

"Existing Collection Strength" and Shelflist Count Correlations in RLG's *Conspectus for Music*

William E. McGrath and Nancy B. Nuzzo

"Existing Collection Strength" (ECS) is an estimate of the extent to which subject areas in Research Libraries Group (RLG) collections meet research level objectives on a scale of 0 to 5. ECS can be interpreted as a proxy for shelflist counts. To test that assumption, two main sets of correlations were computed: 1) 17 correlations for 17 RLG libraries across 138 LC class ranges, indicating how well the scale represents collection strength within a library; 2) 138 correlations for 138 LC ranges across 17 libraries, indicating how well those ECS estimates were made from library-to-library. "Within library" correlations ranged from 0.36 to 0.89, all significant, with a confidence interval of the mean (0.66) from 0.55 to 0.74. "Across library" correlations ranged from -0.34 to 0.94, with none significant below 0.47. The mean was 0.54, with a confidence interval 0.08 to 0.81. As a proxy variable for shelflist within libraries, ECS is expeditious but, necessarily, less than perfect, and is, therefore, a trade-off between convenience and accuracy. However, the wide range of correlations across libraries, from inverse to highly positive, indicates that using ECS to compare libraries is not appropriate, probably because each library interprets and applies the ECS scale differently.



The Research Libraries Group (RLG) *Conspectus* is an unusual source of data about library collections.¹ According to Nancy E. Gwinn and Paul H. Mosher, it was intended to help RLG members evaluate and compare their collections and to facilitate coordinated collection development by summarizing existing and future collection strengths.² Its large

size and widespread recognition by RLG members and others suggest that a close examination of its data would yield interesting insights for collection development and analysis in general.

PURPOSE

A novel concept of the *Conspectus* is the variable "existing collection strength," or ECS. ECS was intended to

William E. McGrath is Professor in the School of Information and Library Studies and Nancy B. Nuzzo is an Associate Librarian in the Music Library at State University of New York at Buffalo, Buffalo, New York 14260. This study was supported by the Council on Library Resources (CLR Grant 4015). The authors wish to thank Leslie Hume and James Coleman of the Research Libraries Group for their interest and support. We also thank the Council on Library Resources and Deanna Marcum, formerly of CLR, for their support. RLG staff and member libraries do not necessarily endorse these findings; all errors, interpretations and misinterpretations are solely our own.

be a shortcut for representing collections by assigning numerical values to predefined "levels" of RLG member library holdings. How well ECS represents collections is an important consideration in its continued use. This study addresses that question by examining the data in one *Conspectus*, that for music libraries.³ In addition to ECS values, the *Conspectus for Music* contains shelflist counts of member library collections. These counts can be directly compared or correlated to ECS values.

RLG itself has conducted numerous "verification" studies of its collections. These interesting studies are available from RLG, and have been reported by Paul H. Mosher and Jim Coleman.⁴ Our study was not affiliated with those verification studies, though we received complete cooperation from RLG.

Rationale for correlating ECS and shelflist counts can be found on page 2-1 of the *RLG Collection Development Manual* containing the following description of ECS.

1. These values [ECS] describe collections or collection policies absolutely, not relatively. They assume, therefore, a national perspective and a broad cognizance of all facets of collecting.
2. When the value describes existing collection strength, it should relate to national shelflist measurement, reflecting what is actually on the shelves.⁵

We interpret these statements to mean that ECS is intended to mirror library holdings, i.e., to reflect actual shelflist counts.

METHOD

ECS values are expert judgments intended to represent predefined levels of collection strength in each library in specified LC classification ranges. These values, or "levels" as they are called in the *Conspectus*, range from 0 to 5.⁶ The designers of the *Conspectus* assigned each of these levels a specific definition of collection strength. For the purpose of

this analysis, these definitions are irrelevant, except to note that 0 is the minimum and 5 the maximum, and that each value signifies the strength of holdings as judged by collection development librarians.

The ECS 6-point scale can be interpreted as an interval scale variable. An interval scale usually has a fixed range, and the difference between two adjacent scores is the same as that between any other adjacent scores. For example, in the Fahrenheit temperature scale, which ranges from 0 to 100, the difference between 60 and 62 degrees is the same as that between 62 and 64 degrees. With ECS, the difference between level 3 and level 4 is taken to be the same as between any other level such as 2 and 3 or 4 and 5, without regard to the actual number of holdings the levels are intended to represent.

A ratio-level scale is one in which 0 represents a true absence of value, rather than merely the smallest number, and the largest number has no theoretical maximum. Any two values can be expressed as a ratio. Thus, ECS is not quite a ratio-scale variable because a level of 4, for example, is not twice the level of 2.

An ordinal scale is one in which scores or counts are not implicit in the ranks; one rank is simply lower or higher than another, while ranks have no theoretical maximum. Thus, ECS is probably more than ordinal because shelflist counts are implicit, even though not expressed, and because it also has a maximum permissible value. Interpretation of the data as ordinal would result in a large number of tie ranks—with most of them tied at 2 or 3. Rank order correlation methods do not handle large numbers of ties very efficiently.

ECS values are intended to "reflect what is actually on the shelves," whether 10, 100, or 10,000 volumes. Most elementary statistics texts explain that ratio scales contain more information than interval, interval more than ordinal, and ordinal more than nominal. A collection containing 10,000 volumes is twice that of 5,000, but an ECS of 4 is not twice that of 2. Thus, by *estimating* existing

strength, much information about *actual*, highly variable shelflist counts is not available. Of course, that is presumably the reason for using ECS in the first place—when shelflist data are not available, or too difficult, expensive, or time-consuming to collect. The important question is, how much information is lost? In this study, that question becomes, how well does the ECS interval scale represent the shelflist ratio scale?

Thus, for the purpose of analysis only, ECS values are treated as interval scale. The danger therein is that the ECS values will be regarded as containing more information than they in fact do. The question of whether they are truly interval, ordinal, or something else is left unanswered.

Another major statistical question is whether each library interprets the scale values in the same way. For example, is a 3 in one library the same as a 3 in another? The question may be addressed in part by computing cross-library correlations.

If ECS values and shelflist counts agree fairly well, then collection development librarians could feel confident in using ECS.

If ECS values and shelflist counts agree fairly well, then collection development librarians could feel confident in using ECS. To measure the extent of that agreement, three sets of Pearson product-moment correlations were computed between music library ECS values and their corresponding shelflist counts (SHELFLIST): 1) correlation within libraries, with LC ranges as the units of analysis; 2) correlation across libraries, with libraries as the units of analysis; and 3) correlations with the units of analysis as a composite of LC ranges and libraries. Correlational methods depend a great deal on how the variables are defined—whether they are continuous or dichotomous, nominal, ordinal, interval or ratio—and on the research objective.⁷

Pearson correlations are appropriate for interval-level data but require that data to be normally distributed and have the same variance. Both ECS and SHELFLIST violate this assumption to some degree. This violation requires that one or both variables undergo some transformation to bring them into line with one another. In this study, a logarithmic transformation of SHELFLIST was made, rendering it fairly normal, prior to computation of correlations. This transformation has the effect of improving the correlations somewhat.

RESULTS

Correlation within Libraries

"Correlation within libraries" tests the hypothesis that the ECS values of a particular library and its corresponding shelflist counts are the same. Complete shelflist counts of 138 LC music classification ranges were available for 17 music libraries. Here, the unit of analysis is the LC music classification range with a sample size of 138 classes (figure 1). Correlation coefficients for these 17 are tabulated in table 1. All coefficients were computed in the LOTUS 1-2-3 microcomputer spreadsheet program.⁸

The computed coefficients are estimates of the true correlation, which we cannot know precisely, but which can be reported as confidence intervals. Their accuracy depends on the sample size: the larger the sample the more accurate the coefficient. The best we can say is that 1) the true correlations lie somewhere between lower and upper limits of the confidence intervals, or that 2) the true correlation must be at least some minimum value to be significant. Table 1 shows confidence intervals as well as the smallest coefficient required to be significant.

Correlations range from 0.36 to 0.89, with mean of 0.66 and confidence interval 0.55 to 0.74. The interval for the computed mean is also shown. The true mean lies somewhere between 0.55 and 0.74, with $n = 17$. The overall correlation of 0.67 agrees with the mean for libraries, 0.66.

LC Classification Range	ECS	SHELFLIST	Log of SHELFLIST
M1	3	94	1.97
M2	4	1920	3.28
M3-3.1	4	486	2.68
M5	3	37	1.57
M6-19	3	810	2.9
M20-39	4	2114	3.25
.	.	.	.
.	.	.	.
.	.	.	.
-etc.-			

Figure 1. Partial tabulation of ECS values, with shelflist counts and their logs within one library, NYBS (SUNYBuffalo), from the *Conspectus for Music*, May 1986. Sample size=138. Correlation coefficients were computed for this library and 16 others.

Random error, sample size, and variables not under a library's control all affect correlations. When confidence intervals for two libraries overlap, it is incorrect to say that one library has done better than another. Their correlation coefficients are unlikely to be significantly different. Because our intent was not to compare libraries, they are listed in table 1 alphabetically rather than ranked by correlation coefficients.

Correlation across Libraries—by LC Classification Range

Correlation by LC range tests the hypothesis that the ECS values and their corresponding shelflist counts are the same for a particular classification range. In a very important sense, what is being tested is whether each library interprets collection strength in the same way. That is, does "existing collection strength" mean the same thing in each library? Is a value of 3 in one library the same as a value of 3 in another library? High correlations would indicate that libraries tend to interpret ECS values in the same way.

Here the unit of analysis is the library, and the sample size is 17 (figure 2). One hundred thirty-eight correlation coefficients between ECS and SHELFLIST were computed for each of the 138 LC classification ranges. These coefficients

are tabulated in table 2. They ranged very broadly from -0.34 to 0.94. It should be noted immediately that a sample of 17 is not large, hence it is not surprising that a substantial number of correlations, about 37%, were not significant. Given the small sample size, only correlations higher than 0.47 can be considered significant at the 95% confidence level.

The small sample size also accounts for the lack of confidence in the mean correlation with a broad interval of .08 to 0.81. This interval indicates that the true mean could have any positive value between those two extremes, from virtually no correlation to a very high correlation. When only 17 libraries have counted their shelflists, the mean ECS across libraries is not especially reliable. Should the number of libraries counting their shelflists be increased substantially, the mean ECS might then be somewhat more reliable.

General Correlation between ECS Values and Shelflist Counts Irrespective of Library or LC Classification

Correlations may also be tested without categorizing ECS and shelflist values by library or music classification ranges, and instead may be tested by treating each ECS/shelflist comparison as an observation either from the overall *Music*

TABLE 1.
 "WITHIN-LIBRARY" CORRELATION COEFFICIENTS* BETWEEN ECS
 VALUES AND LOGARITHMS OF SHEFLIST FOR 17 MUSIC LIBRARIES.

Library	Code	Correlation Coefficient	Confidence Interval
Brigham Young Univ.	UTBG	0.36	(0.16-0.54)
Colorado State Univ.	COSG	0.49	(0.30-0.63)
Columbia University	NYCG	0.56	(0.39-0.69)
Cornell University	NYCX	0.73	(0.61-0.82)
Dartmouth	NHDG	0.54	(0.37-0.68)
Johns Hopkins	MDJG	0.81	(0.71-0.87)
New York University	NYUG	0.67	(0.53-0.77)
Northwestern Univ.	ILNG	0.61	(0.45-0.73)
Notre Dame	INDG	0.89	(0.83-0.92)
Rutgers University	NJRG	0.68	(0.55-0.78)
Stanford University	CSUG	0.46	(0.26-0.61)
SUNY-Buffalo	NYBS	0.76	(0.65-0.84)
Temple University	PATG	0.72	(0.60-0.81)
Univ. Calif.-Berkeley	CUBG	0.63	(0.47-0.74)
Univ. of Michigan	MIUG	0.77	(0.66-0.84)
Univ. of Oklahoma	OKUG	0.71	(0.58-0.81)
Yale University	CtYG	0.77	(0.66-0.84)
Mean correlation (n=17)		0.66	(0.55-0.74)

* All are significant @ 99% confidence level for 137 df. (Smallest correlation needed to be significant = 0.23)

Conspectus, from music scores only, or from music literature only. Here, the unit of analysis is neither "Library" nor "LC range." It is, instead, a composite—i.e., "Library/LC range," as in figure 3. Here, the sample is much larger—virtually the entire population of ECS values and shelflist counts—and the confidence much better. Three coefficients were computed, 0.67, 0.71 and 0.64, for the overall *Music Conspectus*, for Music Scores, and for Music Literature (table 3). Though these correlations may seem moderately high, and though the large sample size may seem to provide high statistical confidence, they should be interpreted with caution for reasons given previously.

DISCUSSION

ECS as a Surrogate Measure

The methods discussed above show that 1) correlations between estimates of collection strength and actual counts of holdings are always somewhat less, and sometimes substantially less, than perfect; 2) correlations vary from library to library and from subject to subject; and 3) confidence in those correlations can be expressed only in terms of a range between some high and some low. The correlational approach measures the extent to which ECS agrees with actual shelflist counts and the statistical confidence we can have in those correlations. Because ECS values are surrogates of shelflist

Library	ECS	SHELFLIST	Log of SHELFLIST
COSG	2	84	1.92
CSUG	3	399	2.6
CTYG	4	692	2.84
CUBG	3	437	2.64
ILNG	3	330	2.52
MDJG	1	95	1.98
MIUG	2	443	2.65
NHDG	2	118	2.07
NJRG	3	198	2.3
NYBY	2	296	2.47
NYCG	3	440	2.64
NYCX	4	512	2.71
NYUG	2	218	2.34
OKUG	1	68	1.83
PATG	3	202	2.31
UTBG	3	113	2.05

Figure 2. Typical array of ECS, with shelflist counts and their logarithms, for LC range ML2900-3275, *Conspectus for Music*, May 1986 (see appendix A); 138 correlations between ECS and log of SHELFLIST, with $N = 17$, were computed from such data.

counts, they are necessarily and by definition less reliable than actual counts. Qualitatively, the agreement between ECS and shelflist counts might be considered moderately low to moderately high—about what should be expected for judgmental data.

Within-Library Correlations

The correlation coefficient in the "within-library" test is for only one library at a time. Thus, a coefficient for one library says nothing about that for any other library.

Furthermore, a within-library coefficient explains nothing about the library or its collection. It is, at best, a statistic describing a tenuous relationship between two ways of measuring the same thing. No inferences can nor should be made from it.

Across-Library Correlations

"Across-library" correlations ranged broadly—from moderately negative to nearly perfect. When a large number of statistics such as correlations and means are computed, a very broad range of results is not unusual in many contexts. But how could estimates of collection strength result in negative correlations? It happens this way: one library estimates its strength to be 5 in some LC range, with a shelflist count of 450 volumes; another estimates its strength in the same range to be 3, with a count of 700 volumes. Obviously these numbers have an inverse relationship, a negative correlation, and obviously both libraries have different perceptions of 3 or 5. In theory, each library estimates the strength of an LC range relative to some absolute strength. In practice, estimates

TABLE 2.
FREQUENCY DISTRIBUTION OF CORRELATION
COEFFICIENTS BETWEEN ECS VALUES AND LOGARITHMS OF SHELFLIST
FOR 138 LC MUSIC CLASSES ACROSS LIBRARIES

Distribution (no. of ranges)	Frequency class limits of correlation coefficients ^{a,b}	Significance ^c
1	-0.34 to -0.20	not sig (95%)
2	-0.19 to -0.05	not sig (95%)
4	-0.04 to 0.08	not sig (95%)
9	0.09 to 0.23	not sig (95%)
13	0.24 to 0.37	not sig (95%)
22	0.38 to 0.51	not sig (99%)
35	0.52 to 0.65	sig (95%)
39	0.66 to 0.79	sig (99%)
13	0.80 to 0.94	sig (99%)
<u>138</u>	Total	

a) These are frequency limits, not confidence intervals.

b) Mean correlation = 0.54, conf.interv. = (0.08-0.81). Median = 0.59.

c) Correlations smaller than 0.47 and 0.59 not significant @ 95% and 99% confidence level, respectively, df = 15.

are probably sometimes made relative to perceptions of strength in other libraries. Because different perceptions are inevitable, and because negative correlations are hardly acceptable relationships between estimated and actual collection strengths, cross-library comparisons should not be trusted and are probably not intended by RLG. The very broad range of correlations may substantiate what RLG meant when it said ECS values "describe collections absolutely, not relatively," i.e., not relative to other libraries.

Because different perceptions are inevitable, cross-library comparisons should not be trusted and are probably not intended by RLG.

General Correlations: Overall, Music Scores, Literature

The moderately high general correlations for the composite library/LC data are somewhat more reliable than those for "within" and "across" because of the

large sample size. They should still be treated with caution, of course, for the same reason that all correlations in this study should be suspect. Because one purpose for making ECS judgments in the first place was to establish a measure for facilitating library cooperation, coefficients having no reference to specific libraries may be academic. Yet, these correlations could be useful as standard indexes, to which all other correlations could be compared. For example, if a new library were to join the Group and wished to evaluate its collection, it would be helpful to know that the overall correlation of 0.67 could be used as a frame of reference.

Questions of Methodology

One question about using logarithmic transformations of SHELFLIST is that ECS is also a transformation of SHELFLIST. ECS is a judgmental transformation from true, essentially continuous, ratio data (SHELFLIST), to perceptual, discrete, interval-level data. In effect, we are, in a circuitous procedure, correlating SHELFLIST with itself; more specifically, two transformations of shelflist

Library/LC range	ECS	SHELFLIST	Log of SHELFLIST
CSUG/ML47-54	3	1064	3.00
CSUG/ML2900-3275	3	399	2.60
CSUG/ML3300-3354	2	3	0.47
.	.	.	.
CTYG/ML47-54	4	806	2.90
CTYG/ML2900-3275	4	692	2.84
CTYG/ML3300-3354	3	3	0.47
.	.	.	.
NYCX/ML47-54	4	1568	2.71
NYCX/ML2900-3275	4	512	1.25
NYCX/ML3300-3354	3	18	3.19
.	.	.	.
.	.	.	.
-etc.-			

Figure 3. Partial tabulation of ECS values, shelflist counts, and their logarithms from two composite library/LC ranges in the *Conspectus for Music*, May 1986. Data for general test contains all ranges and all libraries, for a total sample of 2,346 (17 x 138).

TABLE 3. ONE-TO-ONE CORRELATIONS
BETWEEN ECS AND LOGARITHMS OF SHELFLIST.

	Correlation	Sample Size (n)
Overall	0.67	2346
Music scores only	0.71	952
Literature only	0.64	1394

are made with each other: 1) ECS, the judgmental transformation of SHELFLIST, and 2) the logarithmic transformation of SHELFLIST. A more satisfactory approach would be to identify or define other behavioral variables to correlate with SHELFLIST. Data for behavioral variables, of course, may be hard to come by, whereas ECS is relatively easy to obtain.

Sources of Error

RLG librarians are the first to point out the sources of statistical error in the *Conspectus*: the inherent variability of collections, differences in data tabulation, and different procedures for counting, esti-

ating, and reporting. Some of these sources, particularly those concerning the definition of ECS, are noted in the *Conspectus* itself. Even if these sources were perfectly controlled, random error would still exist. If correlations were perfect or nearly perfect, then both the statistical methods and the statistician would be suspect.

The Question of Quality

At the start of this study we assumed that ECS was defined and adopted as a surrogate for actual strength and that actual strength, if measurable at all, could only be measured by the number of items in a library's shelflist. We were,

of course, very careful not to interpret "strength" as quality, an interpretation which would generate vigorous debate. Some librarians maintain that the number of items in a collection is a sufficient measure of quality, while others say that quality is too nebulous a thing to pin down with counts of anything. Though all would agree that strength, quantity, and quality are different things, we offer no rigorous definitions here.

Many attributes of the collection may affect quality: age, recency, specific titles, current acquisitions, budget, selectivity, accessibility, availability, location, as well as the manner of the collection's organization and administration. Additionally, user impressions, demand, testimonials, reputation of the parent institution, and a host of other internal and external attributes all may have an effect. Whether these attributes do or do not measure quality depends very much on how they are perceived and defined. One thing is certain: the extent to which these attributes enter into the estimates of strength must also affect confidence in the ECS concept as well as trust in shelflist counts.⁹

Shelflist as a Measure of Collection Strength

An alternative way of looking at collection strength is to think of it as a theoretical concept, like Plato's ideal, and that neither ECS nor SHELFIST can be considered actual strength. Rather, both are estimates of that mystical ideal. Strength would then be the sum of all the significant attributes discussed above plus numerical counts and anything else

one might want to throw in. At best, ECS and SHELFIST alone would then be less than perfect indicators of that strength. Practically speaking, because there are as many ideals as collection development librarians, and none of them more attainable than another, perhaps the most sensible approach is to think of a shelflist count as a reasonable measure of the number of books and of ECS as a flawed but expeditious substitute for those counts.

Further Research

Further research might involve more detailed investigation into the ordinal and interval properties of ECS as well as alternative ways of estimating collection strength. Sampling shelflists as an alternative to judgment may be a more reliable and a more orthodox procedure. Identification of other variables which could be correlated with shelflist counts is another possible direction for research. If ECS becomes a more widely used method for estimating collection strength, investigations into the statistical concept of validity should be undertaken. Other issues that require investigation include determining whether low within-library correlations suggest a flaw in *Conspectus* methodology or hint at the presence of collections with unique strengths, or whether they imply characteristics of the literature that are not evident from shelflist counts. Do low between-library correlations indicate subject areas that require special attention? Researchers need to take up such issues.

REFERENCES AND NOTES

1. "Conspectus Points Way to Cooperative Collection Development," *Research Libraries Group News* 5:12-14 (September 1984).
2. Nancy E. Gwinn and Paul H. Mosher, "Coordinating Collection Development," *College & Research Libraries* 44:128-40 (March 1983).
3. Research Libraries Group, *RLG Conspectus; Conspectus for Music* (Stanford, Calif.: Research Libraries Group, July 1986).
4. Paul Mosher, "The Nature and Uses of the RLG Verification Studies," *C&RL News* 46 no. 7:336-38 (July/Aug. 1985); Jim Coleman, "Verification Studies: Design and Implementation," *C&RL News* 46 no. 7:338-40 (July/Aug. 1985).

5. Research Libraries Group, *Collection Development, Manual*, 2d ed. (Stanford, Calif.: Research Libraries Group, unpublished document), p. 2-1.
6. Ibid.
7. Edward G. Carmines and Richard A. Zeller, *Reliability and Validity Assessment*, Sage University Paper Series on Quantitative Applications in the Social Sciences, No. 17 (Beverly Hills, Calif., and London: Sage Publications, 1979), p. 17-27; Walter Dick and Nancy Hagerty, *Topics in Measurement: Reliability and Validity* (New York: McGraw-Hill Book Company, 1971), p. 101-28.
8. *LOTUS 1-2-3 Computer Programs* (Cambridge, Mass.: Lotus Development Corporation).
9. Numerous notes in the *Conspectus* describing special features, emphases, exceptions, restrictions, qualifications, and other factors which contribute to the erosion of statistical confidence testify to the difficulties in establishing consistent measurement.



**COLLECTION
ANALYSIS**

**How does
your collection
measure up?**

You need reliable quantitative data to justify collection management decisions. You need a flexible analysis system designed with your library's goals in mind.

*Introducing.....OCLC/AMIGOS
Collection Analysis Systems.*

Collection Analysis CD compares your holdings against those of similar institutions, using a subset of the OCLC database on compact disc.

Tape Analysis gives you a custom-designed MARC tape analysis for your library or group. A Tape Match against *Books for College Libraries* is also offered.

OCLC/AMIGOS
Collection Analysis Systems

Available exclusively in the U.S. from
AMIGOS Bibliographic Council, Inc.
11300 North Central Expressway, Suite 321
Dallas, Texas 75243
(800)843-8482 (214)750-6130