lq_triad.wp1 7/14/11

RUNNING HEAD: LibQUAL+® Triads

An Introduction to the *LibQUAL+® Triads* Protocol: Using Ipsative Measurement to Assess Highly Desired Outcomes

Bruce Thompson

Texas A&M University and *Baylor College of Medicine*

Martha Kyrillidou

Association of Research Libraries

Paper presented during the LibQUAL+® training session at the 9th Northumbria International Conference on Performance Measurement in Libraries and Information Services, York, England, August 22, 2011.

Bruce Thompson is distinguished professor of educational psychology and CEHD distinguished research fellow, and distinguished professor of library science, Texas A&M University, College Station, TX, and adjunct professor of allied health sciences, Baylor College of Medicine, Houston, TX. He may be contacted via e-mail at: bruce-thompson@tamu.edu.

Martha Kyrillidou is director of statistics and service quality programs at the Association of Research Libraries, Washington, DC. She may be contacted at: martha@arl.org.

Abstract

The paper presents the theory underlying the new LibQUAL+® Triads protocol. LibQUAL+® Triads is the third protocol option within the LibQUAL+® suite (i.e., LibQUAL+® Long, LibQUAL+® Lite, and LibQUAL+® Triads). In addition to providing an inherently more granular examination of performance with respect to primary user service quality desires, LibQUAL+® Triads also builds in an innovative procedure for computing *intraindividual* score reliability coefficients (McCollum & Thompson, 1980) to aid in screening out user responses that are not trustworthy. The present paper also explains the manner in which these *intraindividual* score reliability coefficients are derived within the LibQUAL+® Triads protocol, so that library service quality improvement decisions will be based only on meaningful scores.

Thompson (2007) described the origins of the LibQUAL+® protocol. The development of the protocol, and evidence for the integrity of LibQUAL+® scores, have both been quite extensively documented in the refereed journal literature (cf. Cook, Heath & B. Thompson, 2001, 2002; Cook & Thompson, 2001; Heath, Cook, Kyrillidou & Thompson, 2002; Thompson & Cook, 2002; Thompson, Cook & Heath, 2001, 2003; Thompson, Cook & Kyrillidou, 2005; Thompson, Cook & R.L. Thompson, 2002) and elsewhere (Cook, 2002). The development of the LibQUAL+® Lite protocol is also well documented (Cook, Thompson & Kyrillidou, 2010; Kyrillidou, Cook & Thompson, 2010; Thompson, Kyrillidou & Cook, 2009a, 2009b, 2010a, 2010b), and especially is one dissertation (Kyrillidou, 2009).

As noted in these publications, an important consideration in selecting the final 22 LibQUAL+® core items from the much larger initial item pool was whether or nor the items were very highly rated as "desired" by the vast preponderance of library users. After all, why would librarians care how well they were doing at providing services in areas not highly valued by their users?

However, librarians tend to be aware of the general needs of their customers, and do tend to focus their energies on highly valued outcomes. This means that as an empirical matter, as reflected in a dozen of so years of LibQUAL+® data from around the world, libraries often perform very well across the service criteria that are the most highly valued by library users.

The key paradox of library service quality assessment is that we want to focus our assessment

> *efforts on the service quality criteria that are most highly desired by users, and yet the process of measuring primarily highly-desired behaviors in and of itself inherently creates challenges to differentiating user perceptions of what libraries are most and least successfully doing.*

Purpose of the Present Paper

The purpose of this paper is to present the theory underlying the new LibQUAL+® Triads protocol. LibQUAL+® Triads is the third protocol option within the LibQUAL+® suite (i.e., LibQUAL+® Long, LibQUAL+® Lite, and LibQUAL+® Triads).

In addition to providing an inherently more granular examination of performance with respect to primary user service quality desires, LibQUAL+® Triads also builds in an innovative procedure for computing *intraindividual* score reliability coefficients to aid in screening out user responses that are not trustworthy. The present paper also explains the manner in which these *intraindividual* score reliability coefficients are derived within the LibQUAL+® Triads protocol.

LibQUAL+® Long and LibQUAL+® Lite do incorporate some methods for excluding untrustworthy data from users. First, in Long and Lite users' data are excluded if users rate higher an item on minimally-acceptable service quality level than they rate the same item on desired service quality level. Such inherently illogical "inversions" reflect casual or careless, and thus untrustworthy, responding. Second, data from given users are excluded when users

select an unacceptable number of "not applicable" responses, because such users apparently have insufficient knowledge about library services to be able to offer informed opinions. The LibQUAL+® Triads protocol offers yet another way to evaluate the trustworthiness of a given library's service quality data, so that service quality improvement decisions will be based only on meaningful scores.

"Ipsative" or Forced-Choice Response Formats

Cattell (1944) presented a measurement paradigm that distinguishes between what he called "normative" and "ipsative" measures. **Normative** measurement collects data in such a way that responses to one item do not mechanically constrain responses to the remaining items. For example, an item that is normative in its measurement features might ask,

> Please rate on a 1 (worst) to 9 (best) scale how well *library employees instill confidence in users*.

Another normative item might ask,

> Please rate on a 1 (worst) to 9 (best) scale how well *the library provides space that inspires study and learning*.

The response to the first item does not in any way mechanically constrain the response to the second item. LibQUAL+® Long and LibQUAL+® Lite are both normative measures.

As Thompson (2008) explained, "Ipsative measurement collects data such that responses to a given item constrain choices on other items. Items of this sort have forced-choice features, such

as requirements to rank order choices or to allocate a fixed number of points across a set of items" (p. 21). An example of ipsative measurement might be:

> Please allocate exactly 10 points across the following three items that reflect how important these three elements of library service quality ("0" = not important at all) are to you: (a) employees who deal with users in a caring fashion; (b) a library web site enabling me to locate information on my own; (c) community space for group learning and group study.

Some potential disadvantages to ipsative measurement seem obvious. Two potential disadvantages of ipsative measurement are that (a) "[i]psative data collection tends to make responding more complex, because specific response rules must be followed exactly" and (b) "[r]espondents may resent restrictions that have no obvious basis" (Thompson, 2008, p. 23).

However, ipsative measurement may be very appealing in measurement contexts in which forced-choices most accurately reflect the real-world ecology. For example, Milton Rokeach (1973) studied human values (e.g., beauty, wisdom, justice), and asking respondents to normatively rate values on a 1-to-5 scale would doubtless result in virtually all respondents rating virtually all choices "5." Furthermore, in the real world people must selectively make difficult choices about which values to pursue, because pursuit of all values equally is simply not realistic.

We prefer measurement methods to honor the ecology

> *in the real-world where the measured decisions are made. Forced-choice measurement has appeal when the real-world ecology is itself a forced-choice world.*

Thus, Rokeach (1973) employed ipsative measurement in his studies of people's values.

LibQUAL+® Triads

The LibQUAL+® Triads protocol can be implemented to obtain library users' granular ratings of criteria for either **desired** or **perceived** library service quality. The protocol yields data on six LibQUAL+® core items, here labelled "A" through "F". The six core items are presented in all possible combinations of item triads (i.e., three items at a time).

On the Triads *perception* protocol survey respondents are asked for each of the triads,

> At which one of the following 3 things is your library doing BEST? And at which one of the following 3 things is your library doing WORST?

Table 1 presents the 20 possible item triads for six items.

INSERT TABLE 1 ABOUT HERE

The item triads were randomized in such a manner that apparent redundancy in the triads was obscured. Furthermore, item locations within triads were randomized such that (a) a given item appeared in a given triad location (i.e., first, second, or third) approximately an equal number of times, and (b) a given item did

not appear third in one triad and then first in the immediate subsequent triad. Table 2 presents the final orders of the triads and the items within triads.

INSERT TABLE 2 ABOUT HERE

For each of the 6 items, for each person how many first, second, and third place votes each item received is computed. Also computed for each person are the mean rankings of each item.

Intraindividual Reliability Computations

Within this system of 20 triads each one of the six items "A" through "F" is compared to each other item exactly four times. For example, "A" is compared to "B" in triads 19, 17, 10, and 20:

19	**1**	A,B,C	B,A,C	1	1,3	2,4	AB
17	**2**	A,B,D	A,D,B	2	1,6	5,7	AC
10	**3**	A,B,E	B,A,E	3	2,8	5,9	AD
20	**4**	A,B,F	A,B,F	4	3,8	6,10	AE

This makes it possible for each of the 15 item pairs (e.g., AB, AC, AD, ...EF) to compute two scores for each person (e.g., two AB scores, two AC scores, two AD scores, ...two EF scores).

Each of the two scores in the 15 item pairs will be a 0, a 1, or a 2. For example, if in both triad #19 A>B (i.e., "A" is best while "B" is worst, or "A" is best while "B" is middle, or "A" is middle while "B" is worst) and in triad #10 also A>B, the leftmost AB score would be a 2. But if in triad #17 A>B, but not in triad #20, the rightmost AB score would be a 1.

If the respondent behaves with perfect consistency, the two scores for a given item pair (e.g., AB, AC, AD, ...EF) would both be a 2, both be a 1, or both be a 0. Furthermore, the two scores would match across all 15 item pairs, even though some score pairs might be 2,2, while others would be 1,1 or 0,0.

Then the Pearson correlation coefficient (see Thompson, 2008) between the 2 pairs of 15 scores can be computed for each respondent as a measure of consistency or *intraindividual reliability* (see McCollum & Thompson, 1980; Thompson, 2003). Table 3 presents the item pairs used to compute the two pair scores for each of the 15 item pairs.

INSERT TABLE 3 ABOUT HERE

Finally, persons whose intraindividual reliability coefficients indicate that their scores are not trustworthy (perhaps < 0.6 or 0.7) can be omitted from the data. In this manner librarians will base decisions about improving library service quality only on trustworthy responses!

References

Cook, C.C. (2002). A mixed-methods approach to the identification and measurement of academic library service quality constructs: LibQUAL+™. (Doctoral dissertation, Texas A&M University, 2001). Dissertation Abstracts International, 62, 2295A. (University Microfilms No. AAT3020024)

Cook, C., Heath, F., & Thompson, B. (2001). Users' hierarchical perspectives on library service quality: A "LibQUAL+™" study. College and Research Libraries, 62, 147-153.

Cook, C., Heath, F. & Thompson, B. (2002). Score norms for improving library service quality: A LibQUAL+™ study. portal: Libraries and the Academy, 2, 13-26.

Cook, C., Heath, F., & Thompson, R.L. (2000). A meta-analysis of response rates in Web- or Internet-based surveys. Educational and Psychological Measurement, 60, 821-836.

Cook, C., & Thompson, B. (2001). Psychometric properties of scores from the Web-based LibQUAL+™ study of perceptions of library service quality. Library Trends, 49, 585-604.

Cook, C., Thompson, B. & Kyrillidou, M. (2010, May). Does using item sampling methods in library service quality assessment affect score norms?: A LibQUAL+® Lite study. Paper presented at the 2nd Qualitative and Quantitative Methods in Libraries (QQML 2010) International Conference, Chania (Crete), Greece. http://www.libqual.org/documents/LibQUAL/publications/lq_gr_3.pdf

Heath, F., Cook, C., Kyrillidou, M., & Thompson, B. (2002). ARL

Index and other validity correlates of LibQUAL+™ scores. portal: Libraries and the Academy, 2, 27-42.

Kyrillidou, M. (2009). Item sampling in service quality assessment surveys to improve response rates and reduce respondent burden: The LibQUAL+ Lite randomized control trial (RCT) (Doctoral dissertation, University of Illinois). Retrieved from https://www.ideals.illinois.edu/bitstream/handle/2142/14570/Kyrillidou_Martha.pdf?sequence=3

Kyrillidou, M., Cook, C. & Thompson, B. (2010, May). Does using item sampling methods in library service quality assessment affect zone of tolerance boundaries?: A LibQUAL+® Lite study. Paper presented at the 2nd Qualitative and Quantitative Methods in Libraries (QQML 2010) International Conference, Chania (Crete), Greece. http://www.libqual.org/documents/LibQUAL/publications/lq_gr_2.pdf

McCollum, J., & Thompson, B. (1980, April). Analysis of attitudinal data: Dealing with "response bias". Paper presented at the annual meeting of the American Educational Research Association, Boston. (ERIC Document Reproduction Service No. ED ??)

Rokeach, M. (1973). The nature of human values. New York: Free Press.

Thompson, B. (2003). Understanding reliability and coefficient alpha, really. In B. Thompson (Ed.), Score reliability: Contemporary thinking on reliability issues (pp. 3-23). Newbury Park, CA: Sage.

Thompson, B. (2007). The origins/birth of LibQUAL+®. Retrieved from http://www.coe.tamu.edu/~bthompson/libbirth.htm

Thompson, B. (2008). Foundations of behavioral statistics: An insight-based approach. New York: Guilford.

Thompson, B., & Cook, C. (2002). Stability of the reliability of LibQUAL+™ scores: A "Reliability Generalization" meta-analysis study. Educational and Psychological Measurement, 62, 735-743.

Thompson, B., Cook, C., & Heath, F. (2001). How many dimensions does it take to measure users' perceptions of libraries?: A "LibQUAL+™" study. portal: Libraries and the Academy, 1, 129-138.

Thompson, B., Cook, C., & Heath, F. (2003). Structure of perceptions of service quality in libraries: A LibQUAL+™ study. Structural Equation Modeling, 10, 456-464.

Thompson, B., Cook, C., & Kyrillidou, M. (2005). Concurrent validity of LibQUAL+™ scores: What do LibQUAL+™ scores measure? Journal of Academic Librarianship, 31, 517-522.

Thompson, B., Cook, C., & Thompson, R.L. (2002). Reliability and structure of LibQUAL+™ scores: Measuring perceived library service quality. portal: Libraries and the Academy, 2, 3-12.

Thompson, B., Kyrillidou, M., & Cook, C. (2009a). Equating scores on "lite" and long library user survey forms: The LibQUAL+® Lite randomized control trials. Performance Measurement & Metrics, 10, 212-219.

Thompson, B., Kyrillidou, M., & Cook, C. (2009b). Item sampling in

service quality assessment surveys to improve response rates and reduce respondent burden: The "LibQUAL+® Lite" example. Performance Measurement & Metrics, 10, 6-16.

Thompson, B., Kyrillidou, M., & Cook, C. (2010a, May). Does using item sampling methods in library service quality assessment compromise data integrity?: A LibQUAL+® Lite study. Paper presented at the 2nd Qualitative and Quantitative Methods in Libraries (QQML 2010) International Conference, Chania (Crete), Greece. http://www.libqual.org/documents/LibQUAL/publications/lq_gr_1.pdf

Thompson, B., Kyrillidou, M., & Cook, C. (2010b, October). Does using item sampling methods in library service quality assessment compromise data integrity or zone of tolerance interpretation?: A LibQUAL+® Lite study. Paper presented at the 3rd biennial Library Assessment Conference, Baltimore, MD.

Table 1
20 Triads for the Six Items "A" Through "F"

6 items		AB	AC	AD	AE	AF	BC	BD	BE	BF	CD	CE	CF	DE	DF	EF			Score Pairs for Intraindividual Reliability Coefficients		
1	A,B,C	1	1				1										3	1	1,3	2,4	AB
2	A,B,D	1		1				1									3	2	1,6	5,7	AC
3	A,B,E	1			1				1								3	3	2,8	5,9	AD
4	A,B,F	1				1				1							3	4	3,8	6,10	AE
5	A,C,D		1	1							1						3	5	4,9	7,10	AF
6	A,C,E		1		1							1					3	6	1,12	11,13	BC
7	A,C,F		1			1							1				3	7	2,14	11,15	BD
8	A,D,E			1	1									1			3	8	3,14	12,16	BE
9	A,D,F			1		1									1		3	9	4,15	13,16	BF
10	A,E,F				1	1										1	3	10	5,17	11,18	CD
11	B,C,D						1	1			1						3	11	6,17	12,19	CE
12	B,C,E						1		1			1					3	12	7,18	13,19	CF
13	B,C,F						1			1			1				3	13	8,17	14,20	DE
14	B,D,E							1	1					1			3	14	9,18	15,20	DF
15	B,D,F							1		1					1		3	15	10,19	16,20	EF
16	B,E,F								1	1						1	3	16			
17	C,D,E										1	1		1			3	17			
18	C,D,F										1		1		1		3	18			
19	C,E,F											1	1			1	3	19			
20	D,E,F													1	1	1	3	20			
	Sum	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4					
		AB	AC	AD	AE	AF	BC	BD	BE	BF	CD	CE	CF	DE	DF	EF					

Table 2
Triad and Item-Within-Triad Orders
in the Final LibQUAL+® Triads Protocol

Final Item Order			Final Choice Order	A	B	C	D	E	F	Sum
1	11	B,C,D	C,D,B		3	1	2			6
2	8	A,D,E	E,D,A	3			2	1		6
3	15	B,D,F	D,F,B		3		1		2	6
4	7	A,C,F	A,F,C	1		3			2	6
5	17	C,D,E	D,C,E			2	1	3		6
6	12	B,C,E	C,E,B		3	1		2		6
7	18	C,D,F	D,C,F			2	1		3	6
8	14	B,D,E	E,B,D		2		3	1		6
9	13	B,C,F	F,B,C		2	3			1	6
10	3	A,B,E	B,A,E	2	1			3		6
11	9	A,D,F	F,A,D	2			3		1	6
12	19	C,E,F	E,C,F			2		1	3	6
13	6	A,C,E	C,A,E	2		1		3		6
14	16	B,E,F	B,E,F		1			2	3	6
15	5	A,C,D	C,D,A	3		1	2			6
16	20	D,E,F	E,F,D				3	1	2	6
17	2	A,B,D	A,D,B	1	3		2			6
18	10	A,E,F	F,E,A	3				2	1	6
19	1	A,B,C	B,A,C	2	1	3				6
20	4	A,B,F	A,B,F	1	2				3	6
				A	B	C	D	E	F	
			1's	3	3	4	3	4	3	
			2's	4	3	3	4	3	3	
			3's	3	4	3	3	3	4	
			Sum	10	10	10	10	10	10	

Table 3
15 Pairs of Scores Used to Compute
Intraindividual Reliability Coefficients for the Triad Data

	Triad			15	Pairs of Scores		
19	**1**	A,B,C	B,A,C	1	1,3	2,4	AB
17	**2**	A,B,D	A,D,B	2	1,6	5,7	AC
10	**3**	A,B,E	B,A,E	3	2,8	5,9	AD
20	**4**	A,B,F	A,B,F	4	3,8	6,10	AE
15	**5**	A,C,D	C,D,A	5	4,9	7,10	AF
13	**6**	A,C,E	C,A,E	6	1,12	11,13	BC
4	**7**	A,C,F	A,F,C	7	2,14	11,15	BD
2	**8**	A,D,E	E,D,A	8	3,14	12,16	BE
11	**9**	A,D,F	F,A,D	9	4,15	13,16	BF
18	**10**	A,E,F	F,E,A	10	5,17	11,18	CD
1	**11**	B,C,D	C,D,B	11	6,17	12,19	CE
6	**12**	B,C,E	C,E,B	12	7,18	13,19	CF
9	**13**	B,C,F	F,B,C	13	8,17	14,20	DE
8	**14**	B,D,E	E,B,D	14	9,18	15,20	DF
3	**15**	B,D,F	D,F,B	15	10,19	16,20	EF
14	**16**	B,E,F	B,E,F				
5	**17**	C,D,E	D,C,E				
7	**18**	C,D,F	D,C,F				
12	**19**	C,E,F	E,C,F				
16	**20**	D,E,F	E,F,D				