Assessing information literacy among German psychology students[1]

Nikolas Leichner, Johannes Peter, Anne-Kathrin Mayer, and Günter Krampen

Leibniz Institute for Psychology Information, Trier, Germany

Structured Abstract:

Purpose: This paper reports on the development of an information literacy test for German psychology students.

Design: The test consists of 22 items covering two facets of information literacy (searching for information and evaluating information) which are mapped on Standards Two and Three of the information literacy framework provided by the Association of College and Research Libraries (ACRL, 2000). A sample of N = 64 German psychology students including n = 22 freshmen, n = 21 advanced students and n = 21 PhD-students completed the test. The freshmen and advanced students also performed an academic literature search task.

Findings: The test has acceptable internal consistency (Cronbach's Alpha between .73 and .82). An analysis of variance showed that PhD-students scored higher than advanced students, who, in turn, scored higher than freshmen. Additionally, the score on the first scale (Evaluating) showed a significant relationship with the performance on the academic literature search task.

Practical Implications: The test can be used to determine training needs among German psychology students, or to evaluate instruction programs.

Originality: The new measure is the only standardized information literacy test for German-speaking populations for which psychometric properties have been reported.

Introduction

The concept "information literacy" describes the ability to recognize when information is needed and subsequently to find, evaluate, and use information to solve the problem. Information literacy is an individual competency, and it is considered the key competency to enhance higher education and to improve workplace effectiveness (National Forum on Information Literacy [NFIL], n. d.). Several definitions of information literacy for higher education settings have been developed; the most important ones have been provided by the Society of College, National and University Libraries in the United Kingdom (SCONUL, 1999), the Association of College and Research Libraries in the USA (ACRL, 2000) and by the Australian and New Zealand Institute for Information Literacy (Bundy, 2004). As a comparison (Boon *et al.*, 2007) showed, all definitions are similar; therefore, only the framework provided by the ACRL (2000) was considered as a basis for the development of the test used in this

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

study. This framework includes five standards which differentiate information literate persons from those who do not possess this competency. The framework is especially useful for the development of an information literacy test, because it includes performance indicators for each standard. The standards are:

(1) Determining the nature and amount of information needed;
(2) Assessing information effectively and efficiently;
(3) Evaluating information and incorporating information into one's knowledge system;
(4) Using the information effectively to accomplish a specific purpose;
(5) Understanding many of the economic, legal, and social issues related to the use of information.

The standards above have been adapted to the domain of Psychology (ACRL, 2010). The Psychology-specific standards are based on earlier work establishing information literacy guidelines for Psychology students (e.g. Lampert, 2005; Merriam *et al.*, 1992; Thaxton *et al.*, 2004).

The aim of the present study is to develop an information literacy test for a German-speaking college and university student population which satisfies psychometric criteria. As far as we know, information literacy among college students in Germany has never been determined using a standardized test, most likely due to the lack of such a test. The only available indicators are surveys (e.g., Franke and Schüller-Zwierlein, 2008). The development of an information literacy test would have several uses: first it could be used to determine information literacy proficiency on an aggregate level, e.g. for planning education. Second, it could be used to determine individual training needs, so students can be assigned to the appropriate course level. Finally, a test could be used to evaluate information literacy instruction.

The measurement of information literacy appears to be more advanced in English-speaking countries, as several tests have already been developed (e.g., Wise *et al.*, 2009; Noe and Bishop, 2005; Ondrusek *et al.*, 2005), most of them to evaluate information literacy courses. Many tests use a multiple-choice format for practical reasons (e.g., ease of use, reliability). Tests using other formats, however, can also be created as a review by Walsh (2009) shows. For only few of these tests, psychometric properties have been determined and reported; one of few exceptions is the Information Literacy Test (ILT, Wise *et al.*, 2009), which is also based on the ACRL framework. The ILT was developed to measure information literacy Standards One, Two, Three, and Five among English-speaking college students. The ILT uses items in multiple-choice format, or similar formats, and allows the reporting of scores for each of the standards and a total score for each participant. A similar test has been developed by Project SAILS (Kent State University, 2000-2012). This test also uses the multiple-choice format and has acceptable psychometric properties. Developing a tool for assessing information literacy is part of a research project currently underway at ZPID, the psychology information center for the German-speaking countries. The main objective of this research project is to develop a blended learning course (combination of E-Learning and traditional courses) to improve information literacy that is targeted at undergraduate psychology students. There is indication that universities should provide training in information literacy: A survey among students at German colleges and universities showed that only around 33% of the participants used databases when conducting a literature search (Franke and Schüller-Zwierlein, 2008). Additionally, a study carried out in the United Kingdom made clear that, initially, incoming students are not sufficiently information literate to study effectively and that students do not become information

literate during the course of their studies (Warwick *et al.*, 2009). A major drawback of most of the existing courses is that only few of them are subject specific. There were only a few efforts to create Psychology-specific information literacy instruction (e.g. Lampert, 2005; Thaxton *et al.*, 2004). This is not surprising, as relevant ACRL Psychology Information Literacy Standards have just become established recently (ACRL, 2010).

As far as we know, there is no standardized Psychology-specific information literacy test publicly available for which psychometric properties have been reported. Thaxton *et al.* (2004) mentioned that an evaluation instrument has been used to evaluate the information literacy training; however, neither the instrument nor its psychometric properties are reported. Another study evaluated an information literacy training for Psychology students by asking students to prepare a literature search plan (Larkin and Pines, 2005). Thus, the current project adds to the field an information literacy test tailored to the field of Psychology for which psychometric properties are available. After reviewing relevant literature, the decision was made to focus mainly on the ACRL (2000) Standards Two and Three in the development of the information literacy course for German psychology students. This restriction was introduced to ensure that the content of the course can be applied in nearly all academic situations. Information literacy instruction directed at the remaining three standards can hardly be given without considering the circumstances, e.g. how to make appropriate use of the found information (Standard Four) will differ substantially depending on the situation. Using the information to write a paper requires abilities and skills differing from those required to prepare a presentation. The test presented in this article will be used to evaluate the success of the course; therefore the scope of the test is also constricted to Standards Two and Three.

Development of items

The items were taken from the literature (Noe and Bishop, 2005; Ondrusek *et al.*, 2005) or were developed by our research group. When developing the items, we took care to consider Psychology-specific aspects of information literacy. For example, one of the items deals with bibliographic databases relevant for psychology students, and other items contained citations which were formatted according to the Standards of the American Psychological Associations. Therefore, the test covers the demands formulated by the ACRL Psychology Information Literacy Standards (ACRL, 2010) quite well. The initial version of the test consisted of $k = 35$ items, of which all but one were multiple-choice items. One item required the participants to write down their response using catchwords, or to write a brief paragraph. All multiple-choice items had three alternatives with one exception; in this case, it seemed more appropriate to use two alternatives only. The items were scored in a way that the score varied from 0 to 1 for each item (e.g., for each of three alternatives which was marked or omitted correctly, a score of 0.33 was given). As ACRL (2000) Standards Two and Three should be covered by the test, each item was assigned to one of the standards: 23 items were assigned to Standard Two, and 12 items were assigned to Standard Three. An example of items representing the two standards is provided below; the Appendix contains the complete test.

Example for Standard Two item:

Which option is most effective when resources (e.g., a book) are not available at your local library?

a) *Contacting libraries in nearby cities, eventually going there*
b) *Use interlibrary loan*

c) *There is no alternative except to buy the literature*

Example for a Standard Three item:

Which statement is true?
The Journal Impact Factor (JIF) indicates…

a) *…how often articles published in this journal have been cited by other authors during a certain period of time*
b) *…how many libraries have subscribed to the journal*
c) *…the relevance ascribed to this journal by a group of experts*

Method

Sample

A sample of *N* = 64 psychology students completed the test at their own pace. The sample consisted of *n* = 22 first-year students (*M* = 21.77 years; 77% female), *n* = 21 advanced students (in their third and fourth year; *M* = 23.90 years, 86% female), and *n* = 21 PhD students (*M* = 28.48 years, 71% female). Students from different stages of their studies were included because we expected upper level (i.e., PhD) students to be more information literate than the advanced students who, in turn, should be more information literate than the first-year students. The participants were paid € 20 for their participation.

Measures and Procedure

The participants first filled out a questionnaire concerning demographic data, e.g. age and prior study duration. The information literacy test was completed next. After that, they completed several other tests and questionnaires, e.g. a questionnaire concerning epistemological beliefs. The results concerning the additional instruments are not dealt with here as this would be beyond the scope of this article. Finally, the first-year students and the advanced students worked on an academic literature search task. The PhD students were not included in this part of the study because of time constraints. The task was completed in our lab, where each participant was seated in front of a computer which allowed them to search the Internet, as well as psychological databases available through the university library's website like PSYNDEX^{plus} and PsycINFO. The participants were given 30 minutes to complete three literature search tasks: for the first task, they were asked to find three recent meta-analyses on the predictive validity of assessment centers. For the second task, a book or a book chapter giving an overview of the assessment center technique had to be found. The final task (which was denoted optional) required the participants to search for a recent empirical study informing about an assessment center with a sample of business students. The participants were asked to document their search results using a Word document. Only the outcome of the first task was analyzed, as the remaining two tasks were filler tasks to ensure that the participants would be busy for around 30 minutes.

Hypotheses

As this was the first use of the test, it was expected that some of the items would not meet psychometric criteria and would, therefore, have to be discarded from the final version of the test.

To meet psychometric criteria, items should neither be too easy nor too difficult for the participants. We expected strong intercorrelations among the items, as the information literacy items require knowledge from one broad domain which can be described as academic information retrieval and evaluation. Consequently, it was expected that a factor analysis would reveal one factor. Finally, to achieve a homogenous test, several items would have to be eliminated due to low factor loadings (low correlations with the remaining items).

After the removal of several items, it was planned to compare the three groups of participants. As mentioned above, there should be substantial group differences on the final version of the information literacy test. PhD students were expected to score higher than advanced students, who, in turn, were expected to score higher than first-year students. These differences between groups were expected because students gain experience in literature search during the course of their studies as they write papers or prepare presentations. Especially, some of the advanced students had already completed their bachelor's thesis, an endeavor that usually requires extensive literature search. So, even if there was no formal instruction in literature search, we assumed that students would become more information literate during the course of their studies. It was also anticipated that PhD students would score higher than the two other groups combined, as pursuing a PhD requires far more information literacy than taking undergraduate or master's level courses. This would be a first step in the validation process. In addition, it was expected that there would be a positive relationship between the performance on the academic literature search task and the information literacy test score. Participants who were successful on the literature search task should also have higher scores on the information literacy test. This relationship should be especially strong when only the items assigned to Standard Two (Searching for information) are regarded.

Results

Exploratory statistical analyses of the data from the 35 information literacy items revealed that item 15 ("What is Google Scholar?") was answered correctly by all respondents. This item was excluded from all further analyses because it could not be included in a factor analysis due to a lack of variance. A factor analysis including the remaining 34 items was computed. An inspection of the Eigenvalues showed that the structure could be explained best by one factor which accounts for 15.46 % of the variance. This is in accordance with the expectations: it was assumed that the items can be combined into one factor.

In the next step, only those 22 items with loadings higher than .20 on the factor were retained. This limit was chosen because using a limit of .30 would have led to the exclusion of many more items, narrowing the scope of the test. Additionally, care was taken to include items covering a wide range of difficulty indices to make the test applicable with more and less information literate persons. Fourteen of these items measure knowledge related to Standard Two (Searching), and 8 items measure knowledge related to Standard Three (Evaluating). The 22 items were combined into one scale by calculating the mean score of the 22 items. For that reason, the total score also ranges from 0 to 1. Calculations of split-half reliability (Spearman-Brown coefficient) and internal consistency (Cronbach's α) indicate that the scale has satisfactory psychometric properties (see Table 1). Standard Two, however, is given far more weight, as there are many more items related to this standard than to Standard Three. When the test is used to evaluate the training, this might not be a problem, as the training will cover both aspects of information literacy. When individual feedback is

given, though, this might be a problem. To circumvent this problem, two subscales were formed. Scale One includes the 14 items related to Standard Two (Searching), while Scale Two includes the 8 items related to Standard Three (Evaluating). Both subscales were formed by calculating the mean score of the items included in that scale, and they both exhibit satisfactory psychometric properties (see Table 1). In agreement with the one-factor solution mentioned above, the two subscales are correlated significantly (r = .60, p < .01).

----insert Table 1 here----------

Group differences

As mentioned above, the test was completed by students in three stages of their studies. Table 2 shows the mean scores on the three scales for the three groups of students. It is obvious that PhD students scored higher than advanced students, who scored higher than first-year students.

----------insert Table 2 here----------

To explore the group differences, an analysis of variance (ANOVA) was computed for each of the scales. Each ANOVA revealed highly significant (p < .01) differences between the groups: F(2,61) = 44.50 for the total score; F(2,61) = 27.10 for the Searching Scale , and F(2,61) = 28.29 for the Evaluating Scale. Differences between the groups were further analyzed by contrast calculations, as the hypotheses were straightforward. For the total score and the Searching Scale, each group differed significantly from the two other groups. For the Evaluating Scale, only the differences between first-year students and PhD students and between advanced students and PhD students reached significance. The difference between first-year students and advanced students was not significant; the difference between the mean scores of the groups, albeit nonsignificant, points in the expected direction. Furthermore, as expected, the PhD students scored significantly higher on all three scales than the two other groups combined. Figure 1 illustrates the differences for the total score.

----------insert Figure 1 here----------

Performance on the academic literature search task

Due to technical problems, the data from three participants had to be discarded. As mentioned above, a sample solution of four studies was determined to evaluate participant performance. All studies were published in reputable journals and could be found using psychological databases in a sophisticated way (e.g., using Boolean operators or by narrowing the results to certain kinds of studies). The original plan was that three out of the four studies had to be found to achieve the maximum possible score; however, the scoring approach had to be changed as none of the participants reached the maximum score. Most of the participants only found one or two of the selected studies; several students were not able to find even one of these studies at all. Therefore, we decided to analyze whether the participants were able to find one of the studies: 29 out of 40 participants (72.5%) were able to do this.

Single mean comparisons were computed to determine whether participants who were successful on the literature search task also had higher scores on the information literacy test. As expected, there was a significant difference on the total score of the information literacy test ($t[38] = 2.18$, $p < .05$). There was a significant difference on the Searching Scale ($t[38] = 2.35$, $p < .05$), but no significant difference on the Evaluating Scale ($t[38] = .97$, $p > .15$).

Discussion

As expected, several items had to be eliminated from the test to ensure acceptable psychometric properties. The two resulting subscales measure two facets of information literacy. The subscales correlate significantly with each other; this is not a surprising result, as all 22 items had high loadings on one factor in a factor analysis. The subscales can be used for giving detailed feedback to participants, while the total score can be used to evaluate information literacy instruction programs on an aggregate level. Therefore, the first hypothesis postulating one factor revealed in a factor analysis was confirmed. The question of whether or not information literacy is a one-dimensional construct will be dealt with later.

The second hypothesis could be confirmed with minor constraints: significant differences could be found between most of the groups. As anticipated, PhD students scored far better than the two other groups. One explanation for this appears most likely: only highly achieving students who also have a modest level of information literacy pursue a PhD degree. It might also be the case that only these students are offered the opportunity to pursue a PhD degree. The difference between the PhD students and the two other groups is especially noticeable when looking at the Evaluating scale (Standard three). This might be the case because many items on this scale refer to knowledge far more relevant for PhD students; for example, several items deal with issues like peer-review, or the Journal Impact Factor. The differences between the groups are a first indicator of the validity of the test.

The responses to most items follow a clear response pattern: PhD students score higher than advanced students, who score higher than first-year students. There are few exceptions, e.g. items 8 and 19 where the pattern is reversed. Both items concern the use of Google Scholar. A possible explanation is that PhD students do not use Google Scholar as often as regular students; therefore they have less knowledge about this service. Mean scores for each item can be found in the Appendix.

It should be noted that analyses of variance computed for each single item score revealed significant group differences on 18 of the 22 items (81%). This finding means that the single items can also be used to distinguish between more or less information literate persons. Consequently, item scores can also be used when giving feedback. Feedback based on single items can point even more precisely to the area where improvement is needed.

As expected, participants who performed successfully on the academic literature search task had higher scores on the information literacy test. There were significant differences in the Searching Scale and in the total score. The difference in the Evaluating Scale did not reach significance. This confirms the third hypothesis. Confirming convergent validity for the Searching Scale and discriminant validity for the Evaluating Scale was anticipated, as only the Searching Scale was

assumed to measure knowledge that would be of use in the literature search task. Together with the differences between participant groups, this is a second indication that the test has validity.

To sum up, our analyses show that the information literacy test consisting of 22 items is a good indicator of ACRL (2000) information literacy Standards Two and Three. It should be noted that this research adds to the field a standardized discipline-specific instrument to assess information literacy among psychology students. The test has acceptable reliability coefficients and differentiates between successful and unsuccessful participants on the academic literature search task. A further indication of the validity of the test is that the three groups of participants score differently on the test. These results demonstrate that the test has validity and can be used to evaluate information literacy courses. The test can also be used to give feedback to individuals; however, the test should not be used as a basis for high stakes decisions (e.g., admission to colleges), as the sample was too small to determine reliability coefficients precisely. When feedback is given, it should be based on the two scales, so that the feedback points more precisely to the area where improvement is needed. Additionally, feedback can also be given using single items; so feedback points more precisely to the area where there is room for improvement. Notwithstanding, the analysis of single item scores should by no means be used for evaluating instruction.

Another point of debate is whether information literacy is a one-dimensional construct. As noted above, a factor analysis revealed that the 22 items can be combined into one factor. This indicates that information literacy (constricted to Standards Two and Three of the ACRL (2000) framework) is a one-dimensional construct. This contrasts with the concept of the ILT (Wise *et al.*, 2009) which assumes that the information literacy standards are independent constructs. As mentioned above, the ILT measures four information literacy standards using a separate scale for each standard. A total score is also reported. Unfortunately, the intercorrelations among the scales are not reported. We can only speculate whether there would be more than one factor if our test covered more than two standards of the ACRL (2000) framework. The most effective way to determine the dimensionality of information literacy would be to develop tests covering all facets of information literacy and to have large samples of participants filling out these tests. The question of whether there is more than one dimension might not have practical implications (e.g., for giving feedback or evaluating courses), however, it will lead theory formulation forward. In the next stage of research, it is of interest to determine the relationship of information literacy to other constructs, like learning styles, previous knowledge, etc. For this kind of research, it would be helpful to have assumptions about the number of information literacy dimensions.

References

Association of College and Research Libraries (ACRL) (2010), “Psychology Information Literacy Standards”, available at: http://www.ala.org/acrl/standards/psych_info_lit (accessed 19 March 2013).

Association of College and Research Libraries (ACRL) (2000), “Information Literacy Competency Standards for Higher Education”, available at: http://www.ala.org/acrl/sites/ala.org.acrl/files/content/standards/standards.pdf (accessed 6 June 2012).

Boon, S., Johnston, B. and Webber, S. (2007), “A phenomenographic study of English faculty's conceptions of information literacy”, *Journal of Documentation*, Vol. 63 No. 2, pp. 204–228.

Bundy, A. (Ed.) (2004), *Australian and New Zealand Information Literacy Framework: principles, standards and practice,* 2nd ed., Australian and New Zealand Institute for Information Literacy, Adelaide.

Franke, F. and Schüller-Zwierlein, A. (2008), "Wie informationskompetent sind die bayerischen Studierenden im Jahr 2007?", *Bibliotheksforum Bayern No.* 02, pp. 36–39.

Kent State University (2000-2012), *Project SAILS (Standardized Assessment of Information Literacy Skills)*, Kent, OH.

Lampert, L. (2005), ""Getting Psyched" About Information Literacy: A Successful Faculty-Librarian Collaboration for Educational Psychology and Counseling", *The Reference Librarian*, Vol. 43 89-90, pp. 5–23.

Larkin, J.E. and Pines, H.A. (2005), "Developing information literacy and research skills in introductory psychology: A case study", *The Journal of Academic Librarianship*, Vol. 31 No. 1, pp. 40–45.

Merriam, J., LaBaugh, R.T. and Butterfield, N.E. (1992), "Library Instruction for Psychology Majors: Minimum Training Guidelines", *Teaching of Psychology*, Vol. 19 No. 1, pp. 34–36.

National Forum on Information Literacy (NFIL) (n. d.), "What is Information Literacy?", available at: http://infolit.org/?page_id=3172 (accessed 6 June 2012).

Noe, N.W. and Bishop, B.A. (2005), "Assessing Auburn University Library's Tiger Information Literacy Tutorial (TILT)", *Reference Services Review*, Vol. 33 No. 2, pp. 173–187.

Ondrusek, A., Dent, V.F., Bonadie-Joseph, I. and Williams, C. (2005), "A longitudinal study of the development and evaluation of an information literacy test", *Reference Services Review*, Vol. 33 No. 4, pp. 388–417.

SCONUL (1999), *Information skills in higher education*, SCONUL, London, UK.

Thaxton, L., Faccioli, M.B. and Mosby, A.P. (2004), "Leveraging collaboration for information literacy in psychology", *Reference Services Review*, Vol. 32 No. 2, pp. 185–189.

Walsh, A. (2009), "Information literacy assessment: Where do we start?", *Journal of Librarianship and Information Science*, Vol. 41 No. 1, pp. 19–28.

Warwick, C., Rimmer, J., Blandford, A., Gow, J. and Buchanan, G. (2009), "Cognitive economy and satisficing in information seeking: A longitudinal study of undergraduate information behavior", *Journal of the American Society for Information Science and Technology*, Vol. 60 No. 12, pp. 2402–2415.

Wise, S.L., Cameron, L., Yang, S.-T., Davis, S.L. and Russell, J. (2009), *The Information Literacy Test (ILT): Test Manual*, Harrisonburg, VA.

Table 1.

Descriptive statistics for the three information literacy scales resulting from the study with *N* = 64 participants.

Scale	*k*	*M* (*SD*)	α	$\rho^*_{xx'}$
Total Scale	22	.64 (.18)	.82	.78
Scale One (Searching)	14	.68 (.19)	.73	.65
Scale Two (Evaluating)	8	.58 (.44)	.73	.73

Table 2.

Mean Scores and standard deviations (in parentheses) of the Information Literacy Test for the three groups.

Scale	First-year students	Advanced Students	PhD students
Total Score	.54 (.09)	.62 (.10)	.78 (.07)
Standard Two	.59 (.11)	.68 (.09)	.79 (.06)
Standard Three	.47 (.13)	.53 (.13)	.74 (.11)

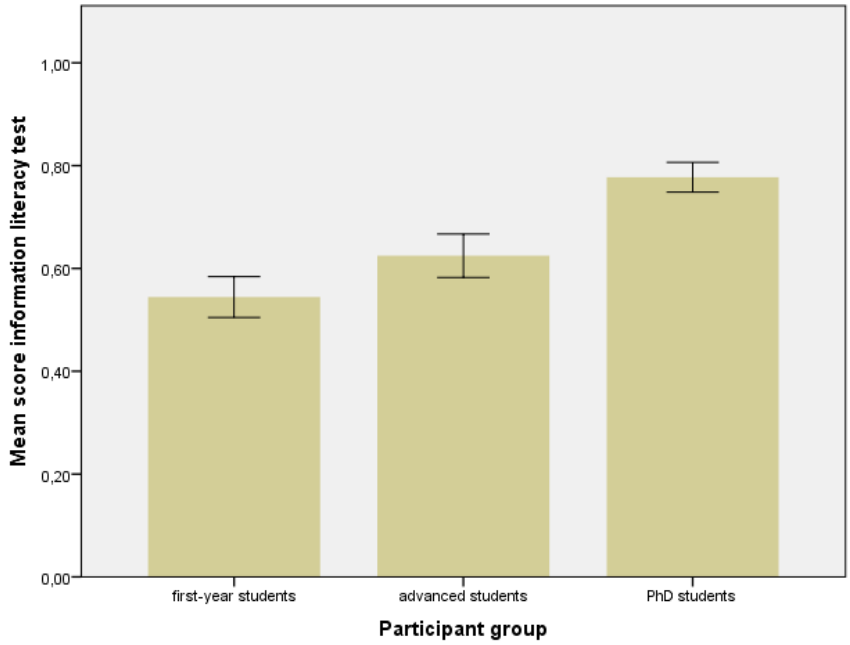

Figure 1.

Group means for the information literacy test total score. Error bars: 95% confidence interval.

Appendix A Information Literacy Test

Item	Question
1 (*)	You come across the following citation: Tulving, E. (1972). Episodic and semantic memory. In E. Tulving & W. Donaldson (Eds.), *Organization of memory.* New York: Academic Press. Which publication type are you dealing with? • Article in a scientific journal • Monograph • **Article in an edited book**
2	Which of the following are important databases in the domain of Psychology? • **PSYNDEX** • **PsycINFO** • EconLit
3 (*)	Which option is most effective when resources (e.g., a book) are not available at your local library? • Contacting libraries in nearby cities, eventually going there • **Use interlibrary loan** • There is no alternative except to buy the literature
4 (*)	Which types of literature are indexed in an electronic library catalog? • **Books available at the library** • **Journals available at the library** • Journal Articles
5 (*)	Which of the following statements about the electronic library catalog are true? • **The library catalog indicates whether a book is currently on loan** • **The library catalog indicates whether the library has a subscription of a certain journal** • Books which are unavailable or currently on loan can be ordered via Interlibrary loan using the library catalog
6(*)	You come across the following citation: Tulving, E., & Watkins, M. J. (1974). On negative transfer: Effects of testing one list on the recall of another. *Journal of Verbal Learning & Verbal Behavior*, *13*(2), 181–193. Which opportunities exist to obtain the publication? • **Searching the electronic library catalog for the journal** • **Searching the electronic periodicals index for the journal** • Searching the electronic library catalog for the journal article
7	Which publication types are indexed in Bibliographic databases? • **Journal articles** • **Books** • Relevant Internet resources
8	Google Scholar is a search engine provided by Google Inc. which can be used to find scientific documents. Is the following statement true? “Google Scholar is a faster and more comfortable alternative to databases” • Yes • No, databases should be preferred • **Basically, yes; however, databases should also be used to make a search comprehensive**

9 (*)	What are the advantages of using Google Scholar instead of specialized databases? • **Google Scholar searches each entry (mostly journal articles) in its entirety for the search term, whereas specialized databases only search the keywords that are linked to the entries** • Google Scholar provides full-text-access to all journal articles • **Google Scholar provides, when available, links to websites where journal articles can be downloaded free of charge**
10	Which of the following statements concerning the use of commercial search engines (e.g. Google, Yahoo!) to find information are true? • The order of the results mirrors the quality of the information found • Because information on the internet is read by many users, wrong information is corrected quickly in most cases • Using search engines, every scientific publication can be found if it was published on the internet
11 (*)	You found an internet forum which provides relevant information. Is it allowed to use this information (properly cited) for your scientific presentation? • **Eventually, this depends on the forum. There are scientific internet forums which have the same reputation as scientific journals** • No, internet forums are by no means appropriate resources for your presentation, because they do not fulfill requirements for scientific resources • No, because there are no guidelines on how to cite internet forums. Therefore, forums cannot be used as scientific resources
12 (*)	Which are the features of a scientific internet forum? • **The forum is provided by a reputable organization** • **Entries are archived, so that the discussion can be retraced** • **Authors publish using their real names instead of pseudonyms**
13 (*)	While searching for literature you came across a seven years old book. Using databases you could find several more recent journal articles. Which are the reasons to read the book first? • **The book conveys probably a better overview of the topic and various research directions** • Books are written only when knowledge is established • **Journal articles often deal with specific aspects of an issue which might not be relevant for me**
14 (*)	You are trying to find information concerning a specific subject. Which are appropriate approaches to identify search terms? • **Summarizing the issue using a few sentences. The central concepts of the summary can be used as search terms** • **Finding synonyms for the central concepts, e.g. by brainstorming, or the use of a lexicon** • **Reading an relevant article and noting related concepts**
15	What is Google Scholar? • **A search engine provided by Google Inc. which can be used to find scientific documents** • A scholarship program of Google Inc. • An interactive tutorial of Google Inc.

16	You are trying to find information about a topic which is relatively unknown to you. As sources of information, textbooks (which report briefly on your topic), Wikipedia and databases are available. In which order should these resources be used to efficiently obtain an overview of the topic, and then to find relevant literature? Not all sources have to be used. • **Wikipedia -> Textbooks -> Databases** • **Textbooks -> Databases** • Exclusively using databases
17 (*)	You are trying to find information concerning the influence of personality characteristics on false memories. Which of the following queries do you consider helpful? • **„Personality AND false memories“** • „Personality influence false memories“ • **„Extraversion AND false memories“**
18 (*)	Why does the query “Personality AND false memories” provide fewer results than the query “Personality OR false memory”? • **Combining two search terms using AND restricts the results to entries associated with both terms. When combining the terms using OR, all entries are found that are linked to one of the terms, or linked to both** • More entries are associated with the word “OR” than with the word “AND” • “OR” leads to the inclusion of additional resources which would otherwise not be used
19	For which reasons might using Google Scholar be more productive than using databases? • Google Scholar does not allow the user to limit the results to items published during a certain period of time • Databases index journals only, while Google Scholar does also contain other sources, e.g. books • **Entries in a database can only be found when searching for the relevant key words. Google Scholar, contrastingly, searches the complete entry for the search terms**
20 (*)	Which of the following statements is correct? Searching the thesaurus can be done to • Identify subdisciplines of Psychology • Identify important scientists in a certain field • **Find additional search terms**
21 (*)	Which of the following statements are correct? The SFX-Button located next to the results of a search engine can be used to • **find out whether a certain journal is available at your university library** • **find out whether the journal article in question is available online for students of your university** • **find out whether a certain book is available at your university library**
22	Is the following statement correct? After reading the results page of a database, you can be sure that you missed only few relevant sources • Yes, the important databases, like PsycINFO, or PSYNDEX, cover the current state of research comprehensively • If the topic is of interdisciplinary nature, it might be helpful to include findings from databases from both fields • **Basically, both options above are correct. When searching a database, the user should notice which period which is covered by the database. Some databases go back only until the 80s when computers became popular**

23	The results page of a database provides additional information about each entry. Which information can be used to evaluate the quality and relevance of the entry? • **Name of the author** • **Name (reputation) of the journal** • **Abstract**
24	A new hypothesis is put up in a journal article. Select the two most promising options to find out how other researchers think about the new hypothesis. • Waiting until a review article is published • **Searching for journal articles that cite the article in question** • **Using a database to find more information about the topic in question**
25	Which of the following is usually not a relevant criterion to evaluate the quality of an article? • Publication date • **Number of authors** • Reputation of the journal
26 (*)	Which criteria can be used to judge the quality of a book or a journal article? [students were asked to write down their response using catchwords For each of the following responses, a score of 0.33 was given. No more than three correct responses were scored, to make this item equivalent to the other items • Is the publication peer-reviewed? • Reputation of the author • Reputation of the publisher • Journal impact Factor • Publication year]
27 (*)	Which of the following problems can occur during the peer-review process? • None, therefore peer-review is the best solution • **Reviewers might reject articles that contradict their own findings** • **Often it is not possible to find an adequate number of reviewers**
28 (*)	Which reasons exist not to use the reputation of an author to judge the quality of a publication? • **Qualified younger authors might get disadvantaged, because they did not have the chance to gather reputation** • **Unknown authors might also publish articles of quality which would be overlooked** • **The peer-review process should make sure that all articles fulfill certain standards, no matter who wrote them**
29	Is the following statement true? "Publications which are older than ten years are generally irrelevant." • Yes • Yes, except for epochal books • **No, older literature can also provide important information**
30 (*)	Which statements about the order of authors of a publication are true? • **Sometimes, the order is determined randomly** • **Often, those authors who contributed more to the study are listed first** • **Established authors are often listed last**
31	How can you check the reputation of an author in a reliable way? • **Using the number of citations of his works** • Examining the homepage of the author • **Whether or not the author is cited in relevant textbooks is a good indicator**

32 (*)	Which statement is true? "The reputation of a publisher is more important when evaluating the quality of • Journals, because the reputation of the publisher is more important than the Journal Impact Factor" • **Monographs or edited books because there is no Impact Factor available for these publications"**
33 (*)	Which options are available to find out about the reputation of a publisher? • **Asking peers (e.g. colleagues)** • **Using information on Wikipedia** • **Asking the relevant contact person at your library**
34 (*)	Which statement is true? The Journal Impact Factor (JIF) indicates… • **…how often articles published in this journal have been cited by other authors during a certain period of time** • …how many libraries have subscribed to the journal • …the relevance ascribed to this journal by a group of experts
35 (*)	The Journal Impact Factor is not available at every institution due to license fees. Which options are feasible for students? • Asking their universities for reimbursement • **Eigenfactor.org is an alternative available free of charge** • **The Journal Impact Factor values can be found on the websites of the journals**

Remarks: Correct response options are marked using bold font.
The first 23 items are assigned to ACRL standard two; the remaining items are assigned to standard three. The 22 items which were retained after conducting psychometric analyses are marked with an asterisk (*)

Appendix B Psychometric properties of the items

Item	Total sample (N = 64)			Subgroups					
				First-year students (n = 22)		Advanced students (n = 21)		PhD students (n = 21)	
	M	*SD*	r_{it}	*M*	*SD*	*M*	*SD*	*M*	*SD*
Item 1	0.86	0.25	0.47	0.71	0.30	0.89	0.24	0.98	0.07
Item 2	0.94	0.17	-	0.89	0.22	0.97	0.10	0.95	0.16
Item 3	0.95	0.14	0.33	0.89	0.19	0.95	0.12	1.00	0.00
Item 4	0.75	0.28	0.38	0.70	0.27	0.63	0.33	0.92	0.15
Item 5	0.62	0.22	0.36	0.55	0.19	0.60	0.23	0.71	0.22
Item 6	0.65	0.33	0.51	0.45	0.30	0.63	0.33	0.86	0.23
Item 7	0.78	0.22	-	0.74	0.27	0.76	0.19	0.83	0.20
Item 8	0.76	0.32	-	0.80	0.32	0.75	0.31	0.73	0.33
Item 9	0.71	0.28	0.28	0.52	0.27	0.81	0.23	0.83	0.23
Item 10	0.73	0.25	-	0.71	0.31	0.75	0.18	0.75	0.23
Item 11	0.52	0.34	0.32	0.44	0.30	0.41	0.31	0.71	0.34
Item 12	0.51	0.36	0.25	0.41	0.38	0.57	0.34	0.54	0.36
Item 13	0.79	0.23	0.38	0.73	0.27	0.75	0.23	0.89	0.16
Item 14	0.68	0.25	0.18	0.61	0.28	0.75	0.21	0.70	0.23
Item 15	1.00	0.00	-	1.00	0.00	1.00	0.00	1.00	0.00
Item 16	0.65	0.12	-	0.64	0.14	0.68	0.07	0.63	0.15
Item 17	0.54	0.33	0.35	0.45	0.36	0.48	0.34	0.68	0.25
Item 18	0.97	0.11	0.34	0.92	0.18	0.98	0.07	1.00	0.00
Item 19	0.77	0.23	-	0.80	0.24	0.75	0.21	0.75	0.23
Item 20	0.84	0.19	0.55	0.79	0.16	0.79	0.20	0.94	0.17

Item 21	0.23	0.32	0.50	0.06	0.17	0.22	0.22	0.43	0.42
Item 22	0.65	0.22	-	0.64	0.23	0.65	0.22	0.65	0.22
Item 23	0.58	0.27	-	0.59	0.27	0.57	0.26	0.57	0.30
Item 24	0.85	0.20	-	0.82	0.22	0.86	0.23	0.89	0.16
Item 25	0.76	0.25	-	0.79	0.22	0.78	0.24	0.71	0.28
Item 26	0.56	0.35	0.52	0.32	0.32	0.57	0.28	0.79	0.29
Item 27	0.66	0.19	0.36	0.58	0.23	0.70	0.18	0.71	0.12
Item 28	0.63	0.29	0.50	0.50	0.30	0.59	0.26	0.79	0.25
Item 29	0.96	0.15	-	0.94	0.20	0.95	0.16	1.00	0.00
Item 30	0.41	0.23	0.50	0.29	0.21	0.33	0.11	0.60	0.23
Item 31	0.81	0.24	-	0.80	0.20	0.84	0.25	0.79	0.27
Item 32	0.72	0.34	0.58	0.61	0.26	0.57	0.43	0.98	0.11
Item 33	0.39	0.33	0.38	0.30	0.32	0.35	0.31	0.52	0.33
Item 34	0.78	0.27	0.21	0.74	0.20	0.70	0.31	0.90	0.24
Item 35	0.49	0.22	0.33	0.41	0.18	0.46	0.22	0.60	0.23

Remarks: M = Mean, SD = Standard deviation, r_{it} = item-total correlation

For those items that were not included in the final version of the test, no item-total correlations are reported. The item-total correlations reported here refer to the correlations with the total scale (consisting of 22 items).