Common patterns and interrelationships in Psychometrics and research methods in Library and Information Science suggest a need for systematized methodologies for research and assessment of information processing and human perception and use. This exploratory study shows where psychometrics may be practically applied to usability assessment of library online databases, bibliographic management software, and digital/virtual learning tools.

Usability, user perception, competencies, assessment are placed within various contexts of psychometric concepts, such as validity and reliability, and applied to an Information Search Process (ISP) model. Triangulation, integration, and broadening of commonalities are analyzed in contexts of usability of library online resources.

The wide variety of theories related to research methods in psychometrics, combined with the changing technological landscape of LIS brings expanding subjectivity to both quantitative and qualitative inquiry; review of literature shows a need to reconcile the qualitative/quantitative dichotomy for development of practical applications.

This research provides background for development of assessment tools, such as instruments and rubrics related to human-computer-interaction (HCI) and usability studies. Such tools facilitate design and redesign of software packages used in LIS and interdisciplinary social sciences and related cost-benefit analyses.

**Keywords:** Library and Information Science, Human-Computer, Psychometrics, Research Methodology

Library and Information Science (LIS), a social science metadiscipline, occupies a practical/theoretical mid-ground where social sciences and Information Technology (IT) meet. Research in IT draws from the social sciences for the development of research databases, and the social sciences draws from IT for analyses of HCI, sociometrics, and related areas of psychometrics. LIS draws from both. This paper seeks to identify and flesh out areas in which psychometric measurement testing is well-suited to the research methods of LIS.

IT, the tools of which increasingly use common language search terms, is a participant in the growing practical demand for assessment that can combine empirical, quantitative analyses with meaningful and useful qualitative interpretations. Budd (2010), verifying need in the literature for postpositivist social science (including LIS) methods, offers a syntheses of critical realism and phenomenology by providing theoretical base for the ontological reality of aspects of perception. The need for psychometric quantification of information-seeking behavior research is particularly applicable to modern LIS and HCI, with wide-ranging implications. Groenen (2006), in his general discourse on the current state of psychometrics, notes the expanding and ever-broadening nature of psychometrics, and focuses on the dual interest—by academics, researchers in statistical software packages, and educational and social psychologists. That several disciplines are concerned with modeling of choice behavior—utilizing research areas of as test theory, item response theory, factor analysis, structural equation modeling, and multidimensional scaling—is of great interest to educational software marketing, which is itself contributes to sociotheoretical contexts.

Systematized bridging of psychometrics and LIS will bring clarity to the decades-old qualitative and quantitative debate. Practical problems associated with the debate are described in Hanson (2008), who asserts that the basis of separation is not supported by...
theoretical examination of the concepts of subjectivity/objectivity; that systematization and quantification adapt to generalization, and that sociopolitical traditions may lurk behind the split. Hanson suggests that the advancement of inquiry and method involves transcending the qualitative/quantitative issue, to move towards more pertinent issues of validity and causality. Movement toward such transcendence is shown in studies such as White (2007), which developed the pennant diagram—a visual ‘cloud’ display of word clusters that combines bibliometrics and psychometrics—using Sperber and Wilson’s (2001) relevance theory (RT)—in attempts to facilitate problems of perception and intent in information-seeking behavior. Such research will be enhanced by systematized, qualitative methodology, and is further supported by Madill (2008), who adds that qualitification will lend to coherence and further interactions across different methodologies.

This exploratory paper presents an exploratory model graphic (Fig.1) showing where need in LIS research to extrapolate both the qualitative and quantitative methodological rigor from psychology research is applicable to modern LIS and HCI. Important implications for technological advances, for example in research database design studies using predictive information-seeking behavior psychometrics, provide supportive justification.

**PRACTICAL/THEORETICAL RELEVANCE**

Psychometrics is rigorously applicable to assessment of library services. LIBQual (Heath, 2002) is a measurement tool of library services which can be administered in-tandem with assessment of user perceptions, satisfaction, and other patron-centered measurements. Multidimensional scaling is noted by McGrath (1984) as applicable to assessment of library circulation, collection management, funding, and online retrieval. Information Literacy assessment is of great relevance. Classical test theory can be applied to LIS, and is itself in need of regular assessment and updating. Mery & Newby (2011) found that test questions tailored for localized institutions were successful according to application of both classical test theory and item response theory to evaluate the validity and reliability of the Standardized Assessment of Information Literacy (SAILS). SAILS utilizes the Association of College Research Libraries (ACRL) *Information Competency Standards for Higher Education*, which is comprised of the following eight skills areas: developing a research strategy; selecting finding tools; searching;
using finding tool features; retrieving sources; evaluating sources; documenting sources; understanding economic, legal, and social issues.

Competencies, usability, and user perception—in application to information-seeking behavior—are foundations of LIS and information literacy, and are highly relevant areas for psychometric measurement. For example, Information literacy involves multi-dimensional information-seeking behaviors. For measuring multi-dimensional concepts such as practical knowledge, Meijer (2002), in attempting to increase testing instrument validity, applied the technique of multi-method triangulation. O’Brien (2010), in developing and evaluating a survey to measure user engagement, says that user experience is a vital factor in interactive system design. Reliability analysis and exploratory factor analysis may identify attributes of engagement, such as perceived usability, aesthetics, focused attention, felt involvement, novelty, and endurability. It was found that these attributes have complex interrelationships across user and systems. Rigorous usability and user perception cohort testing projects are in planning stages; the authors of this paper will use regression analysis and analysis of variance in the development of a research question leading to a hypothesis formulation testing stages. This research will result in a contribution to the systematization integration of psychometrics with LIS.

INFORMATION-SEEKING BEHAVIOR AND PSYCHOMETRICS

Information-seeking behavior is a fertile and widely-applicable ground for research and psychometric quantification in LIS. Kuhlthau (2004) created a six-stage theory—the Information Search Process (ISP)—which isolates stages an information seeker experiences, and notes accompanying feelings (apprehension, uncertainty, optimism, relief). Information professionals such as librarians formulate appropriate response strategies and also anticipate subsequent behaviors. These factors may be used in experimental psychological research measurement, particularly involving predictive behavior. These measurements will be of great interest to the designers of keyword-based research databases, where the fluid, multidimensional mental processes of an information search, if quantitatively measured, can be of great use in HCI experimentation. Of note is that nearly twenty years after the six-stage ISP model became disseminated into the literature, in the context of modern keyword-search based research databases, Kuhlthau’s stages have been tested and empirically shown to be still useful for LIS research. (Kuhlthau, 2008).

In the initial stages of a search, where awareness of a need for information to accomplish a task or goal may be accompanied by apprehension, perception surveys indicating dissatisfaction about source knowledge and search techniques will be lent perspective from systematized psychometric application. When assessing, for example, the effectiveness of a particular database, dissatisfaction may bear unique implications if seen as a stage in a process. User engagement, as asserted by O’Brien (2010), plays a major role in the exploration stage. User engagement is related to perception and user appraisal of information sources; all involve skill sets involving conscious and unconscious mental processes. Perceived usefulness—before, during, and after an information search—an important area for psychometric measurement, LIS, and assessment in every stage of an information search—is relevant to psychological experimentation in predictive behavior, and its measurement.

The exploration stage of information-seeking behavior is most readily associated with keyword searching in research databases. Tann & Sanderson (2009) find that search engine queries fall within categories including navigational, informational, and transactional. Their findings suggest that whereas in the past some queries would have been classified as informational, web searching has evolved to where they are now classified as navigational. It is surmised that modern large websites specializing in a particular type of information, such as Wikipedia and Facebook, contribute to such changes in searching characteristics. This attention to public websites and the development of skill sets outside of formalized learning environments—involving tracking data—is of great use to psychometrists; the use of information-seeking behavioral statistics—of primary interest to marketing research for decades, will benefit from the development of a systematized integration with LIS research. Relevance Theory (RT) may be applied to the study of human-computer dialogues to account for, for example, the human tendency to apply the least possible amount of effort in order to obtain a result. This is important for the designers and re-designers of software, if they are to be cognizant of human information-seeking processes. In keyword searches, randomness—mismatch irrelevancy and hit-and-miss—are cited as contributing factors to less-than-satisfactory HCI (White, 2007). These factors are in need of further research into evolving software, psychology, and HCI.
TEST VALIDITY, RELIABILITY AND LIS

As Mery & Newby (2011) found with the SAILS testing, validity and reliability are central concerns for LIS. There is a need for continual revisiting of the concept of validation, as technology evolves, and conditions change. Validity clarification—the establishment that a test measures what it is intended to measure—is a topic for continual revisiting. Kim (2009) in a study on psychometric research instrument validation finds a need for standards on validation processes and reporting, which will provide structural ballast for a noted increase in attentions to specific validation issues. Kim’s article offers ways for researchers to implement improved validity reporting. Tojib & Sugianto (2011) assert that research literature in information science suffers from a lack of agreement about definition of terms—in the case of their article, the terms are convergent and discriminant validity. This lack of agreement evidences the need for a broad-based, systematized integrated foundation of research methodology bridging psychometrics and LIS, whereas LIS is at the forefront of the human-computer interface and of information-seeking behavior, and whereas the measurement of skill sets for the use of technological tools for information searches inevitably bleed into each other.

Among the many aspects of validity, instrument construct validity—the primary form of validation underlying the trait-related approach to psychometrics—wherein the entity the test is measuring is normally not measurable directly, but is evaluated by examining the relationship between the test and the various phenomena that a theory predicts—is of great complexity in LIS and HCI, owing to the multidimensional nature of human neuroprocessing and computer processing of digital information.

All aspects of validity are readily applicable and useful for psychometric considerations of LIS, HCI, and information-seeking behavior. For example, face validity—the assumption of the validity of a test according to its appearance, content, or context—addresses the complicated assessment of internet and database keyword searching, whereby there are so many approaches to an information search. There are no defined right or wrong approaches, and if a test-taker has searched a certain way for certain kinds of information, if presented with choices outside of their experience, may assume a test is invalid. This can affect the outcome of a test, and is particularly problematic for HCI and psychometrics. Content validity is of relevance whereas the relationship between the content and the purpose of a test may be unclear regarding search techniques. A searcher naturally has more aptitude for familiar topics, and the level of information-seeking engagement changes during the process, depending on topic familiarity. Furthermore, there may be completely fluid and changing goals and tasks from one minute to the next. Criterion-related validity establishes the relationship between scores on a questionnaire and a criterion measure which should be indicative of something. For example, in an information search for a biology reference table, if a user is unable to identify source material, the test ought to show whether it was because of subject ignorance or database ignorance. Predictive validity and its measurement bear particularly important implications for database design. Database searching is a skill that develops according to different personal mental processes, and therefore the scores of a given assessment should accurately predict, for example, the potential for a patron to use a given database that is designed in a similar way to the one the patron was tested on. This should take into account how knowledge cross-pollinates over time and experience.

An estimate of the accuracy of a test, the concept of reliability is important to library and information science and, in particular, the assessment of article database and bibliographic management software. For example, if users get similar or very different results from taking tests in different circumstances needs to be noted. Several ways of measuring reliability—test-retest, parallel forms, and split half—are pertinent to library and information science, and are shown to involve aspects the measurement of which beg for systematized integration into the research methodology.

Test-retest, whereby respondents are given the same test, may be problematic for psychometrics in that not only the respondents of a given test may remember their responses on the second occasion (the common solution being a long interval between the two administrations of the test). In the case of LIS, and, for example, internet or database searching, such searching skills are developed in a variety of ways in everyday internet use, as well as use of a variety of databases. There would appear to be a continuum of skill-set development that transcends the concept of reliability in any traditional sense, in terms of Library and information science. This special consideration could be taken into consideration in a systematization of psychometrics applications to research methodology. Parallel forms, whereby two different tests are created with questions that are meant to be equivalent—for correlation—involve problems

Among the many aspects of validity, instrument construct validity—the primary form of validation underlying the trait-related approach to psychometrics—wherein the entity the test is measuring is normally not measurable directly, but is evaluated by examining the relationship between the test and the various phenomena that a theory predicts—is of great complexity in LIS and HCI, owing to the multidimensional nature of human neuroprocessing and computer processing of digital information.

All aspects of validity are readily applicable and useful for psychometric considerations of LIS, HCI, and information-seeking behavior. For example, face validity—the assumption of the validity of a test according to its appearance, content, or context—addresses the complicated assessment of internet and database keyword searching, whereby there are so many approaches to an information search. There are no defined right or wrong approaches, and if a test-taker has searched a certain way for certain kinds of information, if presented with choices outside of their experience, may assume a test is invalid. This can affect the outcome of a test, and is particularly problematic for HCI and psychometrics. Content validity is of relevance whereas the relationship between the content and the purpose of a test may be unclear regarding search techniques. A searcher naturally has more aptitude for familiar topics, and the level of information-seeking engagement changes during the process, depending on topic familiarity. Furthermore, there may be completely fluid and changing goals and tasks from one minute to the next. Criterion-related validity establishes the relationship between scores on a questionnaire and a criterion measure which should be indicative of something. For example, in an information search for a biology reference table, if a user is unable to identify source material, the test ought to show whether it was because of subject ignorance or database ignorance. Predictive validity and its measurement bear particularly important implications for database design. Database searching is a skill that develops according to different personal mental processes, and therefore the scores of a given assessment should accurately predict, for example, the potential for a patron to use a given database that is designed in a similar way to the one the patron was tested on. This should take into account how knowledge cross-pollinates over time and experience.

An estimate of the accuracy of a test, the concept of reliability is important to library and information science and, in particular, the assessment of article database and bibliographic management software. For example, if users get similar or very different results from taking tests in different circumstances needs to be noted. Several ways of measuring reliability—test-retest, parallel forms, and split half—are pertinent to library and information science, and are shown to involve aspects the measurement of which beg for systematized integration into the research methodology.

Test-retest, whereby respondents are given the same test, may be problematic for psychometrics in that not only the respondents of a given test may remember their responses on the second occasion (the common solution being a long interval between the two administrations of the test). In the case of LIS, and, for example, internet or database searching, such searching skills are developed in a variety of ways in everyday internet use, as well as use of a variety of databases. There would appear to be a continuum of skill-set development that transcends the concept of reliability in any traditional sense, in terms of Library and information science. This special consideration could be taken into consideration in a systematization of psychometrics applications to research methodology. Parallel forms, whereby two different tests are created with questions that are meant to be equivalent—for correlation—involve problems
whereby equivalencies in database keyword searching are difficult in that search processes are nonlinear and related to prior subject knowledge. Parallel form reliability applied to multidisciplinary keyword-search based databases seems naturally adaptable to LIS. Research methodologies in both psychology and LIS therefore would stand to benefit from an integrated, systematized supportive framework when conducting interdisciplinary information-seeking behavioral research.

Mery (2010), noted that SAILS testing involved the need to establish complicated criteria framework design/redesign, including external validation related to other similar tests, item reliability related to level of difficulty, individual and cohort-score correlation, and other editing of the instrument. This was achieved over a three-year period, following which the final skill sets (Developing a research strategy; selecting finding tools; searching; using finding tool features; retrieving sources; evaluating sources; documenting sources; understanding economic, legal, and social issues) were configured successfully to ensure that what was being was measured conformed to the construct of information literacy. This complex design/redesign is further indicative of the need for a systematized integration of psychometric and LIS/HCI research methodology.

SUMMARY, CONCLUSIONS, FUTURE RESEARCH

The literature review suggests a strong need for a systematized integration of psychometrics into the research methods of the field of Library and Information Science. Demonstrated affinities are evidence that they can be efficiently and effectively applied and integrated. Applied to usability studies in conjunction with fundamentals of information-seeking behavior and human-computer-interaction—both components of the field of library and information science, psychometrics applications will broaden, triangulate, and further highlight commonalities, and facilitate their integration.

The expansive volume of theoretical literature related to research methods in psychology—and the various ways in which psychometrics may be applied to them, combined with the changing landscape of LIS, bring broadening subjectivity—and a need for flexibility—to this inquiry. Whereas theoretical literature may tend to complicate and frustrate goals for objectivity and practical implications and applications, this is evidence of the need to anchor such related theoretical fields with practical controls.

The development of assessment instruments and tools—such as rubrics related to human-computer-interaction and usability studies—relies on practical systematization; to integrate psychometrics with LIS and HCI will contribute to the kinds of cost-benefit analyses research institutions and institutions of higher education are increasingly in need of in times of economic uncertainty. To plan for the future design and redesign of software packages and electronic research databases utilized in the field of Library and Information Science—in-tandem with psychometrics and other tools of the social sciences—is the driving force behind this project.

The authors are developing a theoretical research question leading to a hypothesis that will be rigorously tested with regression analysis and analysis of variance, using cohorts of primarily undergraduate students in an urban college environment. The results will be submitted for publication, with the goal of wide use of the hypothesis in the development of systematized psychometric application tool for LIS. The results will benefit the broadening interdisciplinary fields involving digital information and its processing, usability and competencies involved in its use and assessment—in theoretical, practical, technological, and psychological contexts.

What is happening inside the human mind during information-seeking behavior? Are there correlations to what is happening inside a computer? Can any such correlations be measured, analyzed, and put to practical use through the development of tools? Can those practical tools and their uses be assessed? Changing perceptions of Library and Information Science, human-computer interaction, and social sciences in general call for new applications of psychometrics that can facilitate integration across such related fields. This calls for new rigorous structural methodologies in research, which will result in practical development of widely-applicable assessment tools, and suggestions for the building of navigational tools for the information highways and byways ahead.
References


