Article

Information literacy as a measurable construct: A need for more freely available, validated and wide ranging instruments

Helena Hollis, Academic and Reader Services Coordinator & Academic Liaison Librarian for Psychology and film, Media & Performance, Regent's University London.
Email: hollish@regents.ac.uk

Abstract

Though definitions of Information Literacy (IL) vary, there is consensus in the literature that it is one coherent variable or measurable construct so important it has been called a human right (Sturges & Gastinger, 2010). This paper reviews existing IL measures, treating them as psychometric tests, aiming to address two questions: do existing IL tests sufficiently meet the needs of researchers and what do the tests tell us about whether or not there truly is a single variable identifiable as IL?

Existing instruments for testing IL will be assessed. Only validated, freely available, testing instruments are considered. All of the current IL tests of this type are context bound to education, and are often domain-specific. These are only useful for researchers looking to assess these particular populations. However, if IL is a human right it ought to be a construct that exists across all human populations. The context and domain specificity of existing IL tests is therefore surprising; it implies that IL is something specific to higher education, not a factor in the population at large.

This paper therefore argues for a need to develop validated testing instruments for IL that are context and domain independent and made freely available. The review of IL tests will benefit researchers looking to investigate IL. The broader argument of this paper aims to be of interest to a wider library and information science audience seeking to understand whether IL is indeed a coherent singular construct.

Keywords

construct validity; higher education; information literacy; information literacy measures; psychometric testing

_____

1. Introduction

Information Literacy (IL) has been argued to be of prime importance across many contexts, from school through to the workplace (Bruce, 1999; Bucher, 2010). In fact, it has been deemed so important that it is considered by many to be a human right (Sturges & Gastinger, 2010). It is therefore interesting for researchers and information professionals to consider whether IL is in fact one underlying variable manifested across all of these contexts, such that we could truly consider it a human right, or whether IL in school may be something different to IL in the workplace or any other context, and thus if it would be more accurate to talk about many different information *literacies*. There is a wealth of literature debating and defining IL, and the

question of whether it is a singular or plural concept could be discussed through the lens of various theories. However, the question can also be approached from an empirical standpoint. This paper takes IL measures and treats them as psychometric tests in order to evaluate them, and thereby identify the nature of what they are measuring. The aim is to gain a better insight into what we are empirically testing when we attempt to measure IL.

There are different ways to frame a definition of IL, and definitions can change over time, but they are consistent in describing IL as a combination of knowledge and ability; the knowledge of when and what kind of information is needed, with abilities such as finding, evaluating and using that information (CILIP, 2004, 2018b). Knowledge about information and the abilities pertaining to it can be tested. This leads to the question of whether testing participants' information-related knowledge and abilities is testing a series of disparate individual elements, or if it is testing one underlying measurable construct that we can identify as IL.

If IL is a singular construct, then it is a common underlying determinant that can manifest itself in many different observable instances. For instance, being effective at locating information for school work, thinking critically about information in the media and writing illuminating blog posts could be three entirely unrelated abilities, or they may be manifestations of a construct that underpins them all – IL. Messick (1974) writes 'this problem of developing evidence to support an inferential leap from an observed consistency to a construct that accounts for that consistency is a generic concern of all science' (p.1). In library and information science, IL is discussed as a singular construct. However, the same scientific evidence to support the leap from observations of consistency to dealing with it as a construct is required. It is therefore valuable to look at measures of IL and evaluate whether or not responses to them demonstrate construct validity, thus giving evidence for the claim that these responses are being determined by the participants' underlying IL.

2. Exclusion criteria for measures

This paper aims to provide a comprehensive review of existing IL measures. However, some measures will be excluded from further analysis based on three considerations; test validity, self-report and cost.

2.1 Validity

The validity of a test refers to its consistency and accuracy in measuring what it purports to measure. Messick (1979) highlights this with the simple question 'is the test any good as a measure of the characteristics it is interpreted to assess?' (p.1). Construct validity, as discussed above, can be broken down into distinct ways of measuring how good a test is as a measure of the characteristics it is interpreted to assess. Face validity is the subjective, common-sense relation of the test to what it should measure. Criterion validity is a measure of how comparable a test is to others measuring the same construct, and how predictive it is of outcomes associated with that construct (Suen, 2009; Walsh & Betz, 1995). Strict standards for validity in educational and psychometric testing exist (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014), and holding IL measures to the same standards is essential if they are to have the same value. Therefore only fully validated tests are considered.

2.2 Self-report

There are many self-report measures of IL available, in which participants are asked to assess their own knowledge and abilities. These can be useful for researchers looking to understand how their participants feel about their IL, but are not suitable for actually measuring IL. Asking participants to self-report assumes they will have an accurate perception of their own abilities, and will correctly remember what they did well at or struggled with; these assumptions are not borne out by wider research (Tourangeau, 2009).

Additionally, the Dunning-Kruger effect demonstrates that participants with very low ability often have overinflated self-perceptions (Kruger & Dunning, 1999). Kruger and Dunning comment on the fact that certain domains require a baseline understanding of the domain in order to be able to evaluate one's performance in it. For example, if someone has little knowledge of grammar they may well rate themselves as excellent at it because they do not have enough understanding of the complexities of grammatical rules in order to accurately assess their own usage. It seems likely that participants with very low IL may believe themselves to be highly information literate, in the same way Kruger and Dunning observed with other subject areas. To mirror the grammar example, if someone believes that their preferred newspaper is veracious and habitually takes their information from it, and is simply naïve of issues surrounding press integrity, they may rate themselves as very good at selecting appropriate sources and yet this person would likely perform poorly at a test of source evaluation. Kruger and Dunning (1999) postulated that this effect can also run the other way, where one's high levels of knowledge in a subject area lead them to underestimate their ability. In the source-assessment case, a participant who is highly knowledgeable about the complexities of journalism may be reluctant to rate themselves as excellent at evaluating sources precisely because they understand how fraught the field can be. Thus the uncritical news consumer may score very highly, while the cautious reader scores lower, on a self-report IL question.

A case of these issues with self-evaluation can be seen in an interview study in Canadian libraries, which found that patrons reported themselves as being highly capable, while library staff reported them as being lacking in capability in internet usage (Julien & Hoffman, 2008). A self-report instrument would be of no use in determining which group was more accurate in their assessment; a testing measure would be needed. It is therefore important to separate these two kinds of tests, and be clear about what kinds of conclusions can be drawn from them. Self-report may be a useful tool in assessing participants' feelings about their IL, but should not be used as a measure of IL.

2.3 Cost

Finally, measures that require payment will be excluded from further consideration in this analysis. As library budgets come under increasing strain (CIPFA, 2017; Nicholas, Rowlands, Jubb, & Jamali, 2010), few will have access to funding for costly research projects. Furthermore, if we are to treat IL as a human right, we should aim to measure it in all populations and not merely where funding is greatest.

3. Evaluation of existing measures

3.1 Method

Tests of IL were found through systematic literature searching. In a comprehensive review of IL research tools, Catalano (2016) provides a useful overview of existing IL measures that was used as a starting point. Beyond this, papers discussing the development and validation of IL tests were identified. Where the full test was not provided requests were sent to authors. Papers discussing the findings of IL testing were also reviewed to identify which tests they used, and these were then searched for and/or requested. Tests of skills that can be seen as analogous to IL, but not using IL terminology, were also researched (for instance "information evaluation", "information skills" etc.). Searching was conducted in English only, and therefore IL tests in other languages are not included.

The identified tests were evaluated based on a series of criteria that will be indicated in the following section. Firstly, the rationale behind the development of the test was read and evaluated for the dependencies inherent in the test's purpose. Then the full test was evaluated with a thorough categorisation of all test items. The reviewer used the specified criteria shown in the tables in the following section as a checklist for each test, and checked off test items against

each criterion. It should be noted that a single reviewer made these evaluations, and this is a limitation of the research.

3.2 Evaluation

Table 1 below shows a breakdown of all freely available, validated testing measures of IL, noting their context and domain dependencies. What is immediately visible is that there is no general measure of IL; all existing measures are specific to context and some also to domain.

Table 1: Dependencies in IL tests

Measure	Source	Context Dependency	Domain Dependency
B-TILED	(Beile O'Neil, 2005) (Jesse, 2012; Robertson, 2018)	HE	Education; Seminary; Nursing
IL Test for Chemistry Students	(Emmett & Emde, 2007)	HE: Postgraduate	Chemistry
Information Evaluation Pre- and Post- Test	(Catalano, 2015)	HE	No
Information Literacy (Psychology)	(Leichner, Peter, Mayer, & Krampen, 2013)	HE	Psychology
Information Literacy Survey	(Ferguson, Neely, & Sullivan, 2006)	HE	No
Information Literacy Test for Higher Education	(Boh Podgornik, Dolničar, Šorgo, & Bartol, 2016)	HE	No
Information Search Tasks	(Leichner, Peter, Mayer, & Krampen, 2014)	HE	Psychology
Information Skills Survey	(Clark & Catts, 2007)	HE	Law; Social Sciences; Medicine
Locally Developed IL Test	(Mery, Newby, Peng, Bowler, & MacMillan, 2013)	HE	No
PIKE-P	(Rosman, Mayer, & Krampen, 2015)	HE	Psychology
Project Trails	(Schloman & Gedeon, 2007)	High School (USA)	No
VOILA	(Ondrusek, Dent, Bonadie-Joseph, & Williams, 2013)	HE	No

It is clear that the primary context of IL measurement is higher education (HE). This is not surprising, given that research into IL is most likely to be instigated by those working in research intuitions. Furthermore, IL is increasingly taught in HE and recognised as an important factor in student outcomes (Dold, 2013) and so there is a need to measure the outcomes of that teaching and the impact of IL in student journeys. There is therefore a clearer motivation for testing IL in HE than in other contexts. Beyond HE, Project Trails is designed to test IL in high schools where similar motivators to test the effectiveness of instruction, and a similarly explicit understanding of the need for IL exist (Bucher, 2010). At the time of research, no freely available test of IL goes beyond educational contexts. It is interesting to note that iSkills was developed to test IL and later adapted for assessing ICT-specific IL skills in the workplace (Katz, 2007). However this measure required payment and it is no longer available.

Domain dependency is less prevalent in IL tests than context dependency. Where tests are domain-dependent, it is to a field of academic study within HE and this is often motivated by specific IL instruction incorporated into programmes of study. This is a logical approach where instructors are interested in measuring their students' knowledge of subject-specific information resources. However, it renders the test unusable outside of its domain.

Methodologically, the tests do not differ drastically. Almost all IL tests rely upon multiple choice questions. Almost all IL tests rely upon multiple choice questions. Information Search Tasks/PIKE-P are notable exceptions, using open-ended tasks with scoring rubrics. This test therefore requires more effort in assessment, but arguably provides a more naturalistic assessment of IL (Leichner, Peter, Mayer, & Krampen, 2014). Catalano's (2015) Information Evaluation test uses hypothetical research situations to give the test more realism, but has multiple choice answers for ease of assessment. VOILA uses a virtual library tour and video instruction, followed by multiple choice questions. Additionally, some measures such as B-TILED and Information Literacy Survey combine self-report and testing questions, giving additional data beyond an IL score.

The Information Skills Survey and Project Trails tests are available in full only upon request. At the time of writing, neither had been provided following requests and therefore these two measures are omitted in the following in-depth analysis.

Some tests include questions specific to their own institutional libraries. For example, tests may ask about specific resources such as databases, journals or systems the library subscribes to. Furthermore, tests can include questions relating to specific classification schemes, which will exclude participants accustomed to other schemes. Library policy on areas such as inter-library loans, fines, access etc. can be included in test items. Some questions can also be country-specific, either in terms of regulations or culture. Therefore, the wider application of many of these tests is further limited beyond the restrictions of domain and context dependency. Table 2 provides a breakdown of these limitations in the discussed IL measures.

This analysis demonstrates a lack of generalisability in most IL measures. It highlights how the primary purpose of many IL tests is in-house assessment of specific student populations, not the general measurement of IL more widely. Researchers can select a test that matches the specificities of their own institution, for instance Information Literacy Test for Higher Education if they are located in the EU and use endnotes for referencing. However, Information Search Tasks is the only test that could be used to draw comparisons between the IL scores of students at different institutions regardless of their resources, policies and geographic location, but only those studying psychology. There is no truly general IL measure for all of HE. Beyond their in-built specificities, it is also interesting to investigate the extent to which these measures test all the aspects of IL, that is how well the construct is represented in them. Construct under-representation, where a test only covers a limited subset of the characteristics of what it is supposed to measure, undermines its validity (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014). For the purposes of this analysis, the construct of IL was defined using the existing CILIP definition of IL at the time of analysis (CILIP, 2004). However, there has since been a newly released definition (CILIP, 2018a) which came out between the time of evaluation and the publication of this paper. Upon review, the new definition has not been found to radically diverge from the previous version so as to change how the construct of IL should be construed. In this evaluation IL is taken to be 'knowing when and why you need information, where to find it, and how to evaluate, use and communicate it in an ethical manner' (CILIP, 2004).

Table 2: Further specificities in IL tests

Measure	Classification scheme	Resources	Library Policy	Referencing style	Country specificities
B-TILED	No	Yes	No	No	Yes: US copyright regulations
IL Test for Chemistry Students	Yes	Yes	Yes	No	Yes: US copyright regulations
Information Evaluation Pre- and Post- Test	No	No	No	No	Yes: US based examples
Information Literacy (Psychology)	No	Yes	Yes	No	No
Information Literacy Survey	No	No	No	No	Yes: US news sources; US copyright law
Information Literacy Test for Higher Education	No	No	No	Yes (endnotes)	Yes: EU copyright regulations
Information Search Tasks	No	No	No	No	No
Information Skills Survey	Full test not available				
Locally Developed IL Test	No	Yes	Yes	Yes (MLA)	No
PIKE-P	No	Yes	No	No	No
Project Trails	Full test not available				
VOILA	Yes	Yes	Yes	No	No

This definition has been broken down into its constituent parts, encompassing knowledge and ability, with ethics falling in between. Table 3 shows to what extent the construct is represented in each IL measure.

Table 3: Construct representation

IL aspects	Knowledge			Ability	
	when/why you need information	*where to find it*	*information ethics*	*how to evaluate information*	*how to use and communicate it*
B-TILED	Yes	Yes	Yes	Yes	No
IL Test for Chemistry Students	Yes	Yes	Yes	Yes	No
Information Evaluation Pre- and Post- Test	Yes	No	No	Yes	No
Information Literacy (Psychology)	Yes	Yes	No	Yes	No
Information Literacy Survey	Yes	Yes	Yes	Yes	Yes
Information Literacy Test for Higher Education	Yes	Yes	Yes	Yes	Yes
Information Search Tasks	Yes	Yes	No	Yes	No
Information Skills Survey	Full test not available				
Locally Developed IL Test	Yes	Yes	Yes	Yes	No
PIKE-P	Yes	Yes	No	Yes	No
Project Trails	Full test not available				
VOILA	No	Yes	No	No	No

In this analysis, the referencing of sources is considered to be part of information ethics (though it could be arguably a form of communication). The decision was made due to the primary focus of referencing in HE being the avoidance of plagiarism.

From the measures available at the time of analysis, only two fully represent the construct of IL; Information Literacy Survey and Information Literacy Test for Higher Education.

Using and communicating information is the most under-represented facet of IL. This is likely due to the prevalence of multiple choice design, as it is arguably harder to phrase multiple choice questions to capture usage and communication of information.

It is important to note that not all of these measures were designed as complete IL tests; VOILA only assesses the finding of information as it is designed to accompany a virtual library tour; Information Evaluation Pre- and Post- Test is intended as an information evaluation measure, though it does also capture some considerations of when and why information is needed. Therefore, this assessment is not a criticism of these individual measures, it is merely identifying an under-representation of IL as a construct across the breadth of existing measures.

4. Recommendations

This analysis of existing freely available, validated testing measures of IL has identified a lack of domain and context independent measures that could be used in a general population. In fact, all current measures are limited to educational contexts and so allow us to measure IL only in participants currently within educational institutions. If IL is to be treated as relevant outside of education, it should surely be measurable outside of this context.

Context and domain dependency are not necessarily negatives. It is highly useful, for instance, to measure psychology students' IL within their subject area, testing their knowledge of psychology-specific resources and their ability to use that information in a way that meets the conventions of publishing in psychology. This is especially pertinent if the results are to be used to improve IL instruction for those students. Similarly, it would be useful to measure IL in other contexts where IL instruction could be designed to meet the needs of a specific population, such as in prisons, schools, businesses and so on. Therefore creating IL measures for these contexts would be valuable. However, the main recommendation of this paper is that a general measure of IL, neither context nor domain dependant, would be valuable as it would enable researchers to make comparisons between populations.

The analysis of existing measures also provides some useful insights into the design of IL measures, their validation, and their uses, from which recommendations can be formulated for the creation of future IL measures.

4.1 Instrument Design

Table 2 shows how tests can have inbuilt specificities beyond their context and domain dependencies. A general IL measure should avoid inclusion of test items that are dependent upon any of these forms of specific knowledge. Though knowledge is part of IL (CILIP, 2018b), it should be a wide-ranging form of knowledge that is tested, for example knowledge of how to use a library catalogue in general, rather than knowledge of a specific classification scheme. Multiple choice questions are used in almost all existing IL measures. However, it can be argued that multiple choice is only testing participants' knowledge and not actual ability, and thus only measures half of how we define IL (CILIP, 2018b). Liechner et al.'s (2014) Information Search Tasks stands out as an alternative to multiple-choice, providing open questions that require online information finding, along with a scoring rubric. This requires participants to actively engage in information tasks, utilising both their knowledge and abilities. However, this then creates a dramatically increased workload in scoring the participants' responses. Thus this design is not recommended for a general measure of IL, or for measures of IL in contexts where librarians have limited research time allocated. Catalano's (2015) Information Evaluation test uses multiple choice, making it easier to assess, but presents questions as hypothetical situations. Though this is not directly testing ability, it does perhaps encourage participants to engage with the questions *as if* they were tapping into their abilities. There is evidence that imagined scenarios produce emotional responses very close to real situations (Robinson & Clore, 2001) and that the same neural circuits are involved in imagining, performing and observing actions (Decety & Grèzes, 2006; Jeannerod, 2001). This implies that having participants imagine performing a task may engage similar neural circuitry to them actually performing the task, and therefore may be a more efficacious form of multiple choice testing.

4.2 Validation

Many of the existing IL measures (VOILA, B-TILED, IL Test for Higher Education) are based upon the ACRL standards (Association of College and Research Libraries, 2000), and their face validity has been examined with relation to those standards. These, or similar standards provide a good basis for face validity testing. New measures for further contexts should have their face validity assessed by librarians and users in those contexts. For a general measure of IL, ideally a wide range of library and information practitioners should be consulted, from across as many contexts as possible including public libraries, prisons, schools, universities and other

organisations. Furthermore, it would be interesting to see what participants from outside of the library and information professions would judge the tests to be measuring. If we are judging whether or not IL is in fact a measurable construct, the face validity of IL should be investigated with appeal to those who do not already know what it is supposed to be.

Additionally, any new IL measures should have their validity statistically tested to rigorous standards just as educational and psychometric tests do (American Educational Research Association et al., 2014). As information professionals, it is important to lead by example and to use the most robust possible methodologies to gather information and draw conclusions from it. New IL tests can be validated in comparison to those examined in this paper. However, this may not be adequate given that all existing tests are context and domain dependent. Further validation is needed against measures from outside of HE. An example of this can be seen in the iSkills test, which was validated in comparison to business writing (Katz, Haras, & Blaszczynski, 2010). Similar measurable skills or outcomes reasonably assumed to be related to IL should be used in the validation of a general IL measure.

4.3 Uses

A general test of IL would be useful even within the HE context. Currently, IL measures assess students' knowledge and ability that is directly tied to their academic work, for instance their knowledge of how to correctly reference a source in an essay. However, when these IL scores are then compared to academic outcomes, there is a risk of tautology. It is not informative that a student who performs well on a HE IL measure also attains good grades, if the HE IL measure is largely testing the same things that are required learning outcomes in order to obtain the good grades. Therefore finding these correlated is less meaningful than if the IL measure did not cover the same ground as the students' learning outcomes. It would be interesting to see if the way students understand and use information in their everyday life outside of university would correlate with academic outcomes.

Beyond this, a general measure of IL would allow research into IL levels beyond educational contexts where it is relatively lacking. Crucially, it would permit for the comparison between different populations. A wealth of research questions could be explored, such as:
- Are there regional differences in IL?
- Do these correlate with local library provision?
- Do prison inmates have lower IL levels than the general population?
- Are there differences between prisons with more or less IL training provision?
- Do businesses with more information literate employees perform better?
- Do university students and vocational apprentices have different IL levels?

All of these questions come with the potential for identifying needs and then being able to present a stronger case for increased resourcing to fulfil those needs.

IL has been argued to be of critical importance for a wide of range of outcomes. It has been argued that IL is essential for democracy and social power (Cope, 2018) and that therefore public libraries must include it as part of their mission (Hall, 2010). With a test of IL usable in the general population, comparisons could be made between IL scores and other variables of importance to democracy and social power, such as levels of bias, political engagement, socio-economic status and even whether IL correlates with voter registration. If such correlations were identified, the findings would support the claims of Cope, Hall and others in the societal importance of IL.

Finally, the development, validation and use of a general IL measure would strengthen the claim that IL is a single measurable construct. As things stand, we can argue from the empirical evidence that there are such a things as 'IL in HE', and 'IL in High School', which could merely

be seen as student skills. However, IL beyond those two kinds remains untested through validated instruments.

5. Conclusion

This study has its limitations; only English language measures have been evaluated; a single assessor judged test items against the selected criteria; other IL tests may exist that were not located. However, there is an overall picture that emerges from investigating IL measures of HE specificity. IL has been argued to be of vital importance across contexts and in practically all domains, so much so that it is considered by some to be a human right. There is a wealth of theoretical research that argues for IL as one identifiable factor. There are also empirical observations pointing to an observed consistency in individuals' knowledge and abilities pertaining to information. However, there is a need to prove whether IL is truly a singular construct that accounts for that consistency in a measurable sense. If IL is a singular construct, it should be possible to design general measures of IL that can be validated. Currently, the only IL measures are context and often domain dependent; all existing measures are for educational contexts. This means that we can currently only conclude that 'IL in education', possibly just a set of student skills, is a measurable construct. Furthermore, it leaves researchers outside of educational contexts without freely available, validated measures to use in investigating IL.

Therefore additional measures of IL are needed. Measures for contexts outside of education would be valuable. Critically, there is currently no general measure of IL that could be used across populations, allowing for comparisons to be made. Such a measure should be developed by library and information professionals from as wide a range of sectors as possible, should aim to contain as little specificity as possible and should be validated to rigorous standards. The creation of such a measure would open the possibility of richer IL research, identifying needs with greater clarity and leading to more meaningful recommendations and actions in IL provision.

References

American Educational Research Association, American Psychological Association, & National Councilon Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Association of College and Research Libraries. (2000). Information literacy competency standards for higher education. Available at: https://alair.ala.org/handle/11213/7668 [Accessed 10 February 2018]

Beile O'Neil, P. (2005). *Development and validation of the Beile test of information literacy for education (B-TILED).* (Doctoral thesis), University of Central Florida. Available at: http://purl.fcla.edu/fcla/etd/CFE0000749 [Accessed 10 February 2018]

Boh Podgornik, B., Dolničar, D., Šorgo, A., & Bartol, T. (2016). Development, testing, and validation of an information literacy test (ILT) for higher education. *Journal of the Association for Information Science and Technology, 67(10)*, 2420–2436. https://doi.org/10.1002/asi.23586

Bruce, C. S. (1999). Workplace experiences of information literacy. *International Journal of Information Management, 19(1)*, 33–47. https://doi.org/10.1016/S0268-4012(98)00045-0

Bucher, K. T. (2010). The Importance of Information Literacy Skills in the Middle School Curriculum. *The Clearing House, 73(4)*, 217–220. https://doi.org/10.1080/00098650009600955

Catalano, A. (2015). The effect of a situated learning environment in a distance education information literacy course. *The Journal of Academic Librarianship, 41(5)*, 653-659. https://doi.org/ 10.1016/j.acalib.2015.06.008

Catalano, A. (2016). *Streamlining LIS research: A compendium of tried and true tests, measurements, and other instruments*. Santa Barbara, CA: Libraries Unlimited.

CILIP. (2004). Information Literacy. Available at: https://archive.cilip.org.uk/research/topics/information-literacy [Accessed 10 February 2018]

CILIP. (2018a). ILG launch new CILIP information literacy definition. Available at: https://infolit.org.uk/new-il-definition [Accessed 10 February 2018]

CILIP. (2018b). What is information literacy?. Available at: https://infolit.org.uk/ [Accessed 10 February 2018]

CIPFA. (2017). CIPFA's annual library survey for 2016/17. Available at: http://www.cipfa.org/about-cipfa/press-office/archived-press-releases/2017-press-releases/spending-on-public-libraries-falls-by-%C2%A366m-in-a-year?crdm=0 [Accessed 10 February 2018]

Clark, C., & Catts, R. (2007). Information Skills Survey: Its application to a medical course. *Evidence Based Library and Information Practice, 2(3)*, 3–26. https://doi.org/10.18438/B81G65

Cope, J. (2018). Information literacy and social power. In M. T. Accardi, E. Drabinski, & A. Kumbier (Eds.), *Critical library instruction: Theories and methods* (pp.13–27). Sacramento, CA: Library Juice Press.

Decety, J., & Grèzes, J. (2006). The power of simulation: Imagining one's own and other's behavior. *Brain Research, 1079(1),* 4–14. https://doi.org/10.1016/j.brainres.2005.12.115

Dold, C. J. (2013). The role of librarians in academic success. *Journal of Systemics, Cybernetics and Informatics, 11(2)*, 1–5.

Emmett, A., & Emde, J. (2007). Assessing information literacy skills using the ACRL standards as a guide. *Reference Services Review, 35(2)*, 210–229. https://doi.org/10.1108/00907320710749146

Ferguson, J. E., Neely, T. Y., & Sullivan, K. (2006). A baseline information literacy assessment of biology students. *Reference & User Services Quarterly, 46(2)*, 61–71. https://doi.org/10.5860/rusq.46n2.61

Hall, R. (2010). Public praxis: A vision for critical information literacy in public libraries. *Public Library Quarterly, 29(2)*, 162–175. https://doi.org/10.1080/01616841003776383

Jeannerod, M. (2001). Neural simulation of sction: A unifying mechanism for motor cognition. *NeuroImage, 14(1)*, S103-S109. https://doi.org/10.1006/nimg.2001.0832

Jesse, M. (2012). Subject specific information literacy curriculum and assessment. *The Christian Librarian, 55(1)*, 2–16.

Julien, H., & Hoffman, C. (2008). Information literacy training in Canada's public libraries. *The Library Quarterly, 78(1)*, 19–41. https://doi.org/10.1086/523908

Katz, I. R. (2007). Testing information literacy in digital environments: ETS's iSkills assessment. *Information technology and Libraries, 26(3)*, 3–12. https://doi.org/10.6017/ital.v26i3.3271

Katz, I. R., Haras, C., & Blaszczynski, C. (2010). Does business writing require information literacy? *Business and Professional Communication Quarterly, 73(2)*, 135–144. https://doi.org/10.1177/1080569910365892

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77(6)*, 1121–1134. https://doi.org/10.1037/0022-3514.77.6.1121

Leichner, N., Peter, J., Mayer, A.-K., & Krampen, G. (2013). Assessing information literacy among German psychology students. *Reference Services Review, 41(4)*, 660–674. https:doi.org/10.1108/RSR-11-2012-0076

Leichner, N., Peter, J., Mayer, A.-K., & Krampen, G. (2014). Assessing information literacy programmes using information search tasks. *Journal of Information Literacy, 8(1)*, 3–20. https://doi.org/10.11645/8.1.1870

Mery, Y., Newby, J., Peng, K., Bowler, M., & MacMillan, M. (2013). Assessing the reliability and validity of locally developed information literacy test items. *Reference Services Review*, *39(1)*, 98–122. https://doi.org/10.1108/00907321111108141

Messick, S. (1974). The standard problem: Meaning and values in measurement and evaluation. *ETS Research Report Series, 1974(2)*, i–37. https://doi.org/10.1002/j.2333-8504.1974.tb01034.x

Messick, S. (1979). Test validity and the ethics of assessment. *ETS Research Report Series, 1979(1)*, i–43. https://doi.org/10.1002/j.2333-8504.1979.tb01178.x

Nicholas, D., Rowlands, I., Jubb, M., & Jamali, H. R. (2010). The impact of the economic downturn on libraries: With special reference to university libraries. *The Journal of Academic Librarianship, 36(5)*, 376–382. https://doi.org/10.1016/j.acalib.2010.06.001

Ondrusek, A., Dent, V. F., Bonadie-Joseph, I., & Williams, C. (2013). A longitudinal study of the development and evaluation of an information literacy test. *Reference Services Review*, *33(4)*, 388–417. https://doi.org/10.1108/00907320510631544

Robertson, D. S., & Felicilda-Reynaldo, R. F. D. (2018). Evaluation of graduate nursing students' information literacy self-efficacy and applied skills. *Journal of Nursing Education, 54(3)*, S26–S30. https://doi.org/10.3928/01484834-20150218-03

Robinson, M. D., & Clore, G. L. (2001). Simulation, scenarios, and emotional appraisal: Testing the convergence of real and imagined reactions to emotional stimuli. *Personality and Social Psychology Bulletin, 27(11)*, 1520–1532. https://doi.org/10.1177/01461672012711012

Rosman, T., Mayer, A.-K., & Krampen, G. (2016). Measuring psychology students' information-seeking skills in a situational judgment test format: Construction and validation of the PIKE-P test. *European Journal of Psychological Assessment*, *32(3)*, 220–229. https://doi.org/10.1027/1015-5759/a000239

Schloman, B. F., & Gedeon, J. A. (2007). Creating TRAILS. *Knowledge Quest, 35(5)*, 44–47.

Sturges, P., & Gastinger, A. (2010). Information literacy as a human right. *Libri, 60(3)*, 195–202.

Suen, H. K. (2009). *Principles of test theories*. New York, NY: Routledge.

Tourangeau, R. (2009). Remembering what happened: Memory errors and survey reports. In A. A. Stone, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman, & V. S. Cain (Eds.), *The science of self-report: Implications for research and practice* (pp.29–48). Mahwah, NJ: Lawrence Erlbaum Associates.

Walsh, W. B., & Betz, N. E. (1995). *Tests and assessment* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall, Inc.