (kind = 1 and randtype = 2).
compute nm1 = 1 / (ncases-1).
compute vcv = nm1 * (sscp(raw) - ((t(csum(raw))*csum(raw))/ncases)).
compute d = inv(mdiag(sqrt(diag(vcv)))).
compute realeval = eval(d * vcv * d).
compute evals = make(nvars,ndatsets,-9999).
loop #nds = 1 to ndatsets.


```

compute x = raw.
loop #c = 1 to nvars.
loop #r = 1 to (ncases -1).
compute k = trunc( (ncases - #r + 1) * uniform(1,1) + 1 ) + #r - 1.
compute d = x(#r,#c).
compute x(#r,#c) = x(k,#c).
compute x(k,#c) = d.
end loop.
end loop.
compute vcv = nm1 * (sscp(x) - ((t(csum(x))*csum(x))/ncases)).
compute d = inv(mdiag(sqrt(diag(vcv)))).
compute evals(:,#nds) = eval(d * vcv * d).
end loop.
end if.

```

```

* PAF/common factor analysis & random normal data generation.
do if (kind = 2 and randtype = 1).
compute nm1 = 1 / (ncases-1).
compute vcv = nm1 * (sscp(raw) - ((t(csum(raw))*csum(raw))/ncases)).
compute d = inv(mdiag(sqrt(diag(vcv)))).
compute cr = (d * vcv * d).
compute smc = 1 - (1 &/ diag(inv(cr)) ).
call setdiag(cr,smc).
compute realeval = eval(cr).
compute evals = make(nvars,ndatsets,-9999).
compute nm1 = 1 / (ncases-1).
loop #nds = 1 to ndatsets.
compute x = sqrt(2 * (ln(uniform(ncases,nvars)) * -1) ) &*
      cos(6.283185 * uniform(ncases,nvars) ).
compute vcv = nm1 * (sscp(x) - ((t(csum(x))*csum(x))/ncases)).
compute d = inv(mdiag(sqrt(diag(vcv)))).
compute r = d * vcv * d.
compute smc = 1 - (1 &/ diag(inv(r)) ).
call setdiag(r,smc).
compute evals(:,#nds) = eval(r).
end loop.
end if.

```

```

* PAF/common factor analysis & raw data permutation.
do if (kind = 2 and randtype = 2).
compute nm1 = 1 / (ncases-1).
compute vcv = nm1 * (sscp(raw) - ((t(csum(raw))*csum(raw))/ncases)).
compute d = inv(mdiag(sqrt(diag(vcv)))).
compute cr = (d * vcv * d).
compute smc = 1 - (1 &/ diag(inv(cr)) ).
call setdiag(cr,smc).

```

```

compute realeval = eval(cr).
compute evals = make(nvars,ndatsets,-9999).
compute nm1 = 1 / (ncases-1).
loop #nds = 1 to ndatsets.
compute x = raw.
loop #c = 1 to nvars.
loop #r = 1 to (ncases -1).
compute k = trunc( (ncases - #r + 1) * uniform(1,1) + 1 ) + #r - 1.
compute d = x(#r,#c).
compute x(#r,#c) = x(k,#c).
compute x(k,#c) = d.
end loop.
end loop.
compute vcv = nm1 * (sscp(x) - ((t(csum(x))*csum(x))/ncases)).
compute d = inv(mdiag(sqrt(diag(vcv)))).
compute r = d * vcv * d.
compute smc = 1 - (1 &/ diag(inv(r)) ).
call setdiag(r,smc).
compute evals(:,#nds) = eval(r).
end loop.
end if.

* identifying the eigenvalues corresponding to the desired percentile.
compute num = rnd((percent*ndatsets)/100).
compute results = { t(1:nvars), realeval, t(1:nvars), t(1:nvars) }.
loop #root = 1 to nvars.
compute ranks = rnkorder(evals(#root,:)).
loop #col = 1 to ndatsets.
do if (ranks(1,#col) = num).
compute results(#root,4) = evals(#root,#col).
break.
end if.
end loop.
end loop.
compute results(:,3) = rsum(evals) / ndatsets.

print /title="PARALLEL ANALYSIS:".
do if (kind = 1 and randtype = 1).
print /title="Principal Components & Random Normal Data Generation".
else if (kind = 1 and randtype = 2).
print /title="Principal Components & Raw Data Permutation".
else if (kind = 2 and randtype = 1).
print /title="PAF/Common Factor Analysis & Random Normal Data Generation".
else if (kind = 2 and randtype = 2).
print /title="PAF/Common Factor Analysis & Raw Data Permutation".
end if.

```

```

compute specifics = {ncases; nvars; ndatsets; percent}.
print specifics /title="Specifications for this Run:"
  /rlabels="Ncases" "Nvars" "Ndatsets" "Percent".
print results
  /title="Raw Data Eigenvalues, & Mean & Percentile Random Data Eigenvalues"
  /clabels="Root" "Raw Data" "Means" "Prcntyle" /format "f12.6".

do if (kind = 2).
print / space = 1.
print /title="Warning: Parallel analyses of adjusted correlation matrices".
print /title="eg, with SMCs on the diagonal, tend to indicate more factors".
print /title="than warranted (Buja, A., & Eyuboglu, N., 1992, Remarks on parallel".
print /title="analysis. Multivariate Behavioral Research, 27, 509-540.)".
print /title="The eigenvalues for trivial, negligible factors in the real".
print /title="data commonly surpass corresponding random data eigenvalues".
print /title="for the same roots. The eigenvalues from parallel analyses".
print /title="can be used to determine the real data eigenvalues that are".
print /title="beyond chance, but additional procedures should then be used".
print /title="to trim trivial factors.".
print / space = 2.
print /title="Principal components eigenvalues are often used to determine".
print /title="the number of common factors. This is the default in most".
print /title="statistical software packages, and it is the primary practice".
print /title="in the literature. It is also the method used by many factor".
print /title="analysis experts, including Cattell, who often examined".
print /title="principal components eigenvalues in his scree plots to determine".
print /title="the number of common factors. But others believe this common".
print /title="practice is wrong. Principal components eigenvalues are based".
print /title="on all of the variance in correlation matrices, including both".
print /title="the variance that is shared among variables and the variances".
print /title="that are unique to the variables. In contrast, principal".
print /title="axis eigenvalues are based solely on the shared variance".
print /title="among the variables. The two procedures are qualitatively".
print /title="different. Some therefore claim that the eigenvalues from one".
print /title="extraction method should not be used to determine".
print /title="the number of factors for the other extraction method.".
print /title="The issue remains neglected and unsettled.".
end if.

compute root    = results(:,1).
compute rawdata = results(:,2).
compute percntyl = results(:,4).

save results /outfile= 'screedata.sav' / var=root rawdata means percntyl .

end matrix.

```

* plots the eigenvalues, by root, for the real/raw data and for the random data;
This command works in SPSS 12, but not in all earlier versions.
GET file= 'screedata.sav'.
TSPLLOT VARIABLES= rawdata means percnyl /ID= root /NOLOG.

Appendix L

Validation Study Survey

Non-statistical validation study

Thank you for your interest in taking this survey! This survey is part of a research project at the University of Michigan. The following questions concern your perceptions of the trustworthiness of digitized documents (e.g., JPEGs). Because these questions ask about your perceptions, there are no right or wrong answers. Please try to provide an answer for every question. The survey will take approximately 5 minutes to complete. If you decide to provide your contact information at the end of this survey, it will only be used by the researcher responsible for this survey, University of Michigan Ph.D. student Devan Ray Donaldson (devand@umich.edu), so that he can contact you to discuss some of your responses. Otherwise, your responses will be kept anonymous and your contact information will be kept confidential. Would you like to continue with this survey?

- YES
- NO THANKS

If NO THANKS! Is Selected, Then Skip To End of Survey

In the following screens, you will view two digitized documents. Please examine each document and consider whether you think each document is either untrustworthy or trustworthy. Then, at the bottom of each page, list as many reasons as possible for why you think each document is either untrustworthy or trustworthy. Afterwards, you will answer a set of questions regarding each digitized document.

The following is a digitized marriage certificate. The source of this document is the Washington State Digital Archives and is available at the following link: <http://www.digitalarchives.wa.gov/DigitalObject/ViewAsJpeg?ID=6D0CEF4B6CBA02706B3CFC205052042C&ViewAsJpegPage=0> Examine the digitized document below. Do you think it is trustworthy? If so, why? If not, why not? Please list as many reasons as possible. Please provide your answers in the text box below the document toward the bottom of this page.

Recall the digitized marriage certificate you just looked at. Please indicate whether the circumstance described in each statement below would cause you to perceive that digitized document as untrustworthy or trustworthy. Select the button below the number that most closely represents your perception for each statement. To view the document again, click the "previous" button at the bottom of the page. If you click the "previous" button, your responses will be saved.

	Very untrustworthy -3	Untrustworthy -2	Slightly untrustworthy -1	Neither untrustworthy or trustworthy 0	Slightly trustworthy 1	Trustworthy 2	Very trustworthy 3
The document includes verifiable data.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is credible.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The information is correct.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document appears free from error.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is readable.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is factual.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document resembles what an official document should look like.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<p>matches other information I know about.</p> <p>The document is from the time period it claims to be.</p> <p>The document includes documentation of where it came from.</p>	<input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/>
---	--	--	--	--	--	--	--

Recall the digitized marriage certificate you just looked at. Please indicate whether the circumstance described in each statement below would cause you to perceive that digitized document as untrustworthy or trustworthy. Select the button below the number that most closely represents your perception for each statement. To view the document again, click the "previous" button at the bottom of the page. If you click the "previous" button, your responses will be saved.

	Very untrustworthy -3	Untrustworthy -2	Slightly untrustworthy -1	Neither untrustworthy or trustworthy 0	Slightly trustworthy 1	Trustworthy 2	Very trustworthy 3
The digitized document resembles what the original document should look like.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document appears complete.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document matches research I have found using other sources.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document accurately	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

reflects what happened.							
The document was created using responsible and accepted practices.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document has all the appropriate fields, even though all the fields are not completely filled out.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is what it claims to be.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is the same every time I download or view it.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<p>nt is official. The document was written at the time of the event.</p>	○	○	○	○	○	○	○
---	---	---	---	---	---	---	---

The following is a digitized birth certificate. The source of this document is the Hawaii State Department of Health and is available at the following link: http://www.whitehouse.gov/sites/default/files/rss_viewer/birth-certificate-long-form.pdf Examine the digitized document below. Do you think it is trustworthy? If so, why? If not, why not? Please list as many reasons as possible. Please provide your answers in the text box below the document toward the bottom of this page.

Recall the digitized birth certificate you just looked at. Please indicate whether the circumstance described in each statement below would cause you to perceive that digitized document as untrustworthy or trustworthy. Select the button below the number that most closely represents your perception for each statement. To view the document again, click the "previous" button at the bottom of the page. If you click the "previous" button, your responses will be saved.

	Very untrustworthy -3	Untrustworthy -2	Slightly untrustworthy -1	Neither untrustworthy or trustworthy 0	Slightly trustworthy 1	Trustworthy 2	Very trustworthy 3
The document includes verifiable data.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is credible.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The information is correct.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document appears free from error.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is readable.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is factual.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document resembles what an official document should look like.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<p>matches other information I know about.</p> <p>The document is from the time period it claims to be.</p> <p>The document includes documentation of where it came from.</p>	<input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/>	<input type="radio"/> <input type="radio"/>
---	--	--	--	--	--	--	--

Recall the digitized birth certificate you just looked at. Please indicate whether the circumstance described in each statement below would cause you to perceive that digitized document as untrustworthy or trustworthy. Select the button below the number that most closely represents your perception for each statement. To view the document again, click the "previous" button at the bottom of the page. If you click the "previous" button, your responses will be saved.

	Very untrustworthy -3	Untrustworthy -2	Slightly untrustworthy -1	Neither untrustworthy or trustworthy 0	Slightly trustworthy 1	Trustworthy 2	Very trustworthy 3
The digitized document resembles what the original document should look like.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document appears complete.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document matches research I have found using other sources.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document accurately	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

reflects what happened.							
The document was created using responsible and accepted practices.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document has all the appropriate fields, even though all the fields are not completely filled out.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is what it claims to be.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document is the same every time I download or view it.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
The document	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

<p>nt is official. The document was written at the time of the event.</p>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
---	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

What type of document do you access most frequently when using a digital archive?

- Birth records
- Death records
- Marriage records
- Military records
- Census records
- Land records
- Other: please specify _____

What is your primary reason for using documents that you typically find in a digital archive?

- Business
- Genealogy
- Historical research
- Local government
- Title searches
- Other: please specify _____

In the last year, how frequently have you used documents in a digital archive?

- Not at all
- A few times a year
- Monthly
- Weekly
- Daily

How much time do you spend using documents on your typical visit to a digital archive?

- 0-30 minutes
- 31-59 minutes
- 1-2 hours
- More than 2 hours

Name the digital archive you use most frequently:

How satisfied are you with the way that digital archive displays documents?

- Very Dissatisfied
- Dissatisfied
- Satisfied
- Very Satisfied

Do you consider yourself a:

- Democrat
- Republican
- Independent
- Libertarian
- Other: please specify _____

What is your gender?

- Male
- Female
- Alternate self-identification: please specify _____

What is your age range?

- 19 or under
- 20 to 29
- 30 to 39
- 40 to 49
- 50 to 59
- 60 to 69
- 70 or older

Are you willing to be contacted to discuss some of the answers you provided in this survey?

- Yes
- No

If YES Is Selected, Then Skip To Please provide your contact informati...If NO Is Selected, Then Skip To End of Survey

Please provide your contact information below:

First name
Last name
Email address
Phone number

Appendix M Digitized Birth Certificate

STATE OF HAWAII		CERTIFICATE OF LIVE BIRTH		DEPARTMENT OF HEALTH	
		FILE NUMBER 151		61 10611	
1. Child's First Name (Type in full)		2. Middle Name		3. Last Name	
BARACK		MUSSEIN		OBAMA, II	
4. Sex	5. This Birth	6. If Twin or Triplet, Was Child Born	7a. Birth Date	7b. Month	7c. Year
Male <input checked="" type="checkbox"/>	Single <input checked="" type="checkbox"/> Twin <input type="checkbox"/> Triplet <input type="checkbox"/>	In <input type="checkbox"/> Out <input type="checkbox"/>	August	4	1961
8. Place of Birth: City, Town or Rural Location			9. Island		
Honolulu			Oahu		
10. Name of Hospital or Institution (If not in hospital or institution, give street address)			11. Is Place of Birth Inside City or Town Limits? If not, give judicial district		
Kapiolani Maternity & Gynecological Hospital			Yes <input checked="" type="checkbox"/> No <input type="checkbox"/>		
12. Usual Residence of Mother: City, Town or Rural Location		13. Island		14. Country and State or Foreign Country	
Honolulu		Oahu		Honolulu, Hawaii	
15. Street Address			16. Is Residence Inside City or Town Limits? If not, give judicial district		
6085 Kalaniana'ole Highway			Yes <input checked="" type="checkbox"/> No <input type="checkbox"/>		
17. Mother's Mailing Address			18. Is Residence on a Farm or Plantation? Yes <input type="checkbox"/> No <input checked="" type="checkbox"/>		
19. Full Name of Father		20. Race of Father			
BARACK MUSSEIN OBAMA		African			
21. Age of Father (If Birthplace (most born in foreign country))		22. Usual Occupation		23. Kind of Business or Industry	
25 Kenya, East Africa		Student		University	
24. Full Name of Mother		25. Race of Mother			
STANLEY ANN DUNHAM		Caucasian			
26. Age of Mother		27. Birthplace (most born in foreign country)		28. Type of Occupation Outside Home During Pregnancy	
18 Wichita, Kansas		None			
I certify that the above stated information is true and correct to the best of my knowledge.			19a. Signature of Registrar		19b. Date of Signature
			<i>Alvin T. Omaha</i>		8-7-61
I hereby certify that this child was born alive on the date and hour stated above.			20a. Signature of Physician		20b. Date of Signature
			<i>David A. Simola</i>		8-8-61
21. Date Accepted by Local Reg.			22. Signature of Local Registrar		23. Date Accepted by Reg. General
AUG - 8 1961			<i>W. L. Lee</i>		AUG - 8 1961
24. Evidence for Delayed Filing or Alteration					

APR 25 2011

Alvin T. Omaha, Ph.D.
STATE REGISTRAR

Appendix N Digitized Marriage Certificate

DEC 5 1950

FRANK J. GLOVER
COUNTY AUDITOR

Y **Marriage Certificate** No 7246 B

STATE OF WASHINGTON, }
COUNTY OF Spokane } *ss.*

THIS IS TO CERTIFY, that the undersigned, a Lutheran clergyman, Pastor of
Holy Cross Lutheran Church, Spokane, Washington,

by authority of a license bearing date the 1st day of December A. D. 1950,
and issued by the County Auditor of the County of Spokane did, on the second day of
December A. D. 1950, at Spokane in this county and state, join in
(City or Town)


lawful wedlock Carl M. Abitz
of Spokane State of Washington
and Annette Reichel
Spokane State of Washington

with their mutual assent, in the presence of the following witnesses:

1. James H. Bacon and
2. Leona R. Bacon

IN TESTIMONY WHEREOF, witness the signatures of the parties to said ceremony, the witnesses and myself, this second day of December A. D. 1950.

Groom Carl M. Abitz Rt # 1 Spokane Wash,
(Address)
Bride Annette Reichel 2409 N. Alberta
(Address)
Witness James H. Bacon 2926 N. Alberta
(Address)
Witness Leona R. Bacon 2926 N. Alberta
(Address)
Officiating Clergyman or Officer Louis J. Rush

 MAIL THIS Certificate to
FRANK J. GLOVER,
County Auditor
Spokane 11, Washington

This Certificate must be filled out and filed with the County Auditor of the County wherein the license was issued, within 30 days after the ceremony. Failure to make and deliver Certificate to the County Auditor within 30 days is punishable by a fine of not less than \$25.00 or more than \$300.00.

See Chapter 59, Session Laws of Washington of 1947.