Combining Self-Assessments and Achievement Tests in Information Literacy Assessment: Empirical Results and Recommendations for Practice

T. Rosman[a], A.-K. Mayer[b] and G. Krampen[c]

[a]*Leibniz Institute for Psychology Information (ZPID), Universitätsring 15, 54296 Trier, Germany; Phone: +49 (0)651 201-2045; E-Mail: rosman@zpid.de*

[b]*Leibniz Institute for Psychology Information (ZPID), Universitätsring 15, 54296 Trier, Germany; Phone: +49 (0)651 201-2930; E-Mail: mayer@zpid.de*

[c]*University of Trier and Leibniz Institute for Psychology Information (ZPID), Universitätsring 15, 54296 Trier, Germany; Phone: +49 (0)651 201-2967; E-Mail: krampen@uni-trier.de*

This is an **Accepted Manuscript** of an article published by Taylor & Francis in Assessment & Evaluation in Higher Education on August 26, 2014, available online: http://www.tandfonline.com/doi/full/10.1080/02602938.2014.950554

Recommended citation:

Rosman, T., Mayer, A.-K., & Krampen, G. (2015). Combining self-assessments and achievement tests in information literacy assessment: Empirical results and recommendations for practice. *Assessment and Evaluation in Higher Education, 40*(5), 740-754. doi: 10.1080/02602938.2014.950554.

Acknowledgements: This work was supported by the Joint Initiative for Research and Innovation with a grant that was acquired within the Leibniz Competition 2013 under Grant SAW-2013-ZPID-1 195. We would like to thank Prof. Dr. Dirk Wentura from Saarland University for supporting our data collection.

Disclosure statement: The authors declare that there is no actual or potential conflict of interest with regard to the present manuscript.

Abstract

This article examines the significance of information literacy self-assessments in higher education with a special focus on situational conditions increasing their explanatory power. First, it was hypothesized that self-assessments of information literacy correlate higher with factual information literacy if measured *after* the administration of information search tasks (and tests) due to the intrinsic feedback associated with their completion. Furthermore, it was assumed that self-assessments measured after the administration of information search tasks explain incremental variance over standardized information literacy tests in information search performance. A study with *N* = 82 German psychology students was carried out to verify these assumptions. Information literacy self-efficacy (as one form of self-assessment) as well as the PIKE-P information literacy test and several standardized information search tasks were included in the test battery. As both assumptions were largely supported, we recommend that researchers complement their test batteries by self-assessments, but warn them against relying primarily on this method. We further recommend that self-assessments should take place at the end of testing.

Keywords: Self-assessments, achievement test, validity, information literacy

Introduction

Information literacy is defined as the ability to recognize an information need and to subsequently locate, evaluate, and use the needed information (American Library Association 1989). In higher education, due to an exponentially growing body of scientific information, proficiency with the complex methods to retrieve such information has become of central importance for student learning and academic achievement (Johnston and Webber 2003; Bruce 2004; Bowles-Terry 2012). Despite some encouraging efforts, the assessment of information literacy is however still in its infancy (Neely 2006; Radcliff, Jensen, Salem, Burhanna, and Gedeon 2007; Walsh 2009), and researchers have not yet been able to find a consensus on how the measurement of this multifaceted concept should be approached.

In the literature, three fundamentally different conceptions of information literacy assessment can be identified: Achievement tests, information search tasks, and self-assessments. Unfortunately, these methods are mostly employed independent of each other, and comparative studies on their objectivity, reliability, and validity are scarce. Particularly, very little is known about the validity of self-assessments in information literacy measurement. The present article fills this gap by focusing on conditions that influence self-assessment accuracy and by examining the use of self-assessments in enhancing the validity of information literacy assessment batteries.

The assessment of information literacy with achievement tests

Considerable effort has been put in assessing information literacy ‘objectively’ through standardized achievement tests. By drawing on a multiple choice format, the majority of these tests are easy to administer while also maximizing scoring objectivity. For example, the Information Literacy Test (ILT) by Wise, Cameron, Yang, Davis, and Russell (2009) consists of items that deal – among others – with knowledge about publication types, source credibility, or Boolean operators. The main advantage of such tests is that they prevent faking through deliberate overreporting of abilities. However, they mostly measure isolated declarative knowledge and are not well-suited to address higher-order skills (Scharf, Elliot, Huey, Briller, and Joshi 2007). Recently, the Procedural Information-Seeking Knowledge Test (PIKE) has been introduced, which, by drawing on a situational judgment (or scenario-based) test format, focuses both declarative *and* procedural knowledge (Rosman and Birke, forthcoming; Rosman, Mayer, and Krampen 2014).

A second, somewhat underrepresented method, is the assessment by means of standardized information search tasks (or assignments), which are completed using common search tools like web search engines and bibliographic databases (e. g., Ivanitskaya, O'Boyle, and Casey 2006; Kim 2009). Leichner, Peter, Mayer, and Krampen (2014) adopted this approach and designed nine different information search tasks as well as a corresponding standardized scoring key. For example, they asked subjects to find '… *meta-analyses published after 2005 investigating "risk factors" for the development of a "Posttraumatic stress disorder"*' and document their search results. As information literacy is *defined* as the ability to successfully conduct information searches (American Library Association 1989), the structure of these tasks corresponds with the demands of scholarly information searches in real life. Therefore, performance in information search tasks may be interpreted as a highly representative (or ecologically valid) measure of this rather complex set of competencies. Of course, assessments of information literacy through evaluating students' portfolios or scientific papers (combined with standardized observations during the required information searches) would be the best indicators for information literate behaviour. Due to its extremely time-consuming nature (Walsh 2009), this method is however not an option in practice. Even information search tasks are of limited practical value as their administration and scoring are very time-consuming (Chang, Zhang, Mokhtar, Foo, Majid, Luyt, and Theng 2012; Leichner et al. 2014), a fact that may well serve as an explanation for their low popularity.

Self-assessments of information literacy

As the comprehensive review by Walsh (2009) shows, many intend to measure information literacy with self-assessment inventories. This practice seems especially prevalent in Europe, where most studies solely rely on assessing information literacy using self-reports (e. g., Gapski and Tekster 2009; Heinze, Fink, and Wolf 2009; Pinto 2010). Most inventories focus on subjective information search ability and subjective (prior) performance (e. g., domain-specific self-efficacy; Kurbanoglu, Akkoyunlu, and Umay 2006; Behm, forthcoming). For example, Kurbanoglu et al. asked students about their confidence in selecting information most appropriate to the information need. From the point of view of educational psychology, this approach is promising: First, subjective ability (in terms of self-efficacy) is often considered a '*… core belief [that] is the foundation of human motivation, performance accomplishments, and emotional well-being*' (Bandura 2010, 1534) and can therefore positively influence effort expenditure and task persistence, especially in the case of obstacles (Schunk 1984). Due to these positive effects of self-efficacy on self-regulation, self-assessments might have a special diagnostic value, as they are suited to investigate *why* students fail in objective tests (e. g., due to their low self-efficacy), shedding light on the motivational reasons of success and failure.

Second, the *assessment* of subjective ability and performance as such can have positive effects on subsequent performance. This is especially true if the measurement instrument draws on different criteria that define competency in a certain domain, because self-assessments relating to such criteria enable learners to reflect on their abilities and strategies and to monitor their learning progress (Boud 1995). Conscious control over learning increases and students become metacognitively aware of their knowledge and thought (see already Vygotsky 1962). This, in turn, enables them to identify their strengths and weaknesses, and to direct their behaviour to act accordingly on eventual weaknesses (Lew, Alwis, and Schmidt 2010). For example, a student recognizing that he/she has difficulties in selecting information sources adequate to the inquiry might put more emphasis on learning about the scope of different information sources. Besides these direct effects on learning, self-assessments also enable students to become more independent learners as they are actively involved in planning and managing their own learning, thus facilitating the shift from passive reception to active construction of knowledge. Self-assessments can even promote *'meta-cognitive competencies such as self-reflection and self-evaluation'* (Dochy, Segers, and Sluijsmans 1999, 332) and thus enhance students' ability to self-assess their learning in the future (Oscarson 1989).

Finally, including self-assessments in test batteries might be suited to expand the range of the assessment (Oscarson 1989), because, for some aspects of information literacy, learners own appreciations of their abilities are superior to those made by achievement tests. For example, an information literacy achievement test might be limited to more concrete and specific aspects of information literate behaviour (e. g., the use of certain databases) and neglect metacognitive aspects like monitoring and reflection. Self-assessments, on the other hand, are well-suited to investigate these more subjective aspects. Oscarson (1989) and Falchikov (2005) provide overviews over these and other benefits of self-assessments.

The accuracy of information literacy self-assessments

Despite these advantages, the measurement of ability with self-assessment inventories is a slippery task and many studies question their validity. A rather consistent finding, for example, is that self-assessment accuracy is a function of individual ability (Boud and Falchikov 1989; Leach 2012): Lower ability students tend to overestimate their abilities because they lack experience in the respective domain, and thus don't have an empirical basis for their judgements: "*The same knowledge that underlies the ability to produce correct judgment is also the knowledge that underlies the ability to recognize correct judgment.*" (Kruger and Dunning 1999, 1121-1122). Moreover, low experience in a domain might make self-assessment instruments more vulnerable for the biasing effects of general self-beliefs (e. g., general self-efficacy: Schwarzer, Bäßler, Kwiatek, Schröder, and Zhang 1997; or global self-esteem: Betz and Klein 1996).

With this in mind, it comes as no surprise that empirical evidence on the relationship between self-reported and actual ability is rather mixed. For example, Freund and Kasten, while trying to meta-analytically predict IQ scores with self-reports of intelligence, concluded: '*... using self-estimates as proxies for standardized ability tests appears to be a rather imprecise endeavor, suggesting that the validity of self-estimates of cognitive ability is not very high*' (2012, 331). Brackett and Mayer even go so far as to claim that '*... self-report and ability scales only modestly correlate because people are notoriously bad at assessing their own capacities*' (2003, 1155). On the other hand, however, Benton, Duchon, and Pallett (2013) and Boud, Lawson, and Thompson (2013) could find positive relationships between self- and tutor-rated performance as well as course examination results.

Although a few studies have been conducted on that subject in information literacy research, no final conclusion on the validity of information literacy self-assessments can be drawn yet. For example, Coupe (1993) as well as Ivanitskaya et al. (2006) found low but significant relationships between self-reported and actual skill level (measured by an information literacy test). Geffert and Christensen (1998, 279), however, found no correlation between '*students' test scores and their levels of self-confidence, comfort in libraries, or self-assessment of library skills*'. Another study by Gross and Latham (2007) deserves further attention: Before and after completing the ILT (see above; Wise et al. 2009), subjects were asked to estimate how well they would do (or had done, respectively) in the test. Two major findings of this study are relevant: First, many students overestimated their test performance, which is in line with findings of Maughan (2001). Second, estimated performance *prior* to the test did not correlate with test scores, whereas the respective correlation *after* completing the test was significant (and positive). This finding might seem trivial, but it shows that exposure to an information-search related task can actually help students assess their performance more adequately, taking the experience with the task as a judgment base for their self-assessments.

In light of these findings, one can conclude that self-assessments are related to information literacy test performance, but correlations tend to be rather low. Nevertheless, as depicted earlier, self-assessments are a promising way of investigating those parts of information literacy that are subjective in nature, therefore complementing ‘objective’ assessments of the broad and fuzzy concept of information literacy. Moreover, due to their ascribed positive effects on metacognition and learning (see above), they might be a valuable tool to help students develop adequate information-seeking skills.

When applying self-assessment inventories, measurement circumstances seem to play a crucial role for their accuracy. Whereas moderating variables of self-assessment accuracy have been investigated in many different educational domains (e. g., Falchikov and Boud 1989; Lew et al. 2010; Benton et al. 2013; Jackson 2014), only few corresponding studies have been conducted with regard to information literacy. As depicted earlier, Gross and Latham (2007) demonstrated that students give more precise estimations of their test results after finishing an information literacy test than before. This suggests that exposure to an information-search-related task can actually moderate self-assessment accuracy. Even though the simple estimation of test results is a rather narrow indicator, a study by Ackerman and Wolman (2007) suggests that these findings may be generalizable to broader domains (e. g., self-assessed information literacy): While using a test design identical to the one by Gross and Latham (2007; see above), they had their subjects take a wide range of ability tests (e. g., vocabulary tests, mathematics tests, etc.). Before and after completing these tests, participants were asked to self-assess their ability in the corresponding domains (e. g., verbal ability, mathematical ability, etc.). The authors largely confirmed their hypothesis that self-assessment accuracy increases upon completion of the respective ability test. They attributed this to two factors: Both the perceived level of effort demanded by the test and the '*intrinsic feedback*' (Ackerman and Wolman 2007, 61) associated with test completion (e. g., number of items solved etc.) provide information about individual ability. Findings that self-assessment accuracy indeed increases after exposure to feedback provide support for the latter assumption (Boud and Falchikov 1989; Taras 2001). As noted earlier, information search tasks have a high ecological validity. As they additionally provide a large amount of inherent feedback ('Did I find the demanded publication or not?'), we conclude that completing information search tasks should enable subjects to calibrate their self-assessments even more accurately than the completion of standardized tests. Consequently, changes in self-assessments should correlate with search task performance: The higher task performance, the more self-reported ability will increase from the first to the second measurement point; the lower task performance, the more self-reported ability will decrease, respectively.

Hypothesis 1a: Information literacy self-assessments and information literacy – both test and search task scores – will only be associated when the self-assessment takes place after the completion of the information search tasks or tests.

Hypothesis 1b: The change in self-assessed information literacy from the first to the second phase of self-assessment is attributable to information literacy test and information search task performance: The higher test and task performance, the higher the increase in self-assessment scores.

Incremental validity of information literacy self-assessments

With regard to the low correlations between self-reported and actual information literacy levels (see above), it seems naïve to assume that self-assessments can *replace* standardized tests. One might however ask whether they are suited to enhance the explanatory power within an information literacy assessment battery. As no studies have yet been conducted on the incremental validity of self-assessments over achievement tests in information literacy assessment, it remains unclear whether subjective and objective measures explain the same or different variance proportions of information literacy. There are, however, some arguments in favour of the latter assumption. In fact, subjectively misjudging one's objective abilities can have a considerable impact. Overestimation of abilities might reduce motivation by procuring feelings of frustration and failure (Freund and Kasten 2012), while underestimation might impair outcome expectancies and therefore reduce motivation as well (Bandura 1994). Ackerman and Wolman (2007, 61) summarize this by stating that '*… in order to function reasonably successfully with their environments, adults must have a good general sense of what they are and are not capable of doing*'. Doubtlessly, even an information literate student's motivation is compromised if he thinks he lacks information search ability. Thus, adequate information searching supposedly requires both objective skills (as measured by an information literacy test) and cognitions about individual experiences and abilities that enable students to develop their potential while working on the tasks (e. g., confidence or self-efficacy; Bandura 1994). Additionally, due to the complex nature of the concept, it is very difficult to design a test that comprehensively covers all aspects of information literacy (for example, certain bibliographic databases are omitted, etc.). On that account, having subjects report about their prior performance and ability might very well be a good idea. In sum, all these points speak in favour of measuring both objective and subjective components of information literacy. Hence, it would make sense to integrate self-assessment items into an information literacy test battery when striving for a more comprehensive assessment. Thus, a second purpose of this article is to determine whether self-assessments explain incremental variance over a standardized information literacy test in predicting information search task performance. The following hypothesis is suggested:

Hypothesis 2: Self-assessed information literacy predicts search task performance over and above a standardized information literacy test if the self-assessment takes place after assessing information literacy by a standardized test.

Method

Participants and procedure

Participants were *N* = 82 psychology students from Saarland University in Saarbrücken, Germany. Due to technical problems, data from one subject was eliminated from the dataset. For the remaining *N* = 81 participants, mean age was *M* = 22.33 years (*SD* = 2.99). Gender was distributed very unevenly with 91 % female and 9 % male participants, a distribution nevertheless typical for German psychology students (Wentura, Ziegler, Scheuer, Bölte, Rammsayer, and Salewski 2013). Two third (67 %) of the students were enrolled in an undergraduate programme and one third (33 %) in a master’s degree programme. The sample was recruited by means of flyers and mailing lists inviting psychology students to take part in an information literacy study. Participants were paid for their participation. Data was collected in a computer lab in groups of 8 to 15 subjects; the sessions took approximately two hours. All participants were assured that data was collected anonymously.

Measures

Self-reported information literacy was measured with the Self-Efficacy Scale for Information Searching Behaviour (SES-IB-16; Behm, forthcoming). The SES-IB-16 consists of 16 items such as ‘When searching for information on a specific subject, I am able to use different sources of information in a way to obtain a maximum of relevant information.’, which are to be rated on a 5-point Likert-Scale. The scale was designed to comprehensively cover all phases the information search process: Its items relate to seven different criteria that define information literate behaviour (e. g., information searching, selecting relevant information, integrating information, monitoring the search process, etc.), thus facilitating bottom-up processing (processing that relates to concrete experiences with the task) and initiating reflexive and metacognitive processes with regard to students’ information search ability. The SES-IB was administered twice; both at the beginning and at the end of the assessment battery.

The *information search tasks* consisted of three increasingly difficult tasks (e. g., medium level: ‘Find two longitudinal studies of risk factors for generalized anxiety disorder published after 2005.’) that were based on the work by Leichner et al. (2014). Subjects were asked to enter the title and first author of the discovered publications on a separate page. Additionally, they were required to answer some questions about their course of action while searching (e. g., ‘Where did you search and which search keywords did you use?’). With the help of the standardized scoring rubric by Leichner et al. (2014), two different scores were calculated in order to rate subjects’ performance: An *outcome score* reflecting the fit of the documented publications to the task requirements, and a *process score* indicating the quality of the search process itself.

Information literacy test levels were measured by the PIKE-P test (Rosman et al. 2014; Rosman and Birke, forthcoming), an adoption of the PIKE test for psychology students. Its situational judgment (or scenario-based) test format can be described as follows: Subjects are given a short description of a problem situation requiring an information search (e. g., ‘You need the following book: “Richard S. Lazarus - Stress, Appraisal, and Coping”. How do you proceed?'), and four response alternatives which differ in their appropriateness to resolve the problem (e. g., 'I look up the ISBN-Number of the book and enter it into the library catalog.'). Their task is to rate the appropriateness of each response alternative on a 5-point Likert-Scale ranging from 'not useful at all' to 'very useful'. Test scores are obtained by applying a standardized scoring key which is based on the ranking of response alternatives by experts.

Results

Hypothesis 1

As shown in Table 1, correlations between SES-IB-16 and PIKE-P are non-significant at the first measurement point and significant at the second measurement point. The same pattern is found for correlations between SES-IB-16 and the process score of the information search tasks. Regarding the relation between the SES-IB-16 and the outcome score of the information search tasks, the corresponding correlation at the second measurement point narrowly missed statistical significance (p = .06). In light of these findings, Hypothesis 1a can be seen as largely supported.

------ Insert Table 1 here ------

To test Hypothesis 1b, subjects' SES-IB-16 scores of the first measurement point were subtracted from their SES-IB-16 scores of the second measurement point. The resulting difference score is positive for subjects whose self-efficacy increases from T1 to T2, and negative for subjects with decreasing self-efficacy. To avoid the respective relationships being biased by baseline levels of self-efficacy, we controlled for SES-IB-16_T1 while predicting PIKE-P as well as information search tasks scores from the difference score in a multiple regression. As can be seen in Table 2, our hypothesis is supported for both scores of the information search tasks. For the PIKE-P as dependent variable, the respective beta weight was only marginally significant ($p < .10$). Hypothesis 1b is also largely supported.

------ Insert Table 2 here ------

Hypothesis 2

Hypothesis 2 was tested using multiple regression. As shown in Table 3, SES-IB-16 (second measurement point) remained a significant predictor of the process score of the information search tasks when PIKE-P test scores were entered into the equation. It could be found that SES-IB-16 explains 4.5 % of incremental variance over PIKE-P in the process score of the information search tasks. For the outcome score of the search tasks, however, the respective beta weight did not reach statistical significance. In sum, Hypothesis 2 is partially supported.

------ Insert Table 3 here ------

Discussion

The purpose of this article was twofold: First, we aimed to verify that the accuracy of information literacy self-assessments increases after exposure to information-search-related tasks and whether the change in self-assessed information literacy from the first to the second measurement point is attributable to information literacy test and information search task performance. This hypothesis was supported. A second purpose was to determine whether self-assessments of information literacy have incremental validity over standardized tests in the prediction of information literacy when they are measured at the end of an information literacy assessment battery. This assumption could be supported for the process score of the information search tasks, but not for the outcome score.

Effects of information literacy task exposure on self-assessment accuracy

The confirmation of Hypothesis 1 indicates that self-assessments of ability (in terms of domain-specific self-efficacy) are more accurate *after* exposure to information literacy tests and search tasks. This confirms that the findings by Gross and Latham (2007) are not only applicable to self-reported test performance, but that they can indeed be generalized to the broader domain of ability self-assessments. Thus, our research serves as a potential explanation for the inconsistent results on the relationship between self-assessed and actual information literacy. As hypothesized earlier, the increase in accuracy could be due to an *anchor effect*, as students lack information to adequately assess their ability when they have not searched for information in a while. Additionally, self-assessment questions are presumably seen in a different light after working on search tasks: As memory recall is a dynamic process, more specific aspects of one’s information search abilities may come to mind while working on the information search tasks, and form a basis for further self-assessments. This might have been facilitated by the fact that our self-assessment included seven distinct criteria for information literate behaviour (see above), thus inducing more differentiated reflective processes: For example, students might not just have asked themselves whether they ‘did good’ in the information search tasks, but also whether they had selected the correct information sources, evaluated the search results using clear criteria, and metacognitively monitored the search process. These assumptions were tested by calculating a difference score between post- and pre-test self-efficacy levels. Multiple regression analyses showed that the reinforced accuracy of self-assessments at the second measurement point is mainly attributable to performance in information search tasks. This same pattern exists for the PIKE-P test, although its magnitude is rather small. In light of the difference between information search tasks and the PIKE-P test, this finding is easily explained: The PIKE-P test does not provide much feedback if subjects' answers are correct, whereas information search tasks enable subjects to easily figure out whether their search was effective and efficient or not.

Incremental validity of self-assessed information literacy over standardized tests

With respect to Hypothesis 2, our findings indicate that self-assessed information literacy, if measured at the end of the assessment battery, can explain variance in information literacy over and above a standardized information literacy test. This is likely due to search tasks measuring information literacy more comprehensively than standardized tests (see above). In that sense, the incremental variance explained by self-assessment instruments makes up for the imperfectness (or narrower scope) of information literacy tests.

A limitation regarding the second hypothesis is that the findings could not be reproduced with regard to the outcome score of the information search tasks. This might be due to the lower reliability of the outcome scores (see Table 1), which is accounted for by two factors: First, rating the outcome quality of information search tasks can be a challenging task, '*… as the judgment whether an article covers a certain topic is more ambiguous than rating whether a certain resource has been used or not*' (Leichner et al. 2014, 11). Second, outcome scores may be distorted by subjects finding the demanded publications by chance, for example by searching in Google™ and going through an extensive list of results (Leichner et al. 2014). With regard to the process score, however, accessing a bibliographic database or choosing the right keywords 'by chance' is much less likely. Another potential limitation regarding Hypothesis 2 might be that the amount of incremental variance explained by the self-assessments is rather small. This finding, however, is in full accordance with the current state of research, and it would – at least in our opinion – be naïve to assume it were different. In sum, we conclude that self-assessments of information literacy explain a small proportion of incremental variance in information literacy above standardized tests if they are administered at the end of testing.

Limitations

A few potential limitations of our efforts should not be withheld. First, method bias resulting from the repeated measurement of information literacy self-efficacy cannot be ruled out in our design. There is a good case to believe that performance in search tasks is accountable for the changes in information literacy self-efficacy at the second measurement point, but a certain likelihood that the changes are simply caused by the repeated measurement still exists. Adopting a field-experimental design with a control group who solely answers the self-report questions (and gets no 'treatment' in form of tasks or tests) would remedy this limitation.

Second, self-assessments usually refer to both '*the involvement of students in identifying standards and/or criteria to apply to their work and making judgements about the extent to which they met these criteria and standards*' (Boud 1991, 5). Our study however did not involve students in identifying standards or criteria. As our hypotheses primarily focus self-assessment validity (and not its metacognitive benefits), involving students in criteria selection might have induced bias on our measurement because the self-assessment would not measure the same concept throughout all study participants. To ensure comparability of the self-assessment throughout all subjects, we thus opted for a standardized self-assessment measure. Nevertheless, further research should focus on replicating our findings with different self-assessment instruments.

A third limitation concerns our sample: Even though we considered students of all semesters (e. g., seniors as well as freshmen), our sample is limited to psychology students. With a very high percentage of women, gender was distributed unevenly, too. Albeit typical for German psychology students (Wentura et al. 2013), such a distribution negatively affects the generalizability of our findings.

Finally, we have to acknowledge that the present article focuses just one single information literacy test. However, the amounts of explained variance by self-assessments may vary depending on the type of measures and tests administered. As a rather high convergent validity could be demonstrated for the PIKE-P test (see Rosman and Birke, forthcoming), self-report items would supposedly explain even more incremental variance if less sophisticated inventories are used. Nevertheless, further research should focus on replicating our findings with different measures, particularly with different self-assessment inventories (see above).

Implications for research and practice

In sum, our findings do support the existence of a small relationship between self-assessed information literacy and actual information literacy levels that is independent of subjects' test levels. We therefore recommend that researchers and practitioners include self-report measures into their test batteries. In fact, both objective and subjective approaches to assessment have their benefits: Conducting information literacy tests and search tasks might initiate learning processes leading to increases in ability. As calibration is positively related to ability (Kruger and Dunning 1999; Lew et al. 2010), the increased accuracy in the second self-assessment phase may also indicate that individual learning took place in the course of the study. With regard to the self-assessments, their positive effects might not only have led to an increase in calibration, but also to an increase in individual ability: Reflecting over their ability during the first phase of the self-efficacy assessment might have already strengthened participants' metacognitive skills, aided them to recognize their deficits, and thus induced learning processes (Boud 1995; Dochy et al. 1999). We acknowledge these two points being rather speculative, but nevertheless wish to point out that assessment itself can (and should) be seen as an important tool for learning (Dochy et al. 1999). From a methodological point of view, as self-reports have been shown to explain incremental variance in information literacy, aggregating both objective and subjective scores (or including them in a multiple regression) might also be a valuable method to strengthen assessment precision and validity. However, we wish to indicate that one should not expect too much from including self-report items, as the amounts of incrementally explained variance are rather small. But nevertheless, self-assessment items are a quick and easy way to explain a few more percent of variance in a multi-faceted concept that is so difficult to measure comprehensively.

As for the ‘how’ of including self-assessments, we first suggest that one should draw on standardized or otherwise established measures, like the SES-IB-16 for information literacy self-efficacy (Behm, forthcoming). Second, our findings show that test order plays a crucial role: We recommend that self-assessments (e. g., self-efficacy) are administered at the end of the respective test battery, ideally preceded by tasks in which subjects have to conduct information searches themselves. These tasks would not even have to be scored, as their sheer administration already enhances self-report validity. If the administration of information search tasks is not possible (e. g., in paper-pencil testing), self-assessments should nevertheless be administered *after* ability tests.

To summarize, we claim that self-assessments can be used to enhance precision and broaden the scope of many information literacy assessment approaches, given that they are measured at the right moment and with established measures. However, we would like to issue a note of caution at this point: Correlations between self-reported and actual ability tend to be rather small, and our findings show that standardized tests have a much higher predictive value (see also Rosman et al. 2014). Therefore, we urge researchers and practitioners *not* to assess information literacy with self-reports alone. As stated earlier, this practice is still prevalent and our study as well as the valuable research by Gross and Latham (2007; 2009) demonstrate that it is by no means an alternative to standardized tests.

References

Ackerman, P. L., and S. D. Wolman. 2007. “Determinants and Validity of Self-Estimates of Abilities and Self-Concept Measures.” *Journal of Experimental Psychology: Applied* 13 (2): 57-78.

American Library Association. 1989. *Presidential Committee on Information Literacy. Final Report.* Chicago: American Library Association.

Bandura, A. 1994. *Self-Efficacy.* New York: Wiley.

Bandura, A. 2010. “Self-Efficacy.” In *The Corsini Encyclopedia of Psychology,* edited by I. B. Weiner and W. E. Craighead, 1534-1536. New York: Wiley.

Behm, T. forthcoming. “Information Literacy and Self-regulation: The Significance of Domain-Specific Self-Efficacy Expectations.” [Informationskompetenz und Selbstregulation: Zur Relevanz bereichsspezifischer Selbstwirksamkeitsüberzeugungen.] In *Informationskompetenz im Hochschulkontext – Interdisziplinäre Forschungsperspektiven*, edited by A.-K. Mayer. Lengerich, Germany: Pabst Science Publishers.

Benton, S. L., D. Duchon, and W. H. Pallett. 2013. “Validity of Student Self-Reported Ratings of Learning.” *Assessment & Evaluation in Higher Education* 38 (4): 377-388.

Betz, N. E., and K. L. Klein. 1996. “Relationships Among Measures of Career Self-efficacy, Generalized Self-Efficacy, and Global Self-Esteem.” *Journal of Career Assessment* 4 (3): 285-298.

Boud, D. 1991. *Implementing Student Self-Assessment (2nd ed.).* Sydney: Higher Education Research and Development Society of Australasia.

Boud, D. 1995. *Enhancing Learning through Self-Assessment.* London, UK: Kogan Page.

Boud, D., and N. Falchikov. 1989. “Quantitative Studies of Student Self-Assessment in Higher Education: A Critical Analysis of Findings.” *Higher Education* 18 (5): 529-549.

Boud, D., R. Lawson, and D. G. Thompson. 2013. “Does Student Engagement in Self-Assessment Calibrate Their Judgement Over Time?” *Assessment & Evaluation in Higher Education* 38 (8): 941-956.

Bowles-Terry, M. 2012. “Library Instruction and Academic Success: A Mixed-Methods Assessment of a Library Instruction Program.” *Evidence Based Library and Information Practice* 7 (1): 82-95.

Brackett, M. A., and J. D. Mayer. 2003. "Convergent, Discriminant, and Incremental Validity of Competing Measures of Emotional Intelligence." *Personality and Social Psychology Bulletin* 29 (9): 1147-1158.

Bruce, C. 2004. "Information Literacy as a Catalyst for Educational Change. A Background Paper." In *Lifelong Learning: Whose responsibility and what is your contribution?: Proceedings of the 3rd International Lifelong Learning Conference*, edited by P. A. Danaher, 8-19. Yeppoon, Queensland: Queensland University of Technology.

Chang, Y. K., X. Zhang, I. A. Mokhtar, S. Foo, S. Majid, B. Luyt, and Y. L. Theng. 2012. "Assessing Students' Information Literacy Skills in Two Secondary Schools in Singapore." *Journal of Information Literacy* 6 (2): 19-34.

Coupe, J. 1993. "Undergraduate Library Skills: Two Surveys at Johns Hopkins University." *Research Strategies* 11 (4): 188-201.

Dochy, F. J. R. C., M. Segers, and D. Sluijsmans. 1999. "The Use of Self-, Peer and Co-Assessment in Higher Education: A Review." *Studies in Higher Education* 24 (3): 331-350.

Falchikov, N. 2005. *Improving Assessment Through Student Involvement: Practical Solutions for Aiding Learning in Higher and Further Education.* London: RoutledgeFalmer.

Falchikov, N. and D. Boud. 1989. "Student Self-Assessment in Higher Education: A Meta-Analysis." *Review of Educational Research* 59 (4): 395-430.

Freund, P. A., and N. Kasten. 2012. "How Smart Do You Think You Are? A Meta-Analysis on the Validity of Self-Estimates of Cognitive Ability." *Psychological Bulletin* 138 (2): 296-321.

Gapski, H., and T. Tekster. 2009. *Information Literacy in Germany. [Informationskompetenz in Deutschland.]* Düsseldorf, Germany: Landesanstalt für Medien Nordrhein-Westfalen (LfM).

Geffert, B., and B. E. Christensen. 1998. "Things They Carry: Attitudes Toward, Opinions About, and Knowledge of Libraries and Research Among Incoming College Students." *Reference & User Services Quarterly* 37 (3): 279-289.

Gross, M., and D. Latham. 2007. "Attaining information literacy: An investigation of the relationship between skill level, self-estimates of skill, and library anxiety." *Library & Information Science Research* 29 (3): 332-353.

Gross, M., and D. Latham. 2009. "Undergraduate Perceptions of Information Literacy: Defining, Attaining, and Self-Assessing Skills." *College & Research Libraries* 70 (4): 336-350.

Heinze, N., J. Fink, and S. Wolf. 2009. *Information Literacy and Research: Empirical Results and Recommendations on Academic Information Searches in Higher Education. [Informationskompetenz und wissenschaftliches Arbeiten: Studienergebnisse und Empfehlungen zur wissenschaftlichen Recherche im Hochschulstudium.]* Augsburg, Germany: University of Augsburg.

Ivanitskaya, L., I. O'Boyle, and A. M. Casey. 2006. "Health Information Literacy and Competencies of Information Age Students: Results From the Interactive Online Research Readiness Self-Assessment (RRSA)." *Journal of Medical Internet Research* 8 (2): e6. doi:10.2196/jmir.8.2.e6.

Jackson, D. 2014. "Self-Assessment of Employability Skill Outcomes Among Undergraduates and Alignment With Academic Ratings." *Assessment & Evaluation in Higher Education* 39 (1): 53-72.

Johnston, B., and S. Webber. 2003. "Information Literacy in Higher Education: A Review and Case Study." *Studies in Higher Education* 28 (3): 335-352.

Kim, J. 2009. "Describing and Predicting Information-Seeking Behavior on the Web." *Journal of the American Society for Information Science and Technology* 60 (4): 679-693.

Kruger, J., and D. Dunning. 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.

Kurbanoglu, S. S., B. Akkoyunlu, and A. Umay. 2006. "Developing the Information Literacy Self-Efficacy Scale." *Journal of Documentation* 62 (6): 730-743.

Leach, L. 2012. "Optional self-assessment: some tensions and dilemmas." *Assessment & Evaluation in Higher Education* 37 (2): 137-147.

Leichner, N., J. Peter, A.-K. Mayer, and G. Krampen. 2014. "Assessing Information Literacy Using Information Search Tasks." *Journal of Information Literacy* 8 (1): 3-20.

Lew, M. D., W. A. M. Alwis, and H. G. Schmidt. 2010. "Accuracy of Students' Self-Assessment and Their Beliefs About its Utility." *Assessment & Evaluation in Higher Education* 35 (2): 135-156.

Maughan, P. D. 2001. "Assessing Information Literacy Among Undergraduates: A Discussion of the Literature and the University of California-Berkeley Assessment Experience." *College & Research Libraries* 62 (1): 71-85.

Neely, T. Y. 2006. *Information Literacy Assessment: Standards-Based Tools and Assignments.* Chicago: American Library Association.

Oscarson, M. 1989. “Self-assessment of language proficiency: Rationale and applications.” *Language Testing* 6 (1): 1-13.

Pinto, M. 2010. “Design of the IL-HUMASS Survey on Information Literacy in Higher Education: A Self-Assessment Approach.” *Journal of Information Science* 36 (1): 86-103.

Radcliff, C. J., M. L. Jensen, J. A. Salem, K. Burhanna, and J. A. Gedeon. 2007. *A Practical Guide to Information Literacy Assessment for Academic Librarians.* Westport, CT: Libraries Unlimited.

Rosman, T., A.-K. Mayer, and G. Krampen. 2014. “Measuring Psychology Students’ Information-Seeking Skills in a Scenario-Based Test Format: Construction and Validation of the PIKE-P.” Manuscript under review, Leibniz Institute for Psychology Information, Trier, Germany.

Rosman, T. and P. Birke. forthcoming. “Discipline-Specific Assessment of Information-Seeking Ability With Procedural Knowledge Tests: Psychology vs. Computer Science.” [Fachspezifische Erfassung von Recherchekompetenz durch prozedurale Wissenstests: Psychologie vs. Informatik.]. In *Informationskompetenz im Hochschulkontext – Interdisziplinäre Forschungsperspektiven,* edited by A.-K. Mayer. Lengerich, Germany: Pabst Science Publishers.

Scharf, D., N. Elliot, H. A. Huey, V. Briller, and K. Joshi. 2007. “Direct Assessment of Information Literacy Using Writing Portfolios.” *Journal of Academic Librarianship* 33 (4): 462-477.

Schunk, D. H. 1984. “Self-Efficacy Perspective on Achievement Behavior.” *Educational Psychologist* 19 (1): 48-58.

Schwarzer, R., J. Bäßler, P. Kwiatek, K. Schröder, and J. X. Zhang. 1997. „The Assessment of Optimistic Self-Beliefs: Comparison of the German, Spanish, and Chinese Versions of the General Self-efficacy Scale.” *Journal of Applied Psychology* 46 (1): 69-88.

Taras, M. 2001. “The Use of Tutor Feedback and Student Self-Assessment in Summative Assessment Tasks: Towards Transparency for Students and for Tutors.” *Assessment & Evaluation in Higher Education* 26 (6): 605-614.

Vygotsky, L. S. 1962. *Thought and language.* Cambridge: MIT Press.

Walsh, A. 2009. “Information Literacy Assessment: Where Do We Start?” *Journal of Librarianship and Information Science* 41 (1): 19-28.

Wentura, D., M. Ziegler, A. Scheuer, J. Bölte, T. Rammsayer, and C. Salewski. 2013. „German Nationwide Survey of Bachelor's Degree Graduates in Psychology in 2011.“

[Bundesweite Befragung der Absolventinnen und Absolventen des Jahres 2011 im Studiengang BSc Psychologie.] *Psychologische Rundschau* 64 (2): 103-112.

Wise, S. L., L. Cameron, S. T. Yang, S. L. Davis, and J. Russell. 2009. *The Information Literacy Test (ILT): Test Manual.* Harrisonburg: The Center for Assessment and Research Studies.

Tables

Table 1. Means, standard deviations, internal consistencies, and intercorrelations of study variables.

Scale	***M***	***SD***	**1**	**2**	**3**	**4**	**5**	**6**
1 SES-IB-16_T1	3.36	0.50	**(.87)**					
2 SES-IB-16_T2	3.17	0.53	.67***	**(.87)**				
3 SES-IB-16_DIFF	-0.19	0.42	-.34***	.47***	**(.60)**			
4 PIKE-P	53.06	9.82	.02	.15*	.16*	**(.72)**		
5 Information search tasks – Outcome rubric	6.53	2.15	-.09	.13	.28**	.62***	**(.59)**	
6 Information search tasks – Process rubric	6.34	2.16	.08	.30**	.29**	.64***	.66***	**(.75)**

Note. N = 81; SES-IB-16_T1 = Information literacy self-efficacy at the first measurement point; SES-IB-16_T2 = Information literacy self-efficacy at the second measurement point; SES-IB-16_DIFF = Difference between SES-IB-16_T2 and SES-IB-16_T1; PIKE-P = Procedural Information-Seeking Knowledge Test for Psychology Students; *M* = mean; *SD* = standard deviation; values in bold on the diagonal = Cronbach's Alpha.

* $p < .05$. ** $p < .01$. *** $p < .001$.

Table 2. Hierarchical Regression Analyses predicting PIKE-P and information search task scores from the difference between the second and the first measurement point of information literacy self-efficacy.

	PIKE-P			**Information search tasks: Outcome rubric**			**Information search tasks: Process rubric**		
	β	R^2	ΔR^2	β	R^2	ΔR^2	β	R^2	ΔR^2
Block 1		.00			.01			.01	
SES-IB-16_T1	.02			-.09			.08		
Block 2		$.03^{+}$	$.03^{+}$		$.08^{**}$	$.07^{**}$		$.12^{***}$	$.11^{***}$
SES-IB-16_T1	.09			.00			$.20^{*}$		
SES-IB-16_DIFF	$.19^{+}$			$.28^{**}$			$.36^{***}$		

Note. N = 81; Method: Enter; SES-IB-16_T1 = Information literacy self-efficacy at the first measurement point; SES-IB-16_DIFF = Difference between the second and the first measurement point of SES-IB-16; PIKE-P = Procedural Information Literacy Knowledge test for psychology students; β = standardized regression weight; R^2 = total variance explained; ΔR^2 = change in R^2 from block 1 to block 2.

$^{+}$ p < .10; * p < .05. ** p < .01. *** p < .001.

Table 3. Hierarchical Regression Analyses predicting information search task scores from information literacy test scores and information literacy self-efficacy.

	Information search tasks: Outcome rubric			Information search tasks: Process rubric		
	β	***R²***	***ΔR²***	***β***	***R²***	***ΔR²***
Block 1		$.39^{***}$			$.41^{***}$	
PIKE-P	$.62^{***}$			$.64^{***}$		
Block 2		$.39^{***}$	.00		$.45^{***}$	$.05^{*}$
PIKE-P	$.62^{***}$			$.61^{***}$		
SES-IB-16	.04			$.22^{*}$		

Note. N = 81; Method: Enter; PIKE-P = Procedural Information Literacy Knowledge test for psychology students; SES-IB-16 = Information literacy self-efficacy; β = standardized regression weight; R^2 = total variance explained; ΔR^2 = change in R^2 from block 1 to block 2.

* $p < .05$. ** $p < .01$. *** $p < .001$.