

---

# A Factor Analysis Approach to Persona Development using Survey Data

Hae Min Kim and John Wiggins  
Drexel University, USA

---

## Abstract

Managers may be skeptical of incorporating personas in making decisions about service designs or marketing the value of services because such depictions of customers may be based on limited, subjective, or unsupported impressions of customer behaviors. This study explores factor and cross-tab analyses of survey data for suggesting shared characteristics of engagement with library services and user attributes for developing personas. Creation of personas was pursued for use in a business case to engage stakeholders with shaping the libraries' strategic directions and services. Quantitative data describing reported expectations and perceptions of service offerings were acquired via a user survey, and factor analysis was applied to extract common groups of responses. Subsequent cross-tab analysis of the loaded factors with demographic and library location data suggested characteristics that can be used to develop personas. Staff confirm the suggested groupings appeared intuitive based on their experiences with users. This study's results encourage more exploration of factor analysis of quantitative data as an initial step in developing personas for different library management decisions.

## Introduction

An easily communicated understanding of user interactions helps libraries improve services. Personas represent groups of actual users, but are not real people, and help characterize common customer interactions with products and services. Personas are used in design to clarify and predict generalized user behaviors. In libraries, personas are useful for managers and designers to plan services and design spaces, and to communicate with stakeholders.<sup>1</sup> While both qualitative and quantitative methods can be used to collect data for developing personas, many cases have been based on individual and focus group interviews from such patron types as undergraduates, graduates, and faculty from various disciplines. Limiting analysis to qualitative methodologies may result in misinterpreted, stereotyped, or unbelievable personas.<sup>2</sup> In order to avoid these drawbacks, we

designed a quantitative method that utilized factor analysis to identify distinguishing characteristics to use in persona development. Factor analysis has been used in studies to develop personas for employees in a company<sup>3</sup> and restaurant customers.<sup>4</sup> These studies collected data through user surveys and conducted factor analysis which helped identify groups of users.

The purpose of this study is to examine the use of factor analysis based on user ratings of expectations and perceptions of service offerings to identify common characteristics correlated in the survey response data. This is an exploratory case study, guided by the following research questions: (1) to what extent does factor analysis of customer ratings of expectations and perception of services load common characteristics of library users, and (2) to what extent does crosstab analysis of the clustered loads from factor analysis and customer demographic data help create personas?

This study utilized survey responses to questions about most frequently identified library services at one academic library. Factor analysis applied to these responses revealed clusters of significant groups of services within the larger set of observed variables. Within a factor, each variable was triggered by a question that prompted user ratings of a service and thus describes a characteristic—an expectation or perception of a service that has correlation within the factor. Each factor therefore offers a set of correlated characteristics to consider in development of a persona. The factors will be discussed with respect to implications for library service improvement and decision making.

The study setting is the Drexel University Libraries, a mid-sized university with about 26,000 students. The study results will be valuable to library service assessment managers and administrators interested in development of personas with unbiased perspectives and in gaining insights into users' perceptions of library services, to guide strategic planning for service promotion and user outreach.

## Literature Review

Persona development began in the 1990s with use of data collected through qualitative methods. However, it has evolved to include quantitative methods because data gathering and analyses about customers could be fast and cheap, and be objective considering different features in a research scope.<sup>5</sup> Previous studies have reported that factor analysis performs well for identifying new components which are unobservable via qualitative methods.<sup>6</sup> In this section, we reviewed what kinds of methods have been used for developing library personas, and quantitative methods applied in marketing and HCI (human computing interface) studies.

A software designer and programmer, Cooper,<sup>7</sup> initiated personas development to understand target users' needs, behaviors, and attitudes for a product design using ethnographic interviews to obtain data about users. As designers and developers need to understand users with more comprehensive data, the persona development method has evolved. For example, Pruitt and Grudin<sup>8</sup> utilized quantitative as well as qualitative data collecting methods, which enriched descriptions and validated personas for software interface design.

In library settings, many practitioners and researchers have created personas based on qualitative data gathered from interviews with users and librarians that they used in library website design as well as implementation and development of other services. Koltay and Tancheva<sup>9</sup> interviewed academic library website users and identified users' expectations in order to develop a user-focused search interface. In addition to interviews with users, Phillips<sup>10</sup> added a diary study which directed participants to document their activities about their information seeking behavior on an archives website. Toward developing personas, interviews with target users have been conducted to identify perceptions of libraries' services including data curation<sup>11</sup> and to study the information needs of humanities scholars.<sup>12</sup> Library user personas were developed via a group workshop with library staff members, and the gathered assumptions were validated by user data from responses to an assessment survey and factbook statistics data.<sup>13</sup> Zohoorian-Fooladi and Abrizah<sup>14</sup> applied personas to understand how academic libraries have used social media, collecting data via interviews and focus group discussions with librarians. Qualitative data collecting methods and identifying user needs and expectations about

services have uncovered previously unknown issues, which quantitative methods have not validated.

In recent years, a few authors have begun to apply mixed methods including both qualitative and quantitative techniques for creating library personas. Tempelman-Kluit and Pearce<sup>15</sup> developed personas about their library users by extracting variations of users' motivation and information needs through content analysis of chat reference transcripts, and conducted a cluster analysis using those two factors. Four clusters were identified, and qualitative information from library activities was incorporated to finalize personas. Their analysis helped evolve persona characters without including typical academic status such as undergraduate, graduate, and faculty. Zaugg and Rackham<sup>16</sup> identified ten personas for undergraduate students in an academic library based on data gathered from previous assessment projects, observations, focus groups, and interviews with student teams. Then, they conducted a survey that asked students to choose one (out of ten) persona with which they resonate. Using the survey data, factor analysis was performed on the student choices of relevant persona; factor loads found four components among ten personas, which indicated personas that shared a common theme based on a similar component. They used factor analysis for the validation of developed personas via qualitative methods, which is different from our study approaching data collection via survey and conducting factor analysis to identify common characteristics for the design of the persona.

In HCI studies, statistical analysis has been used to identify common types of information needs, which broadened the understanding of target users without relying on demographic classification.<sup>17</sup> Researchers developed surveys to collect data that include a number of dimensions of different context in each study. For example, Sinha<sup>18</sup> and McGinn and Kotamraju<sup>19</sup> used factor analysis to discover underlying groups in their survey data of restaurant customers and one company's workers, and they developed personas with additional information including behavioral data about their user groups. Tu, et al.<sup>20</sup> applied cluster analysis to their survey data of online travel service business finding two clusters, and conducted a user observation and interviews to develop personas for the online travel service business. Brickey, et al.<sup>21</sup> compared factor and cluster analysis for persona development, and recommended factor analysis, because it finds latent components in original variables, rather than cluster

analysis, which identifies groups of cases which could have no underlying structure.

Overall, qualitative methods for creating personas in libraries have been used for design of websites and reference services. However, there are few published articles about development of library personas using factor analysis as an initial approach. Statistical analysis of survey data is a quantitative method to find important groups that are unobservable by interviews. The successful studies using factor analysis in other fields reviewed here support the method as applicable to library studies.

## Methods

A quantitative survey approach was employed in order to quickly gather data from a large population and to apply factor analysis for detecting highly related services in the libraries. The design of the questionnaire was based on service quality assessment, asking for ratings of expectation and perception of specific library services, to yield responses reflecting user reaction related to library services. The questionnaire contained thirty statements about fifteen core Drexel University Libraries services including printing, computing, events organized by the library, assistance from liaisons, staff assistance to locate books and other resources, circulation services and policies, online databases and library catalog, and spaces used for study, collaboration, and events. Each survey respondent was also prompted to report college or school affiliation, academic status, and most frequently visited library location.

To recruit participants, the survey was distributed online—via the libraries' website, through targeted e-mails to faculty, departments and student

organizations, and on Facebook—and on-site at one of the libraries' physical sites and in some classes where faculty agreed in advance to participate. The surveyor tried to gather responses meeting sample size based on colleges (Table 1).<sup>22</sup> The authors calculated sample size with 95% confidence level and  $\pm 10\%$  allowable errors based on the proportion of each college.<sup>23</sup> The proportion of sample size in each college ranges from 1.2% to 1.5% of the total student population.

Survey distribution took place in late February through early March 2016, and the target sample size was about 340. The collected data were analyzed with statistical methods to identify commonly shared characteristics. Factor analysis was performed to identify the underlying components of library services (characteristics of user engagement with services), and crosstab analysis was used to investigate the distribution of demographic values such as college affiliation, academic status, and library location visited in order to match them with the extracted components from factor analysis (user attributes). Data management and all analyses were performed using SPSS 23.0.

## Results

### Responses

The number of received valid responses to the questionnaire was 435. Distribution of respondents is summarized by college affiliation (Table 1) and patron status (Table 3). The total university student enrollment in 2015 was 26,359, distributed across seventeen colleges at Drexel. Although the overall response rate of 1.7% satisfied the overall sample size goal (1.3%), some colleges did not reach the proper ratio. Nonetheless, we used the full data set to explore the potential usefulness of the factor analysis approach.

**Table 1. Number of responses by college affiliation**

College	Enrollment (proportion)	Minimum sample size suggested	Responses	Percent of total responses
Arts and Sciences	3005 (11.4%)	38.80	48	11.1%
Biomedical Engineering Sci & Health Systems	863 (3.3%)	12.17	29	6.7%
Biomedical Sciences & Professional Studies	939 (3.6%)	13.20	21	4.8%
Business	3898 (14.8%)	48.41	39	9.0%

College	Enrollment (proportion)	Minimum sample size suggested	Responses	Percent of total responses
Computing and Informatics	1818 (6.9%)	24.67	13	3.0%
Economics	253 (1.0%)	3.65	2	0.5%
Education	1112 (4.2%)	15.52	20	4.6%
Engineering	4649 (17.6%)	55.81	24	5.5%
Entrepreneurship	13 (0.05%)	0.19	2	0.5%
Hospitality and Sport Management	409 (1.6%)	5.87	28	6.4%
Law	441 (1.7%)	6.32	5	1.1%
Media and Arts Design	2083 (7.9%)	27.96	55	12.6%
Medicine	1083 (4.1%)	15.14	73	16.8%
Nursing and Health Professions	4931 (18.7%)	58.42	48	11.0%
Honors	28 (0.11%)	0.41	0	0.0%
Professional Studies	399 (1.5%)	5.73	14	3.2%
Public Health	435 (1.7%)	6.24	13	3.0%
Missing	-	-	1	0.2%
<b>Total</b>	<b>26359 (100%)</b>	<b>338.49</b>	<b>435</b>	<b>100%</b>

### Factor analysis

First, a factor analysis was conducted. It reduces the number of variables, examines correlations among observed variables, and identifies groups of interrelated variables (each group is a factor). KMO and Bartlett's statistics showed this study data set is suitable for factor analysis, with KMO at 0.83 and Bartlett's indicated significance ( $\chi^2 (df=435) = 6858.91, p < 0.001$ ).

The initial factor analysis output suggested nine factors, which are not much reduced from the original fifteen services for which we gathered the data. The best fit is based on the ratio of factors to

total services along with total variance explained, and the grouping of variables within a factor having a probable explanation to support the persona. In order to find a reasonably small number of interpretable factors explaining the maximum amount of variance in the data, we ran multiple extractions, selecting between four and seven factors, and determined that five factors yielded the most interpretable results (Table 3). These factors suggested grouping services as follows: computer related facilities in the libraries, services providing assistance to patrons, circulation services, online databases and interlibrary loan, and library space use. The grouped services are more related to each other than to other services in different groups.

**Table 2. Factor analysis**

Item	Component loadings					Eigen Values	% of Variance
	1	2	3	4	5		
Importance of the printing service	.815	-	-	-	-	3.989	13.295
Importance of the scanners	.748	-	-	-	-		
Importance of the computers	.742	-	-	-	-		
Importance of the library space	.636	-	-	-	.426		
Satisfaction with the printing service	.510	-	.471	-	-		
Satisfaction with the scanners	.432	-	.421	-	-		
Importance of library events	-	.719	-	-	-	3.777	12.590
Importance of the services provided by library liaisons	-	.698	-	-	-		
Satisfaction with library events	-	.655	-	-	-		
Satisfaction with the services provided by library liaisons	-	.592	.455	-	-		
Importance of the help from library staff in locating books or resources	-	.585	-	-	-		
Satisfaction with borrowing books from the library for leisure reading	-	.557	-	-	-		
Importance of borrowing books from the library for leisure reading	-	.525	-	-	-		

Item	Component loadings					Eigen Values	% of Variance
	1	2	3	4	5		
Satisfaction with the service provided at the circulation desk	-	-	.654	-	-	3.023	10.078
Satisfaction with access to textbooks on Reserve	-	-	.626	.360	-		
Satisfaction with the services provided by the library staffs	-	.444	.603	-	-		
Satisfaction with the desktops and laptops provided by the library	.462	-	.498	-	-		
Frequency of visits to the circulation desk	.388	-	.439	-	-		
Importance of online databases on the library website (including e-journals, e-books, articles...)	-	-	-	.704	-	2.797	9.323
Importance of interlibrary loans	-	-	-	.665	-		
Importance of the library catalog	-	.380	-	.614	-		
Satisfaction with the interlibrary loans	-	-	-	.537	-		
Satisfaction with the library catalog	-	-	-	.525	-		
Satisfaction with the online databases on the library website	-	-	-	.520	-		
Importance of access to textbooks on reserve	.393	-	-	.456	-		
Satisfaction with the library space to organize or host an event	-	-	-	-	.775	2.733	9.109
Frequency of using library space to organize or host an event	-	-	-	-	.708		
Satisfaction with the group meeting spaces	-	-	-	-	.671		
Importance of group meeting spaces	.481	-	-	-	.622		
Satisfaction with the study spaces that the library offers	-	-	.401	-	.530		

Item	Component loadings					Eigen Values	% of Variance
	1	2	3	4	5		
Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization. Rotation converged in 7 iterations.							

When running factor analysis, factor scores were saved as variables for further analyses. Based on the factor scores from the factor analysis extracting five groups, each of the 435 cases was categorized into one of the five groups.

### Crosstab analysis

Crosstab analysis was conducted to test how the five factors might be matched to college affiliation, academic status (Table 3), and frequently visited library location (Table 4) data collected by the survey to suggest attributes that best describe the users

whose responses contributed to the characteristics assigned to the factors. The results of crosstab analysis show the percentage of distribution of each attribute within each factor. In the case of academic status, most freshman students (37.1%) are assigned to factor one as a possible user attribute to suggest for further persona development (Table 3). Among locations, the medical education library is matched to factor three (Table 4). Among the 17 colleges, we selected 10 representative sample colleges (Table 1) which met or surpassed the sample size in each population.

**Table 3. Crosstab analysis by status**

Factor Score Fresh-man		Status							Total
		Soph-omore	*Pre-Junior	Junior	Senior	Graduate /Master	Doctorate		
1	Count	14	6	9	7	22	19	13	90
	% within Status	40.0%	12.2%	29.0%	11.7%	20.2%	18.8%	26.0%	20.7%
2	Count	10	12	5	16	25	24	3	95
	% within Status	28.6%	24.5%	16.1%	26.7%	22.9%	23.8%	6.0%	21.8%
3	Count	5	6	9	12	16	16	16	80
	% within Status	14.3%	12.2%	29.0%	20.0%	14.7%	15.8%	32.0%	18.4%
4	Count	2	5	1	9	25	31	13	86
	% within Status	5.7%	10.2%	3.2%	15.0%	22.9%	30.7%	26.0%	19.8%
5	Count	4	20	7	16	21	11	5	84
	% within Status	11.4%	40.8%	22.6%	26.7%	19.3%	10.9%	10.0%	19.3%
<b>Total Count</b>		35	49	31	60	109	101	50	435
<b>% of Total</b>		8.0%	11.3%	7.1%	13.8%	25.1%	23.2%	11.5%	100.0%

\*Note: \* "Pre-Junior" designates a student in the third year of a five-year undergraduate program.

**Table 4. Crosstab analysis by library location**

Factor Main campus library		Location						Total
		24/7 section of main campus library	Medical education library	Group learning commons	Health sciences library	Don't use space		
1	Count	57	5	14	0	3	11	90
	% within Location	22.5%	38.5%	24.6%	0.0%	11.1%	13.6%	20.7%
2	Count	48	5	6	1	5	30	95
	% within Location	19.0%	38.5%	10.5%	25.0%	18.5%	37.0%	21.8%
3	Count	40	1	11	0	4	30	84
	% within Location	15.8%	7.7%	19.3%	0.0%	14.8%	37.0%	19.8%
4	Count	40	1	11	0	4	30	86
	% within Location	15.8%	7.7%	19.3%	0.0%	14.8%	37.0%	19.8%
5	Count	68	1	4	2	6	3	84
	% within Location	26.9%	7.7%	7.0%	50.0%	22.2%	3.7%	19.3%
<b>Total Count</b>		253	13	57	4	27	81	435
<b>% of Total</b>		58.2%	3.0%	13.1%	0.9%	6.2%	18.6%	100.0%

**Discussion**

The five groups of variables identified from factor analysis and the crosstab analysis of these five

factors with user demographics and reported most frequently visited library locations suggested user attributes for development of five personas (Table 5).

**Table 5. Groupings assigned by factor analysis and cross-tab assignments.**

Factor	Assigned by	
	Factor analysis	Cross-tab analysis
1	<ul style="list-style-type: none"> <li>Printing, scanning and computer use</li> </ul>	<ul style="list-style-type: none"> <li>Freshman, Pre-Junior</li> <li>Medicine, Entrepreneurship, Arts and Sciences</li> <li>Medical educational library, main campus library</li> </ul>
2	<ul style="list-style-type: none"> <li>Liaison's help via web</li> </ul>	<ul style="list-style-type: none"> <li>Senior</li> <li>Arts and Sciences, Education</li> <li>Online, 24/7 section of main campus library</li> </ul>

Factor	Assigned by	
	Factor analysis	Cross-tab analysis
3	<ul style="list-style-type: none"> <li>Borrow reserve books in circulation desk</li> </ul>	<ul style="list-style-type: none"> <li>Pre-Junior, Doctorate</li> <li>Law, Medicine</li> <li>Health education library, health sciences library</li> </ul>
4	<ul style="list-style-type: none"> <li>Online DB and ILL</li> </ul>	<ul style="list-style-type: none"> <li>Master's</li> <li>Biomedical science, Education, Public health</li> <li>Online</li> </ul>
5	<ul style="list-style-type: none"> <li>Space use for group meeting and study</li> </ul>	<ul style="list-style-type: none"> <li>Sophomore, Junior</li> <li>Media arts and design, Hospitality and Sport Management</li> <li>Main campus library, Group learning commons</li> </ul>

The first factor grouping suggests characteristics related to engagement with printing, scanning, and/or computer services, and cross-tab analysis identifies user attributes of first-year students, students in the College of Arts and Sciences, and visitors of the main campus library. The second factor suggests engagement with liaison services via the web, and user attributes of senior-year students and the College of Education. The third factor suggests engagement with reserves and circulation services and staff assistance. User attributes correlating with the third factor suggest doctoral-level students, the College of Medicine, and use of the medical education library. The fourth factor suggests characteristics including use of electronic resources and interlibrary loan services, and user attributes include master's-level students and the School of Public Health. The fifth factor suggests engagement with use of library spaces for study and group meetings with related user attributes of sophomore-level students and enrollment in the College of Media, Arts, and Design. The study authors consulted with library staff to ask to comment on how intuitive each set was to characterize a persona, and confirmed that these groupings made good intuitive sense.

The factors were passed to the team building the library's business case as suggestions for developing personas. The business case team, which determined which services to highlight among library initiatives being presented, can review the factors for suitability for further development of personas in support of the business case. For example, if a persona

describing space use is helpful to highlight in a section of the business case, the characteristics and user demographics suggested by the quantitative analysis may provide helpful elements in developing the persona. The team will need to add other elements found in personas as needed, such as including a name and developing a story that provides context for the persona's interactions with elements of the business case.

However, factor analysis has some limitations. It requires statistical background. KMO and Bartlett's test might in some cases reveal a lack of relationships among some survey data, and so factor analysis would not therefore apply. To apply factor analysis, variables must be measured at least at the interval level; thereby data gathering instruments must appropriately scale response options for questions. The larger the number of variables, the larger the sample size required for accurate factor analysis; for example, for 15 variables, 150 responses, or at least ten samples per variable, are required.<sup>24</sup> After factor analysis, further statistical methods such as cluster analysis, crosstabs, and ANOVA are recommended and may require a larger response sample.

## Conclusions

Drexel and its libraries are committed to using data and evidence for decision making. The quantitative data about library users identified in this study illustrate a data-based approach to development of personas as a tool to help tell the story of the libraries' contributions to the university mission.

The study results suggest that there are several factors to consider in developing personas of Drexel University Libraries users.

Quantitative analysis minimizes researcher bias in the interpretation of results. Staff confirmation of intuitiveness of characteristics grouped as factors suggest that including factor analysis can guide the development of personas. Ultimately, this quantitative approach offers promise to help validate previously developed personas in libraries. Avoiding interpretation and unintended filtering errors that might occur in the collection of qualitative data is the objective, starting with quantitative data to suggest personas. For example, some staff assume that students have access to computers and carry their own laptops, but quantitative data suggest that a cohesive segment of the libraries' users (13% of survey respondents) identify engagement with the libraries' provision of computing, printing, and scanning technology for print and electronic resources. Any conclusions about the validity of groupings of characteristics of service engagement and user attributes should be made with caution, and the factors and associated attributes are offered only as an illustration of the process and not as characters representing the Drexel University Libraries user population.

This exploratory case study addressed responses regarding library services, so the results of the factor analysis could only suggest characteristics and attributes for development of personas to use when planning or improving library services as a subset of the libraries' activities. Factor analysis of different data sets could address different population segments or different library activities. Since the study design is based on a single case study, generalizing findings about library users to possibly create universal personas among academic libraries will require more extensive data gathering of engagement with services and controls for variations across library settings. To define factors and confirm that the suggested factors and attributes might appropriately address library needs, the authors invite others to review the results compared to their perceptions of services and users.

It is recommended that more case studies using factor analysis with survey data be explored in other institutions. Interpretation with more data including circulation, library entrance, and e-resource usage could better identify the characteristics

and attributes suggested for development of personas by factor analysis. Also, further studies are needed to apply the data of expectation and importance ratings by survey respondents to develop personas, and validate and implement the results for library marketing, advocacy, and quality improvement management.

—Copyright 2017 Hae Min Kim and John Wiggins

## References

1. Tara J. Brigham, "Personas: Stepping into the Shoes of the Library User," *Medical Reference Services Quarterly* 32, no. 4 (2013): 443–450.
2. John Pruitt and Jonathan Grudin, "Personas: Practice and Theory," in *Proceedings of the 2003 Conference on Designing for User Experiences* (ACM, 2003), 1–15; Jonalan Brickey, Steven Walczak, and Tony Burgess, "Comparing Semi-Automated Clustering Methods for Persona Development," *IEEE Transactions on Software Engineering* 38, no. 3 (2012): 537–546.
3. Jennifer (Jen) McGinn and Nalini Kotamraju, "Data-Driven Persona Development," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (ACM, 2008), 1521–1524.
4. Rashmi Sinha, "Persona Development for Information-Rich Domains," in *CHI'03 Extended Abstracts on Human Factors in Computing Systems* (ACM, 2003), 830–831.
5. McGinn and Kotamraju, "Data-Driven Persona Development."
6. Brickey, et al., "Comparing Semi-Automated Clustering."
7. Alan Cooper, *The Inmates Are Running the Asylum: Why High Tech Products Drive Us Crazy and How to Restore the Sanity* (Sams-Pearson Education, 1999).
8. Pruitt and Grudin, "Personas: Practice and Theory."
9. Zsuzsa Koltay and Kornelia Tancheva, "Personas and a User-Centered Visioning

- Process,” *Performance Measurement and Metrics* 11, no. 2 (2010): 172–183.
10. Donald Phillips, “How to Develop a User Interface That Your Real Users Will Love,” *Computers in Libraries* 32, no. 7 (2012).
  11. Kathryn Lage, Barbara Losoff, and Jack Maness, “Receptivity to Library Involvement in Scientific Data Curation: A Case Study at the University of Colorado Boulder,” *portal: Libraries and the Academy* 11, no. 4 (2011): 915–937.
  12. Mohammad Khaled Al-Shboul and A. Abrizah, “Information Needs: Developing Personas of Humanities Scholars,” *The Journal of Academic Librarianship* 40, no. 5 (2014): 500–509.
  13. Jennifer L. Ward, “Persona Development and Use, or, How to Make Imaginary People Work for You,” in *Proceedings of the 2010 Library Assessment Conference: Building Effective, Sustainable, Practical Assessment*, October 25–27, Baltimore, Maryland (Washington, DC: Association of Research Libraries, 2011), 477.
  14. Niusha Zohoorian-Fooladi and A. Abrizah, “Personafying Academic Librarians’ Social Media Presence,” *Malaysian Journal of Library and Information Science* 19, no. 3 (2014): 13–26.
  15. Nadaleen Tempelman-Kluit and Alexa Pearce, “Invoking the User from Data to Design,” *College & Research Libraries* 75, no. 5 (2014): 616–640.
  16. Holt Zaugg and Scott Rackham, “Identification and Development of Patron Personas for an Academic Library,” *Performance Measurement and Metrics* 17, no. 2 (2016): 124–133.
  17. Sinha, “Persona Development for Information-Rich Domains.”
  18. Ibid.
  19. McGinn and Kotamraju, “Data-Driven Persona Development.”
  20. Nan Tu, Xiao Dong, P. Rau, and Tao Zhang, “Using Cluster Analysis in Persona Development.” in *International Conference on Supply Chain Management and Information Systems* (2010).
  21. Brickey, et al., “Comparing Semi-Automated Clustering.”
  22. A disciplinary unit within Drexel
  23. Glenn D. Israel, *Determining Sample Size* (University of Florida Cooperative Extension Service, Institute of Food and Agriculture Sciences, EDIS, 1992).
  24. Joe Bruin, “Newtest: Command to Compute New Test,” *UCLA: Statistical Consulting Group*, 2006, accessed December 5, 2016, <http://www.ats.ucla.edu/stat/stata/ado/analysis/>.