Development, testing and validation of an information literacy test (ILT) for higher education

Bojana Boh Podgornik a*, Danica Dolničar a, Andrej Šorgo b, Tomaž Bartol c

aFaculty of Natural Sciences and Engineering, University of Ljubljana, Aškerčeva 12, 1000 Ljubljana, Slovenia, bojana.boh@ntf.uni-lj.si, danica.dolnicar@ntf.uni-lj.si

bFaculty of Natural Sciences and Mathematics / Faculty of Electrical Engineering and Computer Sciences, University of Maribor, Koroška cesta 160, 2000 Maribor, Slovenia, andrej.sorgo@um.si

cDepartment of Agronomy, Biotechnical Faculty, University of Ljubljana, Jamnikarjeva 101, 1000 Ljubljana, Slovenia, tomaz.bartol@bf.uni-lj.si

* Corresponding author bojana.boh@ntf.uni-lj.si, ++386 1 2003257

Abstract 1

A new information literacy test (ILT) for higher education was developed, tested and validated. The ILT contains 40 multiple-choice questions (available in the Appendix) with four possible answers and follows the recommendations of information literacy (IL) standards for higher education. It aims to assess different levels of thinking skills and is intended to be freely available to educators, librarians and higher education managers, as well as being applicable internationally for study programs in all scientific disciplines. Testing of the ILT was performed in a group of 536 university students. The overall test analysis confirmed the ILT reliability and discrimination power as appropriate (Cronbach alpha 0.74; Ferguson’s delta 0.97). The students’ average overall achievement was 66%, and IL increased with the year of study. The students were less successful in advanced database search strategies, which require a combination of knowledge, comprehension and logic, and in topics related to intellectual property and ethics. A group of 163 students who took a second ILT assessment after participating in an IL-specific study course achieved an average post-test score of 78.6%, implying an average IL increase of 13.1%, with most significant improvements in advanced search strategies (23.7%), and in intellectual property and ethics (12.8%).

Introduction

According to the Association of College and Research Libraries (ACRL), information literacy (IL) is defined as an intellectual framework for understanding, finding, evaluating, and using information (ACRL, 2000). For the last two decades, IL competencies and skills have been an important subject in the area of higher education, influencing the design, content, teaching methodology and management of academic courses. Correspondingly, attempts have been made to systematically develop and outline standards and criteria for the evaluation of IL skills. Recent developments in the field have been comprehensively presented at international

conferences on information literacy (Kurbanoglu et al., 2014). Proficiency in IL activities and skills may be accomplished in a combination of fluency with information and communication technology (ICT), investigative methods, logic, critical thinking, discernment and reasoning. The increasingly rapid development in the field of ICT, however, presents a real challenge for any long-term standardization of evaluation methods.

**IL standards in higher education**

A chronological sequence of main IL standards in higher education is presented in Table 1, comparing their structure and contents. One of the earliest attempts to define information skills, the Big Six (Eisenberg and Berkowitz, 1990), first presented in 1988, derives its structure from an earlier taxonomy of educational objectives. The authors envision six stages in solving an information problem, each divided into two further sub-stages.

In a later taxonomy, Bruce (1999) organizes IL into seven categories (Seven Faces of Information Literacy in Higher Education).

One of the more frequently cited models was proposed by the Association of College & Research Libraries, a division of the American Library Association (ACRL, 2000). In this complex model of information literacy competency standards for higher education, the recommended five IL standards are specified by 22 performance indicators and a range of detailed outcomes, which serve as guidelines for professors, librarians and curriculum developers in the preparation of study programs, course materials and various assessment instruments for measuring students’ learning progress in IL.

The ACRL standards were used as a foundation for IL standards by the Council of Australian University Librarians (CAUL) in which IL is presented by a model of seven standards, each divided into sections and subsections. These standards later served as the basis for a new edition, which was issued by the Australian and New Zealand Institute for Information Literacy – ANZIIL (Bundy, 2004). In this edition, where the number of standards is reduced to six, certain ambiguities were addressed and an attempt was made to place the standards in a broader context of generic skills.

In the United Kingdom and Ireland, similarly organized standards were presented in 1999 as “Seven Pillars of Information Skills” by the Society of College, National and University Libraries – SCONUL. In 2011, this model was updated and expanded to better reflect more recent concepts of information literacy (The SCONUL Seven Pillars of Information Literacy, 2011). Each pillar is described by a series of statements relating to a set of skills/competencies and a set of attitudes/understandings.

Another UK-based project in the field of information literacy was the Big Blue report (Joint Information Systems Committee, 2002), which was conceived to reflect the eight stages of the information-seeking process.

In 2008, UNESCO published a conceptual framework paper that provides a set of indicators of IL, defined as the transformation of information into knowledge (Catts & Lau, 2008).
Table 1: Comparison of the main IL standards in higher education

<table>
<thead>
<tr>
<th>BIG SIX – six stages by Eisenberg and Berkowitz (1990)</th>
<th>SEVEN FACES – seven categories by Bruce (1999)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: task definition</td>
<td>1: information technology conception</td>
</tr>
<tr>
<td>2: information-seeking strategies</td>
<td>2: information sources conception</td>
</tr>
<tr>
<td>3: location and access</td>
<td>3: information process conception</td>
</tr>
<tr>
<td>4: use of information</td>
<td>4: information control conception</td>
</tr>
<tr>
<td>5: synthesis</td>
<td>5: knowledge construction conception</td>
</tr>
<tr>
<td>6: evaluation</td>
<td>6: knowledge extension conception</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>ACRL/ALA five standards of IL (2000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: determines the nature and extent of the information needed</td>
</tr>
<tr>
<td>2: accesses needed information effectively and efficiently</td>
</tr>
<tr>
<td>3: evaluates information and its sources critically and incorporates selected information into his or her knowledge base and value system</td>
</tr>
<tr>
<td>4: uses information (individually or as a member of a group) effectively to accomplish a specific purpose</td>
</tr>
<tr>
<td>5: understands many of the economic, legal and social issues surrounding the use of information and accesses and uses information ethically and legally</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1: recognises the need for information and determines the nature and extent of the information needed</td>
</tr>
<tr>
<td>2: finds needed information effectively and efficiently</td>
</tr>
<tr>
<td>3: critically evaluates information and the information-seeking process</td>
</tr>
<tr>
<td>4: manages information collected or generated</td>
</tr>
<tr>
<td>5: applies prior and new information to construct new concepts or create new understandings</td>
</tr>
<tr>
<td>6: uses information with understanding and acknowledges cultural, ethical, economic, legal and social issues surrounding the use of information</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1: recognising an information need</td>
</tr>
<tr>
<td>2: addressing the information need</td>
</tr>
<tr>
<td>3: retrieving information</td>
</tr>
<tr>
<td>4: evaluating information</td>
</tr>
<tr>
<td>5: adapting information</td>
</tr>
<tr>
<td>6: organising information</td>
</tr>
<tr>
<td>7: communicating information</td>
</tr>
<tr>
<td>8: reviewing the process</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>UNESCO six skills (Catts &amp; Lau, 2008)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: definition and articulation of information need</td>
</tr>
<tr>
<td>2: location and access of information</td>
</tr>
<tr>
<td>3: assessment of information</td>
</tr>
<tr>
<td>4: organisation of information</td>
</tr>
<tr>
<td>5: use of information</td>
</tr>
<tr>
<td>6: communication and ethical use of information</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SEVEN PILLARS – seven skills by SCONUL (2011)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: identify</td>
</tr>
<tr>
<td>2: scope</td>
</tr>
<tr>
<td>3: plan</td>
</tr>
<tr>
<td>4: gather</td>
</tr>
<tr>
<td>5: evaluate</td>
</tr>
<tr>
<td>6: manage</td>
</tr>
<tr>
<td>7: present</td>
</tr>
</tbody>
</table>

Although the coverage of overall skills and outcomes in various IL standards is interrelated and similar, different proposals divide IL skills into distinctive sections with further subdivisions. The design of some schemes is occasionally updated in line with new developments in ICT. Furthermore, most standards seem to address similar issues, with different wording and definitions. The main themes are identified somewhat arbitrarily (between five and eight), sometimes subdivided into equally subjective subthemes. Some narrower topics are repeated several times in broader sections. Descriptions of tasks are sometimes needlessly complex and theoretical. Also, many tests based on selected standards, still place perhaps too strong importance on specialized library issues, for obvious reasons. The recent developments in scholarly search engines (the ubiquitous Google Scholar) accompanied with online reference managers and veritable "style generator" templates have rendered, insidiously, many library and information services and functionalities obsolete.
Many once powerful databases and catalogues have not survived. Thus, it has become obvious that the standards will need not only an overhaul or mere revision, but perhaps an entirely new conception will need to be thought out. However, the existing standards are what remain available at the present situation.

Lately, initiatives emerged to revise the current ACRL standards, in order to reflect the changing landscape of information overflowing with a variety of sources (Stewart & Basic, 2014). The ACRL Board approved a unanimous recommendation that the standards be significantly revised, and authorized the creation of a task force. The third draft of the Framework for Information Literacy for Higher Education was released in November 2014 and was promoted broadly, including within the higher education community. The final document was sent to the ACRL Board in January 2015 (ACRL, 2015).

Tests and questionnaires for assessing IL in higher education

Many articles report on IL assessment results in different contexts, presenting and using questionnaires and tests as research instruments. Most of them follow the recommendations and themes presented in various IL standards. The assessments are usually conducted by librarians, sometimes in cooperation with researchers or IL educators. The review of IL assessment tools in this subsection focuses primarily on the design and content of available IL tests and questionnaires.

O’Connor et al. (2002) presented different phases in the construction of a standardized ACRL-based questionnaire, the SAILS project. An internationally benchmarked study by Mittermeyer (2005), based on the same ACRL standards, identified a number of research skills, considered as essential. Multiple-choice questions were developed, mostly with four possible answers (one correct) and some additional options, such as “Other (please, specify)”, “Don’t know” or “None of the above”. A similar design was later applied by Salisbury and Karasmanis (2011).

The IL test by Ondrusek et al. (2005) included multiple-choice questions with four possible answers (one correct), as well as a number of open-ended and true-false questions. The content focused primarily on library-specific issues.

Thornton (2006) employed the SCONUL model in a questionnaire consisting mostly of open-ended questions, requiring the students to provide answers by themselves. Reed et al. (2007) composed an IL test with six possible answers (including “I don’t know”), designed as “Choose only one answer” or “Choose all/any that apply”. A short multiple-choice IL survey by Staley et al. (2010) tackled mostly library-related issues and involved five answers per question (including “Not sure”). McKinney et al. (2011) applied SCONUL’s standards in a test with five answers per question (one correct, including “Do not know”), in which some of the correct answers were conceived as optimum responses.

A comprehensive IL test by Mery et al. (2010), mainly addressing specific library-related issues, was designed as a multiple-choice instrument (four answers per question, one correct) employing SAILS as well as locally developed items.

A hands-on approach was adopted by Kingsley et al. (2011), who designed a content-specific test applicable to the needs of a particular discipline (biomedicine). The instrument evaluated the students’ ability to comprehend overall tasks and was not based on any specific IL standards.

An experiment in a non-native English speaking environment was presented by Al Aufi and Al Azri (2013). Their questionnaire was based on the Big Six model, with an emphasis on specific aspects of the Arabic academic environment.
A short five-answer test (including “I don’t know”), based on only the first two ACRL standards, was designed by Hsieh et al. (2013) and primarily focused on identification of and access to information. A test by Leichner et al. (2013), on the other hand, focused on ACRL standards two and three. This nationally standardized German multiple-choice test differs from most other tests in that it implements only three answers per question, of which one, two or all three can be correct. Gross and Latham (2012) used ACRL-based test to identify the student with below-proficient skills and found that such student often overestimated their abilities.

Some other related examples of research based on IL standards measure additional IL-related psychological parameters and aspects. In such assessments, Likert-type scales are frequently employed, for instance to determine the level of self-efficacy in finding, using and communicating information (Kurbanoglu et al., 2006). IL-related self-efficacy, motivation and self-perception (Pinto, 2011; Carr et al., 2011). Likert-type scales accompanied with open-ended questions are also used, for example in addressing self-efficacy and anxiety in using information resources (Booker et al., 2012).

At the end, we may also mention an alternative methodology by Oakleaf (2009) who proposes the construction of rubrics as a different measure of assessing IL rather than tests and questionnaires which are usually used as primary means of evaluation.

However, as various IL tests and questionnaires have been developed and applied in different environments, for different target groups, and within different educational systems and traditions, there seems to be no general final agreement on which methodology or which specific IL test format serves best to evaluate IL of students in higher education.

Regarding the contents, most IL tests and questionnaires have been designed according to the existing IL standards. Several tests focus only on selected topics in the standards, and omit important dimensions, such as scientific writing or ethical issues. In most IL assessment experiments, the tests are duly validated. The administration of tests may involve on-line access to e-questionnaires, e-mail, regular mail, completion in a class or as a live interview. In most cases, librarians have been strongly involved in IL assessment experiments, usually as promoters of IL tests. In fact, librarians have become increasingly involved in IL instructions and teaching as one of the important aspects of the library work in academic context which has been reviewed on the cases from the United States and the UK by Cox and Corall (2013). Professors and related faculty members have also collaborated with librarians in some IL tests. Many of the existing IL tests, however, have been somewhat 'library-biased', paying strong attention to items such as call numbers, OPACs, or implying prior familiarity with content-specific databases and information services. Also, many tests have been dedicated to a local use, for example a particular library or institution. Reported disadvantages of some existing IL tests, which have been employed in similar experiments, include also cost of standardized assessments, and challenges in tracking changes in skill levels of individual students (Fain, 2011).

Earlier user studies have shown that new technologies and an increasing access to a wide range of information will also bring to the fore issues of optimal use of information which will need to be addressed by the promotion of IL (Bawden et al., 2000). In addition, competent education for IL is especially important, given the constant revisions and re-examination of theoretical foundations of IL (Spiranec & Banek Zorica, 2010). Information technologies, tools and environments rapidly change. Some items in IL tests quickly lose relevance over time, and some new issues emerge that need to be included. IL tests, used for examination purposes in credit-evaluated courses, particularly need constant updates. We have thus preferred to construct an instrument that puts emphasis on items less likely to
change in relevance over a period of time - being aware of a comment by Walsh (2009) that it is important to decide how to balance the need for a test that is easy to administer with one that will truly assess the varied transferable information skills that IL implies.

An important impetus behind the decision to develop a new, freely available ILT instrument was the fact that, to the best of our knowledge, most established online IL tests for higher education have only been offered commercially (e.g., SAILS, the Madison Information Literacy Test, The iSkills™) or have been designed specifically to test a selected population of students (e.g. B-TILED, the IL Questionnaire by Mittermeyer & Quirion, and TRAILS), as specified in Table 2.

Table 2 : Examples of tests and questionnaires for assessing IL

<table>
<thead>
<tr>
<th>IL Test Name</th>
<th>Education Level</th>
<th>Description</th>
<th>Availability</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>SAILS - Standardized Assessment of Information</td>
<td>Higher education, general</td>
<td>ACRL-based; e-version; validity and reliability</td>
<td>Commercial; $5 per student for Cohort test, $6</td>
<td>(Kent State University, 2015a)</td>
</tr>
<tr>
<td>Literacy Skills</td>
<td></td>
<td>ensured</td>
<td>per student for Individual Scores test</td>
<td></td>
</tr>
<tr>
<td>Madison Information Literacy Test</td>
<td>Higher education, general</td>
<td>ACRL-based; e-version; multiple-choice test</td>
<td>Commercial; demo test questions available for</td>
<td>(Madison Assessment, 2014)</td>
</tr>
<tr>
<td>The iSkills™ Assessment</td>
<td>Higher education, general</td>
<td>ACRL-based; e-version; measures IL through 7</td>
<td>Commercial; price $18–20, depending on quantity</td>
<td>(Educational Testing Service, 2014)</td>
</tr>
<tr>
<td>B-TILED - Beile Test of Information Education</td>
<td>Higher education; for teacher education programmes</td>
<td>Specifically designed for students enrolled in</td>
<td>Freely available document</td>
<td>Beile O’Neil 2005, 2015</td>
</tr>
<tr>
<td>Information Literacy Questionnaire by</td>
<td>Higher education; specifically for incoming first-year students in</td>
<td>ACRL-based; paper form; 22 multiple-choice</td>
<td>Questions published within a study</td>
<td>(Mittermeyer, 2005)</td>
</tr>
<tr>
<td>Mittermeyer &amp; Quirion</td>
<td>Quebec</td>
<td>questions, measuring IL skills based on 3rd, 6th,</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRAILS - Tool for Real-Time Assessment of</td>
<td>Primary &amp; secondary education; for school librarians and teachers</td>
<td>Knowledge assessment with multiple-choice</td>
<td>Web-based system, available by log-in, at no</td>
<td>(Kent State University Libraries, 2015b)</td>
</tr>
<tr>
<td>Information Literacy Skills</td>
<td></td>
<td>questions, measuring IL skills based on 3rd, 6th,</td>
<td>charge</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>9th, and 12th grade pre-university standards</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The ILT questions under study have been developed by university professors that hold a basic degree in sciences, and, in addition, a post-graduate degree or specialization in information sciences and/or librarianship. Their publishing activity is transparently evaluated in the Slovenian Current Research Information System SICRIS, which links data both to the Web of Science and Scopus (Bartol et al, 2014). The ILT was designed by the intention to follow the recommendations of the standards in a balanced way, with examples that suit study programs in all scientific disciplines (for example, natural sciences, engineering and technologies, agriculture, medicine, as well as social sciences, and humanities) and facilitate assessment of lower and higher order thinking skills. Also, the test is freely available for further use. To this end, 40 questions are published in this article in full length (Appendix). The evaluation of the results based on this test and the respective questions is presented in the ensuing subsections.
Motivation and research objectives

The main motivation of our research was to develop an information literacy test (ILT) for higher education that:

1. follows the recommendations in the standards in a balanced way, with up-to-date examples;
2. is suited to higher education study programs in all disciplines and fields;
3. is broader in scope than some previous questionnaires, which are frequently strongly focused on specific library issues;
4. is applicable internationally, by not including national or other specific themes;
5. can accommodate and assess lower- and higher-order thinking skills, according to Bloom’s taxonomy of educational objectives;
6. is easy to use both in a paper form as well as in an e-environment;
7. is applicable for small and for large groups of students, and enables fast assessment of IL;
8. is freely accessible to all users, with no fees of other payments required;

In addition, the ILT should serve as a tool for university educators and managers to:

1. design and manage study courses to achieve a sustainable advancement of IL knowledge, skills and competences;
2. evaluate students’ IL prior to the introduction of IL-related university courses, so that more emphasis can be placed on topics where the IL of a specific group of students is lower and needs special attention;
3. verify the IL level after completing IL-related courses (the ILT may also be a constituent part of the examination test), in order to monitor the advancement of individual students, group programs and/or institutions;
4. perform and evaluate educational research, e.g., when testing new teaching and learning materials or educational methodologies related to IL.

ILT development process

ACRL standards remain the most frequently used and cited IL standards for higher education, and have also been employed as the basic guideline in the development of ILT in the current study. Most of the IL content, issues and topics have, however, also been addressed in other similar internationally accepted standards and the ILT therefore follows the general guidelines of most IL-related documents. The process of ILT development is presented schematically in Figure 1.

Setting up a team of experienced university educators (university professors) to design a collection of questions

A team of university educators, who are active researchers both in different fields of science as well as in information science and education for IL, prepared an initial collection of 80 questions, based on the performance indicators in the standards. In the further reassessment process, the questions were compared, verified, discussed and optimized.
A set of 40 questions was selected for the first version of the ILT. The main selection criteria were (1) diversity of content – covering all of the major topics defined by standards, (2) diversity of difficulty – covering lower and higher cognitive levels and thinking skills, (3) clarity and unambiguity, and (4) interesting general topics for diverse groups of university students. The first version of the ILT was preliminarily tested on a group of 45 students of science and technology. Based on the results, a second discussion and reassessment followed, which resulted in style and language optimization, final formatting and improved design. Printed and electronic versions of the ILT were made available for wider testing.
Design of the final version of the ILT

The present version of the ILT has been designed as a multiple-choice instrument composed of 40 questions, each offering a choice of four possible answers (only one correct response), allowing for simple and unambiguous data collection, both in a printed form or as an online assessment tool. The same format of four possible answers was applied successfully in an earlier shorter IL questionnaire, developed and used locally at several departments of the University of Ljubljana (Juvan et al., 2006). The comprehensive ILT presented in the present article covers essential concepts in the field of IL, uses examples that are based on generic comprehension of tasks, and does not involve concepts that would require specialist knowledge, such as the names of particular databases or details from specific library services. The content can thus easily be transposed between different languages and environments. The average time required by students to answer the questions is approximately 30 minutes. The ILT version presented in the present article (Appendix) is a translation of the Slovenian language text.

Grouping of the ILT questions into content-related subscales for statistical analyses

In terms of content, existing IL standards are structured according to different scales. Furthermore, the authors and educators who employ the standards frequently design the scales according to additional criteria based on their own practices and needs. Consequently, the ILT questions cover all five ACRL standards and subordinate performance indicators (Table 3, subscales A). In addition, the ILT items have been experimentally structured into two alternative IL content-related systems, assigned independently by two IL educators, both co-authors of the ILT (Table 3, subscales B and C).

Table 3: Distribution of the ILT questions according to three alternative content-related subscale groupings, with group descriptions: ACRL model, and two subscale models assigned independently by two IL educators/co-authors of the ILT

<table>
<thead>
<tr>
<th>Subscale Grouping A (ACRL standards)</th>
<th>Subscale Grouping B</th>
<th>Subscale Grouping C</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1 - Determine the extent of the information needed</td>
<td>B1 - Information sources and databases</td>
<td>C1 - Bibliographic resources</td>
</tr>
<tr>
<td>A2 - Access the needed information effectively and efficiently</td>
<td>B2 - Search strategies</td>
<td>C2 - Critical evaluation</td>
</tr>
<tr>
<td>A3 - Evaluate information and its sources critically, and incorporate selected information into one’s knowledge base</td>
<td>B3 - Intellectual property and ethics</td>
<td>C3 - Ethics</td>
</tr>
<tr>
<td>A4 - Use information effectively to accomplish a specific purpose</td>
<td>B4 - Heuristic methods and critical evaluation</td>
<td>C4 - Search strategy</td>
</tr>
<tr>
<td>A5 - Understand the economic, legal and social issues surrounding the use of information, and access and use information ethically and legally</td>
<td></td>
<td>C5 - Document structure</td>
</tr>
</tbody>
</table>

The main purpose of the subscales is to identify specific content areas of IL in which particular students have potentially achieved higher or lower scores than the average, and that would therefore need special attention in the design, management or methodological approaches of IL courses. The assignment of each individual ILT question to subscales A, B and C is presented in Table 4.
Table 4: ILT item assignment to subscales A, B and C (the ILT item no. corresponds to the consecutive numbers of the ILT questions; groupings A, B and C correspond to the categories defined in Table 3)

<table>
<thead>
<tr>
<th>ILT item no.</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grouping A</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Grouping B</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Grouping C</td>
<td>2</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>

ILT testing and validation methodology

Test groups and settings

The test groups were comprised of students from six faculties of the two main Slovenian universities (University of Ljubljana and University of Maribor), enrolled in study programmes of life sciences, health, technologies, and the education of teachers. The selection of students was based primarily, but not exclusively, on the enrolment of students in compulsory credit-evaluated courses that contained informatics. The IL-related content was designed and implemented by university professors, with problem-based examples from the domains of study.

The 536 students in test group 1 (pre-test group) completed the ILT before taking any IL-specific classes. In terms of education level, 11% of the students in the sample were college students, 78% were university undergraduates, and the remaining 11% were postgraduate students. A large part of the sample (41%) was made up of first-year students, followed by second- (25%) and third-year (16%) undergraduate students. The remaining 18% were postgraduates - year 4 or above.

The 163 students in test group 2 (post-test group) answered the ILT after participating in an IL-related course (1–3 credit points, i.e., 15–45 contact hours, depending on the study programme). The group consisted mainly of first- (53%) and second-year (36%) students of undergraduate programmes. The remaining participants (11%) were first-year postgraduate students.

Testing was conducted during the period from January 2014 to January 2015, using the electronic ILT version on the open access survey system 1ka (http://www.1ka.si/), or in a printed form. It took place in computer classrooms and was supervised by teaching stuff. The instructors of the students tested were aware of ILT contents, both in the pre-tests and post-tests groups. A unified introductory protocol was applied before testing, including clarification of the purpose, instructions, explanation of voluntary participation and anonymity, and expressing gratitude for the participation.

Statistical tools and analyses
The following statistical analyses were carried out for data analysis and validation of the ILT:
• Descriptive statistics of the test score (mean, standard deviation, sample variance and standard error).
• Item analysis: frequencies of individual answers and item difficulty (percentage of correct answers) were calculated;
• Reliability estimates: the Cronbach alpha reliability test was carried out for the entire test, and separately for each of the subscales, based on different IL standards. The leave-one-out method was also used to identify items with positive/negative contributions to the overall Cronbach alpha.
• Subscale analysis: three different groupings of items (subcales A, B and C) were compared and analysed. For each grouping and each group, mean score and average difficulty was calculated. Correlations between group scores and the total score were explored.
• Cluster analysis was attempted in order to identify groups of students with similar characteristics. The variables used were subscale scores according to the groupings B.
• Pre-test post-test correlation analysis: t-test: the paired two sample for means method was used to compare the results of a group of students before and after taking an IL-themed class. A similar test was performed on subscales B.
• Differences in ILT scores between groups of students based on study years were investigated via one factor ANOVA and two sample t-test, assuming unequal variances.

The Statistical Package for the Social Sciences (SPSS®) was used for the analyses.

Results and discussion

Overall ILT analysis

The basic statistics regarding the ILT score, the overall test reliability and the test’s discriminant power are presented in Table 5.

Table 5: Overall ILT score statistics (sample: pre-test group, N = 536)

<table>
<thead>
<tr>
<th>Analysis</th>
<th>Value (max = 40)</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>26.4</td>
<td>66.0%</td>
</tr>
<tr>
<td>Median</td>
<td>27</td>
<td>68%</td>
</tr>
<tr>
<td>Mode</td>
<td>29</td>
<td>73%</td>
</tr>
<tr>
<td>Span</td>
<td>32</td>
<td>80%</td>
</tr>
<tr>
<td>Minimum</td>
<td>8</td>
<td>20%</td>
</tr>
<tr>
<td>Maximum</td>
<td>39</td>
<td>98%</td>
</tr>
<tr>
<td>Variance</td>
<td>25.6</td>
<td></td>
</tr>
<tr>
<td>Standard deviation</td>
<td>5.1</td>
<td></td>
</tr>
<tr>
<td>Reliability (span/st. dev.)</td>
<td>6.3</td>
<td></td>
</tr>
<tr>
<td>Standard error (estimated)</td>
<td>2.6</td>
<td></td>
</tr>
<tr>
<td>Overall test reliability – Cronbach alpha</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>Discriminant power – Ferguson’s delta</td>
<td>0.97</td>
<td></td>
</tr>
</tbody>
</table>

The total ILT score was calculated as the sum of points awarded for correct answers (correct answer = 1, incorrect answer = 0; max. score = 40). Overall test reliability was measured by calculating the Cronbach alpha. In the case of binary data, this is equivalent to the Kuder
Richardson index KR-20. The value obtained was 0.74, which is regarded as sufficient, considering an acceptability limit of 0.7. The impact of individual items on the test’s Cronbach alpha was tested by applying the leave-one-out method. Seven items had a minor negative impact on the overall Cronbach alpha (leaving any one of the items out would marginally increase the alpha by a tenth of a point at most, but by less in most cases). Therefore, no items were omitted from the questionnaire. Ferguson’s delta was calculated to measure the test’s discriminant power. The calculated value of 0.97 was deemed satisfactory, as the lower limit of acceptability is 0.9. Based on the calculated values of Cronbach alpha (0.74) and Ferguson’s delta (0.97), the ILT overall test reliability and discriminant power were therefore proved as sufficient.

**ILT item analysis**

The score distribution for the 40 questions of the ILT is shown in Figure 2. A slight bias towards higher scores is evident, suggesting that students already possessed some IL knowledge and skills, gained in previous formal and informal education.

![ILT score distribution](image)

Figure 2: ILT score distribution (max. score = 40; sample: pre-test, N = 536)

Basic statistics were calculated for each individual ILT item. Due to the binary nature of the item scores (1 = correct, 0 = incorrect), the item mean reflects the item difficulty (Figure 3). According to the suitability criteria (Ding & Beichner 2009), the difficulty should typically be in the range 30–90%. In the ILT, 36 of the 40 questions meet this criterion, while four questions remain outside the limit. Question 7 (recognition of the document type in a bibliographic database record) could be regarded as too easy, and three questions (12, 20 and 36) as very difficult (assignment of broader search terms; understanding the meaning of a search query; understanding and correct application of author’s rights and ethical principles). However, in designing the ILT, the authors intentionally composed and included these items in order to provide a range of difficulty and to assess a range of lower- and higher-order thinking skills. Questions 7, 12, 20 and 36 have therefore been retained as a constituent part of the test.
Differences in ILT scores according to year of study

As in a study by Cameron et al. (2007), the ILT construct validity was tested via score comparison by the year of study, and by the pre-test and post-test. In our research, the 536 students who took the ILT prior to participating in an IL course (pre-test) were grouped by year of study. The results show that the group summary of ILT score means increased by the year of study (Table 6).

Table 6: Group summary of ILT scores by the year of study

<table>
<thead>
<tr>
<th>Group</th>
<th>No. of Students</th>
<th>Score Mean</th>
<th>Score Standard Deviation</th>
<th>Score Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 1</td>
<td>218</td>
<td>24.94</td>
<td>5.20</td>
<td>27.03</td>
</tr>
<tr>
<td>Year 2</td>
<td>134</td>
<td>26.32</td>
<td>4.45</td>
<td>19.80</td>
</tr>
<tr>
<td>Year 3</td>
<td>86</td>
<td>27.33</td>
<td>5.12</td>
<td>26.18</td>
</tr>
<tr>
<td>Year 4 or above</td>
<td>98</td>
<td>28.89</td>
<td>4.31</td>
<td>18.53</td>
</tr>
</tbody>
</table>

One-way ANOVA was then used to test the null hypothesis of differences between group means being due to chance. There was a significant effect of study year on the ILT score at the p<.05 level for the four groups \(F(3, 532) = 16.26, p = 3.98 \times 10^{-10}\); the null hypothesis was therefore rejected, signifying that at least one of the group means differs significantly from the others. Pairs of means were then compared via t-test analysis, assuming unequal variances, with the null hypothesis that the difference in means was due to chance. In the first pair of year 1 vs. year 2, the null hypothesis was rejected \((t (314) = -2.657, p = 0.008, d = 0.29)\), indicating that the two means were significantly different with the moderate effect size \(d\). Similar results were obtained for scores in year 3 vs. year 4 and above \((t (167) = -2.224, p = 0.028, d = 0.33)\), while a comparison of year 2 with year 3 resulted in the rejection of the null hypothesis \((t (163) = -1.494, p = 0.137)\), suggesting that the difference between year 2 in 3 was not statistically significant. The most significant improvement of IL therefore occurred at two stages: (1) between the 1st and 2nd year of study, when students had to acquire new IL skills and knowledge for studying in the new environment of higher education, and (2) in the final year of study, with an increased quantity of individual assignments and preparation of the graduation thesis. The improvement of IL in the pre-test group by study year also indicates that students improved their IL knowledge and competencies with time, throughout
the study process, even if they were not involved in a specifically designed, IL-targeted study course. In this context, various IL content was incorporated and practised within regular study courses of science and technology. At the same time, the results imply that the ILT served well as an instrument to detect and measure the improvement.

**ILT subscale analysis**

The questions of the ILT were classified according to three different grouping systems, reflecting the content and subject areas of IL (subscales A, B and C, presented in Tables 5 and 6). For each subscale, the number of items (i.e., questions in the ILT) and the average difficulty are presented (Table 7).

**Table 7: Subscale statistics for the groupings A, B and C: number of ILT items and difficulty of questions within the subscale**

<table>
<thead>
<tr>
<th>Subscale A</th>
<th>No. of Items</th>
<th>Difficulty (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>15</td>
<td>67</td>
</tr>
<tr>
<td>A2</td>
<td>10</td>
<td>59</td>
</tr>
<tr>
<td>A3</td>
<td>5</td>
<td>82</td>
</tr>
<tr>
<td>A4</td>
<td>3</td>
<td>79</td>
</tr>
<tr>
<td>A5</td>
<td>7</td>
<td>56</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subscale B</th>
<th>No. of Items</th>
<th>Difficulty (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1</td>
<td>9</td>
<td>75</td>
</tr>
<tr>
<td>B2</td>
<td>9</td>
<td>58</td>
</tr>
<tr>
<td>B3</td>
<td>10</td>
<td>58</td>
</tr>
<tr>
<td>B4</td>
<td>12</td>
<td>72</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subscale C</th>
<th>No. of Items</th>
<th>Difficulty (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>9</td>
<td>67</td>
</tr>
<tr>
<td>C2</td>
<td>8</td>
<td>78</td>
</tr>
<tr>
<td>C3</td>
<td>7</td>
<td>54</td>
</tr>
<tr>
<td>C4</td>
<td>8</td>
<td>59</td>
</tr>
<tr>
<td>C5</td>
<td>8</td>
<td>71</td>
</tr>
</tbody>
</table>

The results of the difficulty of the subscales illustrate that students had a fairly good knowledge and understanding regarding the critical evaluation of information and its sources (A3 = 82%; C2 = 78%). Students knew how to use information effectively in order to accomplish a specific purpose (A4 = 79%), and had a good knowledge of information sources and databases (B1 = 75%). However, they were less successful in advanced database search strategies (B2 = 58%, C4 = 59%, A2 = 59%), which require a combination of knowledge, comprehension and logic. The results exposed the greatest lack of knowledge and understanding in IL topics related to intellectual property and ethics (subscale difficulty levels B3 = 58%, A5 = 56%, C3 = 54%), suggesting that the ethical, legal, economic and social issues surrounding access to and use of information need to be emphasized in IL courses, with examples and discussions that not only provide factual knowledge, but also stimulate comprehension, critical thinking and application of knowledge in different situations. Although this finding is not surprising, it has not been clearly evidenced and exposed in previous works. Studies of the IL skills of first-year students by Salisbury & Karasmanis (2011), for instance, revealed a lack of understanding of journal article citations, peer reviewed articles and referencing, while in a study by Hufford (2011) the basic search abilities of students were found to be satisfactory, but the ability to develop sophisticated search strategies was weak or lacking, as was evaluation and critical thinking regarding the retrieved information (Crawford, 2007).

Inter-group correlations and group correlations with the overall score were also computed. Inter-group correlations were found to be low (below 0.5) in all three subscale groupings (A, B and C), while correlations with the overall score were around 0.7, with only a few exceptions (the lowest being 0.54). This suggests that the questions belonging to each of the subscales measured different IL knowledge and skills that were not in a direct linear
correlation. In contrast, a study by Timmers & Glas (2010) investigating student information-seeking behaviour found a strong correlation between three subscales related to applying search strategies, evaluating information and regulation activities.

In order to determine whether the subscales could reliably be used on their own to assess specific areas of IL independently, Cronbach alpha was calculated for each group of all three groupings. The alphas thus obtained were low, with the highest value below 0.55, and in some cases approached 0.2. Thus, none of the subscales could reliably be used independently and separately from the entire ILT questionnaire. In order to be reliable, the ILT must therefore be used as a whole. This finding was additionally confirmed by a cluster analysis with the Jaccard similarity index as a measure. The results revealed that there was no significant clustering between ILT items.

**ILT cluster analysis**

Cluster analysis was experimentally conducted on the subscale scores according to grouping B (4 groups). K-means clustering was used, resulting in 6 clusters, presented by their means in Table 8. It could be argued that there are six distinct groups of students with a combination of different IL knowledge and cognitive skills. For instance, cluster 6 represents the best students overall; they were the best in all four subscale topics. The students in clusters 1 and 2 achieved similar overall results (both close to the average), the former being significantly better at B2 – Search strategies, while the latter were better at B3 - Intellectual property and ethics and B4 - Heuristic methods and critical evaluation. Clusters 3, 4 and 5 contained samples with below average results. Students in clusters 3 and 4 achieved similar totals. In cluster 3 they were significantly better at B4, but much worse at B2 and B3. Cluster 5 comprised of samples with the lowest scores in all subscales. These results imply that IL is a complex system of multiple skills, knowledge, understanding, strategies, and abilities of critical thinking and application. Accordingly, students’ IL results differ depending on their individual characteristics, interests, and previous education.

<table>
<thead>
<tr>
<th>Cluster #</th>
<th>Cluster size</th>
<th>Mean B1</th>
<th>Mean B2</th>
<th>Mean B3</th>
<th>Mean B4</th>
<th>Overall mean score</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>152</td>
<td>7.4</td>
<td>6.8</td>
<td>7.4</td>
<td>10.3</td>
<td>31.9</td>
</tr>
<tr>
<td>1</td>
<td>118</td>
<td>7.0</td>
<td>6.1</td>
<td>4.9</td>
<td>8.8</td>
<td>26.8</td>
</tr>
<tr>
<td>2</td>
<td>126</td>
<td>6.8</td>
<td>4.0</td>
<td>6.3</td>
<td>9.3</td>
<td>26.4</td>
</tr>
<tr>
<td>4</td>
<td>67</td>
<td>5.6</td>
<td>5.0</td>
<td>5.0</td>
<td>5.6</td>
<td>21.2</td>
</tr>
<tr>
<td>3</td>
<td>62</td>
<td>6.5</td>
<td>3.0</td>
<td>3.7</td>
<td>7.2</td>
<td>20.4</td>
</tr>
<tr>
<td>5</td>
<td>11</td>
<td>3.3</td>
<td>2.4</td>
<td>1.8</td>
<td>4.4</td>
<td>11.9</td>
</tr>
<tr>
<td>Total</td>
<td>536</td>
<td>6.8</td>
<td>5.2</td>
<td>5.8</td>
<td>8.7</td>
<td>26.4</td>
</tr>
</tbody>
</table>

**ILT pre-test vs. post-test correlations**

In order to determine the influence of the IL-specific study subject, the results of 163 students who took both the pre-test (before the IL course) and the post-test (after completing the IL course) were analysed by the t-test, using the paired two sample for means method. The null hypothesis assumed there were no significant differences between the pre-test and post-test scores. The results of the t-test are given in Table 9. The pre-test mean score of was 26.2
(65.6%), and the post-test mean score 31.4 (78.6%), with an average difference of 5.2 (13.1%). The null hypothesis was rejected; therefore, it can be stated with a 95% certainty that the average post-test score was from 4.6 to 5.9 points higher than the pre-test score. This infers that, on average, the IL-related course improved the IL of students by 13.1%, as well as implying that the ILT served as an appropriate measuring tool to detect the difference. The calculated effect size was large (d = 1.48).

Table 9: T-test analysis of the total ILT score

<table>
<thead>
<tr>
<th>M Pre-Test</th>
<th>SD Pre-Test</th>
<th>M Post-Test</th>
<th>SD Post-Test</th>
<th>t(163)</th>
<th>p&lt;</th>
<th>d</th>
<th>Post-Test Mean (%)</th>
<th>Mean Difference (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>26.2</td>
<td>4.95</td>
<td>31.4</td>
<td>3.87</td>
<td>-16.1</td>
<td>1.4E-35</td>
<td>1.48</td>
<td>78.6</td>
<td>13.1</td>
</tr>
</tbody>
</table>

A pre-test and post-test correlation analysis was also undertaken for the subscales of all three item groupings (A, B, C, as defined in Tables 5 and 6). All of the differences were statistically significant; effect sizes, mean differences and confidence intervals as % of improvement are given in Tables 10, 11 and 12.

In subscales A (Table 10), the results of the t-test show a significant improvement in the post-test for all five subscales A. The effect size was large in all of the subscales except for A3 and A5, in which it was moderate. The most significant improvement occurred in subscales A4 and A2, implying that students improved their ability to use information effectively to accomplish a specific purpose, and to access the needed information effectively and efficiently. Understanding of the economic, legal and social issues surrounding the use of information (subscale A5) also improved, but to a lesser extent.
Table 10: T-test analysis of ILT scores, with differences in pre-test and post-test scores by subscales A

<table>
<thead>
<tr>
<th>Subscale A</th>
<th>Pre-Test</th>
<th>Post-Test</th>
<th>t(56)</th>
<th>p&lt;</th>
<th>d</th>
<th>Mean Difference (%)</th>
<th>Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1 - Determine the extent of information needed</td>
<td>10.3</td>
<td>2.21</td>
<td>12.0</td>
<td>1.63</td>
<td>-10.6</td>
<td>2E-20</td>
<td>11.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13.3</td>
</tr>
<tr>
<td>A2 - Access the needed information effectively and efficiently</td>
<td>5.6</td>
<td>1.71</td>
<td>7.3</td>
<td>1.50</td>
<td>-11.3</td>
<td>4E-22</td>
<td>17.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>14.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20.2</td>
</tr>
<tr>
<td>A3 - Evaluate information and its sources critically, and incorporate selected information into one’s knowledge base</td>
<td>4.2</td>
<td>1.10</td>
<td>4.6</td>
<td>0.79</td>
<td>-5.3</td>
<td>4E-07</td>
<td>8.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>11.5</td>
</tr>
<tr>
<td>A4 - Use information effectively to accomplish a specific purpose</td>
<td>2.3</td>
<td>0.77</td>
<td>2.8</td>
<td>0.44</td>
<td>-8.7</td>
<td>3E-15</td>
<td>17.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>13.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>21.3</td>
</tr>
<tr>
<td>A5 - Understand the economic, legal and social issues surrounding the use of information, and access and use information ethically and legally</td>
<td>3.8</td>
<td>1.26</td>
<td>4.7</td>
<td>1.25</td>
<td>-8.0</td>
<td>3E-13</td>
<td>12.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>9.5</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15.7</td>
</tr>
</tbody>
</table>

In the alternative subscale system B (Table 11), the results of the t-test show a significant improvement in the post-test for all four subscales. The effect size was large in all but the first subscale, in which it was moderate. The greatest progress (23.7%) was demonstrated in the subscale B2 Search Strategies (scores increased by 20–27%), indicating an improvement in students’ ability to perform advanced searches in professional databases, using search strategies that are different from simple Google searching. In intellectual property and ethics (subscale B3), an average improvement of 12.8% was achieved; although this may be considered to be an encouraging result, it still leaves room for further improvement. These findings also indicate that students substantially improved their knowledge and skills in topics where their pre-knowledge was lower, while less progress was observed in topics where their pre-knowledge was fairly good (i.e., in B1 Information Sources and Databases).
Table 11: T-test analysis of ILT scores, with differences in pre-test and post-test scores by subscales B

<table>
<thead>
<tr>
<th>Subscale B</th>
<th>Pre-Test</th>
<th>Post-Test</th>
<th>t(56)</th>
<th>p&lt;</th>
<th>d</th>
<th>Mean Difference (%)</th>
<th>Lower Limit (%)</th>
<th>Upper Limit (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>B1 - Information Sources and Databases</td>
<td>6.8</td>
<td>7.4</td>
<td>1.34</td>
<td>1.21</td>
<td>-5.6</td>
<td>1E-07</td>
<td>0.5</td>
<td>6.6</td>
</tr>
<tr>
<td>B2 - Search Strategies</td>
<td>4.9</td>
<td>7.0</td>
<td>1.67</td>
<td>1.24</td>
<td>-14.4</td>
<td>1E-30</td>
<td>1.3</td>
<td>23.7</td>
</tr>
<tr>
<td>B3 - Intellectual Property and Ethics</td>
<td>5.8</td>
<td>7.0</td>
<td>1.78</td>
<td>1.44</td>
<td>-9.6</td>
<td>2E-17</td>
<td>0.9</td>
<td>12.8</td>
</tr>
<tr>
<td>B4 - Heuristic Methods and Critical Evaluation</td>
<td>8.7</td>
<td>9.9</td>
<td>1.95</td>
<td>1.52</td>
<td>-8.2</td>
<td>7E-14</td>
<td>0.8</td>
<td>10.2</td>
</tr>
</tbody>
</table>

In the alternative subscale system C (Table 12), the results of the t-test show a significant improvement in the post-test for all five subscales C. The effect size was moderate in scales C1–C3, and large in scales C4 and C5. The greatest improvement of mean difference (20.4%) was evident in search strategies (subscale C4).

Table 12: T-test analysis of ILT scores, with differences in pre-test and post-test scores by subscales C

<table>
<thead>
<tr>
<th>Subscale C</th>
<th>Pre-Test</th>
<th>Post-Test</th>
<th>t(56)</th>
<th>p&lt;</th>
<th>d</th>
<th>Mean Difference (%)</th>
<th>Lower Limit (%)</th>
<th>Upper Limit (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1 - Bibliographic Resources</td>
<td>6.1</td>
<td>6.9</td>
<td>1.47</td>
<td>1.26</td>
<td>-6.6</td>
<td>5E-10</td>
<td>0.6</td>
<td>6.2</td>
</tr>
<tr>
<td>C2 - Critical Evaluation</td>
<td>6.3</td>
<td>6.9</td>
<td>1.30</td>
<td>1.03</td>
<td>-5.6</td>
<td>9E-08</td>
<td>0.5</td>
<td>4.7</td>
</tr>
<tr>
<td>C3 - Ethics</td>
<td>3.8</td>
<td>4.5</td>
<td>1.41</td>
<td>1.30</td>
<td>-6.4</td>
<td>2E-09</td>
<td>0.6</td>
<td>6.7</td>
</tr>
<tr>
<td>C4 - Search Strategy</td>
<td>4.4</td>
<td>6.3</td>
<td>1.60</td>
<td>1.08</td>
<td>-13.8</td>
<td>5E-29</td>
<td>1.3</td>
<td>20.4</td>
</tr>
<tr>
<td>C5 - Document Structure</td>
<td>5.6</td>
<td>6.9</td>
<td>1.55</td>
<td>1.18</td>
<td>-10.0</td>
<td>1E-18</td>
<td>0.9</td>
<td>12.6</td>
</tr>
</tbody>
</table>
Conclusions

A new information literacy test (ILT) for higher education has been developed, tested on groups of university students, and validated by statistical methods.

The ILT comprises 40 multiple-choice questions with four possible answers, one being correct. In terms of content, the test follows recommendations of IL standards for higher education, particularly the ACRL/ALA five standards, and assesses a range of lower- and higher-order thinking skills, providing different levels of difficulty.

Testing of the ILT in a group of 536 university students confirmed the overall reliability and discrimination power of the ILT (Cronbach alpha 0.74, Ferguson’s delta 0.97). The ILT construct validity was confirmed via score comparison by year of study, and by pre-test and post-test.

In the initial experimental testing of the ILT, involving 536 students with no prior IL-specific study course, the average overall achievement was 66%. The results followed the Gaussian distribution, slightly biased towards higher scores.

The calculated difficulty of questions was in the range of 30–60%, with one question being easier and three questions being particularly difficult.

The IL of students progressed by the year of study, with the most significant improvements evidenced in the first year and the final year of study.

Subscale statistics revealed that students were less successful in advanced database search strategies, which require a combination of knowledge, comprehension and logic, and in topics related to intellectual property and ethics.

The results of cluster analysis suggested that IL involved different sets of skills that combine knowledge, understanding and strategies.

The group of 163 students who also took the ILT assessment after participating in an IL-specific study course achieved an average IL post-test score of 78.6%, and improved their result by an average of 13.1%. The most significant improvements were evidenced in advanced search strategies. Understanding of the economic, legal and social issues surrounding the use of information improved to a lesser extent.

The results therefore suggest that ethical, legal, economic and social issues, including the topics of authors’ rights and industrial intellectual property, need to be more strongly emphasized in IL-related courses, preferably with examples that stimulate comprehension, critical thinking, application of knowledge and problem solving.

The ILT has been designed as a general IL assessment tool for higher education, constructed in a form of a multiple choice test with 40 questions. From this structure derives its main strength – fast and simple evaluation of IL in various groups of students, as well as its main expected limitation: ILT is not meant to evaluate the IL skills related to the specific study fields. In this respect, for deeper evaluation of detailed IL contents, the educators may want to adapt some examples to more specific concepts which are in use in respective fields of science, and add additional open ended questions, examples, or problem solving tasks.
Being a freely accessible and easy to use IL assessment tool, the ILT may find its applicability in diverse areas of higher education and research. Its potential users are university educators, researchers, librarians and higher education managers, undertaking tasks and projects involving:

(a) evaluation of IL of freshmen, and/or of students prior to the introduction of specific IL-related university courses – to adapt or modify accordingly the IL study contents and teaching methodologies;

(b) identification of topics where the IL of a specific group or generation of students is weaker – to tailor the IL instruction according to the needs of specific students’ groups or programs;

(c) design and management of new study courses – to identify areas of IL-related knowledge and understanding that need greater or lesser attention;

(d) evaluation of study courses - to modify/improve of the IL-related contents and methodologies applied;

(e) preparation of IL-assessment tests - to monitor achievements and advancements of individual students, groups, programs and/or institutions.

(f) Last but not least, the ILT can serve as a research tool in educational research, e.g., when developing, testing and evaluating new teaching and learning materials and techniques, strategies, or educational methodologies related to IL.

For instance, in the research project entitled “Development of information literacy of university students as a support for solving authentic science problems”, the ILT has been used by the authors as one of the measuring instruments to evaluate methodologies and teaching styles, such as the traditional instruction, project-based learning, and problem-base learning in relation to the IL instruction. The results will be presented in our future publications.

Acknowledgements

The study was financially supported by the Slovenian Research Agency, project J5-5535 entitled “Development of information literacy of university students as a support for solving authentic science problems”.

The authors would like to thank all of the students who participated in the survey, as well as the university professors/project members who participated in testing the ILT: Saša Aleksej Glažar, Vesna Ferk Savec, Mojca Juriševič, Irena Sajovic and Margareta Vrtačnik from the University of Ljubljana; Alenka Baggia, Mirjana Kljajić Borštnar and Andreja Pucihar from the University of Maribor; and Blaž Rodič from the Faculty of Information Studies in Novo mesto, who also transferred the ILT to a web-based e-form on https://www.1ka.si/. The authors would also like to thank Neville Hall for proofreading the manuscript and the ILT.

References


Appendix

Online Information Literacy Test
Note: correct answers are formatted in bold.

1) The most reliable, verified, concise and comprehensive description of an unknown specialized concept can be found in:
   a) daily newspaper
   b) bilingual dictionary
   c) lexicon or encyclopedia
   d) research article

2) The most manageable and precise level of search criteria that include an object (keyword 1) and aspect (keyword 2) will be retrieved by the search sequence:
   a) keyword 1
   b) keyword 1 AND keyword 2
   c) keyword 1 NOT keyword 2
   d) keyword 1 OR keyword 2

3) If I have difficulty selecting the correct specialized English term when searching for information, I use:
   a) Google translate
   b) only the established native-language terms with which I have become acquainted during lectures
   c) a specialized thematic dictionary
   d) a general bilingual dictionary

4) In my assignment, I wanted to describe the impact of human activities on climate change. My initial search strategy returned an overwhelming number of documents. How do I proceed?
   a) I abandon the topic and ask for a completely different assignment.
   b) I define a more specialized theme within the topic, optimize the search strategy and proceed further.
   c) I look up the topic of climate change on Wikipedia and summarize this information in my assignment.
   d) In the faculty library, I look for a related article written by a well-known author and rework the content of that article.

5) An MSc or PhD thesis requires an original scientific contribution by the student. How do I proceed?
   a) I collect the most interesting recent publications and use them as the basis for my thesis.
   b) I look for experiments in research articles published by other authors and describe these experiments.
   c) I formulate new information and conclusions by combining both my own research results and the existing information.
   d) I collect and discuss conclusions from any available research article, book, patent and Web document.

6) In which list have the information sources been correctly ordered from the least to the most formally established and verified?
   a) blog, daily newspaper, scholarly journal, standard
   b) blog, standard, daily newspaper, scholarly journal
   c) daily newspaper, blog, standard, scholarly journal
   d) standard, scholarly journal, blog, daily newspaper

Look at this record from the bibliographic/catalogue database and answer questions 7–10.
7) The record in this database refers to:
   a) newspaper article
   b) specialized book
   c) video film
   d) scientific journal

8) This information source was issued:
   a) as a self-published book by the author Barbara Zemljic
   b) in 2008 in Ljubljana by the publisher Umanotera
   c) in 2008 in partnership with www.planet-sprememb.si
   d) in the current year on the Webpage www.planet-sprememb.si

9) Who is the author?
   a) The author is www.planet-sprememb.si
   b) There are two authors: Barbara Zemljic and Lucka Kajfez-Bogataj.
   c) There are three authors: Zemljic, Kajfez and Bogataj.
   d) This is a general Webpage with no specific known authors.

10) This information source is best defined as:
    a) documentary movie on ecology and the environmental protection of our planet
    b) science fiction DVD
    c) book with colour photographs and an associated soundtrack on human evolution
    d) Slovenian translation of an English TV series on the future of the planet Umanotera

11) I am investigating the impact of diet and nutrition on human health. The most relevant information will be found in information sources for:
    a) medicine and agriculture
    b) medicine and social sciences
    c) medicine and humanities
    d) medicine and sport

12) Which of the data listed below are “raw” unprocessed data:
    a) share prices published at the end of a trading day
    b) weather maps
    c) population growth data presented in tables
    d) population growth data presented diagrammatically (in graphs)

13) Original scientific articles typically describe:
    a) experience and perspectives acquired during the author’s years of professional activity
    b) a summary of other authors’ research
    c) an overview of the development of a scientific field
    d) the author’s original research results
14) I am exploring two-dimensional animations. Using the keyword ‘animation’, I have retrieved 33,314 documents in a database. Which of the queries listed below is the most appropriate for the next search?
   a) animation AND (2D OR 2-dimension* OR two dimension* OR two-dimension*)
   b) animation AND 2D AND 2-dimension* AND two dimension* AND two-dimension*
   c) animation NOT (2D OR 2-dimension* OR two dimension* OR two-dimension*)
   d) animation OR 2D OR 2-dimension* OR two dimension* OR two-dimension*

15) I want to find information on the medicinal plant oregano, which is also known as wild marjoram in traditional herbal medicine. Its scientific name is Oreganum vulgare. What is the most appropriate search query in a database?
   a) “oregano wild marjoram Oreganum vulgare”
   b) oregano AND wild marjoram AND Oreganum vulgare
   c) oregano OR wild marjoram AND Oreganum vulgare
   d) oregano OR wild marjoram OR Oreganum vulgare

16) I am interested in the topic of sweetening and sweeteners, and I find the appropriate English terms: sweet, sweeten, sweetener, sweeteners, sweetening. What is the most appropriate search strategy?
   a) right-hand truncation, using the term sweet*
   b) an exact search, in this case: “sweet sweeten sweetener sweeteners sweetening”
   c) searching with parenthesis: (sweet sweeten sweetener sweeteners sweetening)
   d) using the operator AND, i.e.: sweet AND sweeten AND sweetener AND sweeteners AND sweetening

17) In Google Scholar, “Find articles with all of the words” is equivalent to the search operator:
   a) AND
   b) AND NOT
   c) NOT
   d) OR

18) In Google Scholar, “Find articles with at least one of the words” is equivalent to the search operator:
   a) AND
   b) AND NOT
   c) NOT
   d) OR

19) How would you formulate a standard search query using the Google Scholar search criteria presented as:

   
<table>
<thead>
<tr>
<th>with all of the words</th>
<th>weather</th>
</tr>
</thead>
<tbody>
<tr>
<td>with the exact phrase</td>
<td></td>
</tr>
<tr>
<td>with at least one of the words</td>
<td>data information</td>
</tr>
</tbody>
</table>

   a) (data OR information) AND weather
   b) (weather AND data AND information)
   c) (weather OR data OR information)
   d) weather AND “data information”

20) A database search interface employs pull-down menus instead of search operators. Which of the Boolean operators substitutes the concept ‘optional’?
21) Compared to a search within the title and abstract, a full-text search in a database results in:
   a) the same number of hits
   b) a smaller number of hits
   c) this has no effect on the number of hits
   d) a larger number of hits

22) In order to obtain original research results regarding the behavior of users in relation to a new technology, it is best to employ:
   a) survey questionnaires and interviews
   b) patents
   c) review articles
   d) technical handbooks

23) What is the most appropriate method for organizing information in an electronic format:
   a) I read the documents in an electronic format, underline the most interesting parts, logically rename the files and assign them to folders according to the subject.
   b) I open a new folder and move the files to the folder using the original file names.
   c) I print out all of the documents in their entirety, read them, and then copy all of the interesting sections directly into my paper.
   d) I open a new file in a word processor and then copy-paste the relevant sections of the document directly into the file. I do not save the complete original documents.

24) I am writing a paper and want to cite findings from other articles. Which tab is used for this purpose in MS Word?
   a) References – Citations & Bibliography
   b) References – Footnotes
   c) Review – Tracking
   d) Review – Comments

25) I need to check the content of a large number of articles in a short time. Which element of an article can I examine quickly?
   a) abstract
   b) materials and methods
   c) discussion
   d) results

Which statement on GMO (Genetically Modified Organisms) is not the author’s personal opinion?
   a) GMO will bring about a global food crisis.
   b) According to inventories, 15 new GMOs were registered in the EU in 2013.
   c) GMO experimentation should be banned.
   d) Most GMO researchers have been paid off by large corporations, such as Monsanto.

27) It has been scientifically established that cholesterol is present in animal organisms but not in plants. How would you best describe a TV commercial which claims that the sunflower oil manufactured by a particular producer contains no cholesterol?
   a) This is a valuable benefit, and it will encourage me to buy this brand of oil.
   b) This is manipulative and misleading information, as plant oils do not contain cholesterol.
   c) This information has medical significance, and I am therefore willing to pay more for this oil.
   d) This is interesting information on the unique composition of this oil.
28) On my blog, I would like to publish a picture of a famous person who is seeking to advance humanitarian principles in his/her own country. However, his/her activities are prohibited in that country. How will I proceed?
a) I will not publish the picture because pictures of the person are banned in his/her own country.
b) I will not publish the picture because our two countries have friendly relations.
c) I cannot publish the picture of the person without her/his permission.
d) I will publish the picture because the international activities of the person are public and are based on universal ethical principles.

29) What is the correct sequence of the elements in a research article?
a) Abstract-Bibliography-Introduction-Material and Methods-Results-Discussion-Conclusions
b) Abstract -Introduction-Material and Methods-Results-Discussion-Conclusions-Bibliography
c) Abstract-Conclusions-Introduction- Bibliography-Material and Methods-Results-Discussion
d) Introduction-Results-Discussion-Conclusions-Material and Methods-Bibliography-Abstract

30) Mary Brown needs to create a password in order to access an information system. Which password is the most secure?
a) ma@r$y3br7OWN_  
b) MaryBrown  
c) MaryBrown123  
d) marybrown28111991

31) After an extensive information search, I learn that natural dyes are used in the production of jelly, candy, ice cream and yogurt; in the dyeing of cotton, wool and silk; and are added to makeup products and hair dyes. How should I best classify these products?
a) pharmacy, biology, food technology  
b) medicine, biology, chemistry  
c) nutrition, textile technology, cosmetics  
d) confectionery, animal husbandry, hairdressing

32) Which of these schemes is the most appropriate for presenting the topics from Question 31?

33) In which document type are citations and a bibliography not obligatory?
a) B.Sc. thesis  
b) scientific paper published in conference proceedings  
c) original scientific article  
d) general interest article

34) What is the typical length of an abstract in scientific articles?
a) 150 to 250 words  
b) 2000 to 3000 words  
c) 50 to 100 words
35) Which of these sections is not a standard part in a scientific article?
a) Materials and Methods  
b) Discussion  
c) Introduction  
d) Acknowledgments

36) I bought some old documents in a second-hand bookshop. Which of the documents can I scan and publish on my Webpage without authorization?
a) anonymous photo published in a women’s magazine  
b) article from a daily newspaper  
c) original manuscript by William Shakespeare  
d) translation of a poem written by a living poet and published by a British publisher

37) In my paper, I want to use some data from an article by another author. How do I proceed according to ethical principles and the protection of author’s rights?
a) I am allowed to make reasonable use of the data as long as I cite the source article.  
b) I can only use the data if I quote the source text word-for-word and cite the source article.  
c) I can only use the data if I obtain written permission from the author.  
d) Under no circumstances can I use the data.

38) Our university subscribes to a journal with a pay-per-license agreement. What am I not allowed to do?
a) Cite an article in my B.Sc. thesis.  
b) Print out an article on my printer.  
c) Download a full article on my computer.  
d) Scan a selected page and publish it on my blog.

39) What is the appropriate procedure for referencing other works in my written assignment or thesis?
a) I only reference the author of a picture from the Web if the picture has been supplied with a copyright sign © 
b) I do not have to reference information from the Web, as such information is freely available and does not have a © sign.  
c) I only need to reference the parts of a document that I quote word-for-word.  
d) I have to reference all of the information that is not a result of my own work.

40) If I refer to the citations in my text with numbering using the format [1], how do I structure the final list of references?
a) in alphabetical order by authors’ last names  
b) in chronological order by year of publication  
c) in the order of library accession numbers  
d) in ascending numerical order with regard to the first reference to the source in the paper