Media literacy evaluation: a critical discussion of reliability and usefulness issues

Pierre Fastrez, Camille Tilleul and Thierry De Smedt, Université catholique de Louvain

Abstract

Keywords: Competence, media literacy, assessment, methodology

Abstract

This article discusses three issues related to the reliability and usefulness of media literacy evaluation, based on our experience of a two-year longitudinal assessment of the competences of students enrolled in a new master's degree in Media Literacy Education, in a French-speaking Belgian college. The three issues we discuss are:

1. If competent individuals are able to face a variety of novel in-context problems, how can their competences be evaluated from their performance on a limited set of controlled tasks?
2. In longitudinal within-subject comparisons, how can one strike a balance between modifying the tasks between tests to avoid practice effects, and maintaining the comparability of results?
3. How can one facilitate the appropriation of the assessment results by the community of the curriculum's teachers?

   The article will discuss the respective merits of the answers to these three questions formulated as a part of our own project and of other notable media literacy assessment studies.

The article will discuss the respective merits of the answers to these three questions formulated as a part of our own project and of other notable media literacy assessment studies.

Introduction

Historically, media education first emerged, some forty years ago, as a domain of pedagogical activity, as pioneer educators envisioned the need for youngsters to develop critical thinking abilities towards mass media. Led by their intuitions, these educators

invented a broad cluster of educational activities that combined the appropriation of concepts borrowed from semiotic, literacy and narrative theories with the production of media such as school newspapers. They evaluated these activities by observing the degree to which youngsters engaged in them and committed to them.

As media education slowly grew into an array of public and private (mostly associative) institutions, its nascent pedagogical approaches and core themes were articulated around the overall ambition of enabling everyone to appropriate all contemporary media both critically and creatively (Jolls & Wilson, 2014). On the one hand, the goal was set to support media users in distancing themselves from their ordinary media experience to reflect on the role and place of media in their lives and in society (Masterman, 1985). On the other hand, the objective was to enable people to appropriate the languages, technologies and social implications of media as media producers. In doing so, media education aimed at preparing individuals to be active, autonomous, critical and responsible citizens, able to use media to contribute to the development of a democratic society (Kellner & Share, 2007; Mihailidis & Tevenin, 2013).

As media education gradually became institutionalized as a public educational project in different countries, the necessity to convert this grand vision of an active and reflective, media-savvy citizen into measurable indicators appeared more and more clearly to decision makers as well as to educators, for several reasons. First and foremost, the development of the media literacy of individuals gradually became a desirable societal goal, as it is now seen as a key condition of the citizens' social participation. As the political world embraced this notion as a project, it created the need for means of reporting on the levels of media literacy of different social groups (from entire national populations to specific categories of individuals). Furthermore, the increasing integration of media education, initially developed by educators, into formal education curricula called for the development of competence frameworks that set standards for the evaluation of the efficiency of media education in developing the students' media literacy. For example, in Finland, both the National Core Curriculum for Basic Education (2004) and the National Core Curriculum for Upper Secondary Schools (2003) include a media skills and communication and a communication and media competence section (respectively) that are based on the general Aspen definition of media literacy, which included four components: access, analyze, evaluate, and produce (Tyner, 1998: 120). In 2006, the French Ministry of Education (2013) launched a certificate called 'B2i' for Brevet Informatique et Internet. This certificate is comparable with the European Computer Driver's Licence. The B2i has two levels, the first at the end of the primary school and the second at the end of middle school. It covers a range of practical competences based on the capacity

of the student to operate computers and the Internet in a way that preserves the quality of information, the rights of the authors and the user's privacy. In 2013, the Conseil supérieur de l'éducation aux médias of the French-speaking community of Belgium adopted a general framework for media education in French-speaking Belgium (CSEM, 2013), distinguishing between four media competences (reading, writing, navigating, organizing) on media considered as informational, technical and social objects. The Dutch competencies model, adopted in 2013, consists of four groups of competencies, centered on four specific media user positions: passive (competence to understand), active (competence to make use), interactive (competence to exchange with other users) and effective (competence to act with media).

The many uses of media literacy assessment

One common point of the frameworks presented above (and many others) is that they rely on a definition of media literacy in terms of competences, either explicitly or implicitly. Although the definition of the concept of competence may be subject to discussion, the following traits can be outlined. An individual is deemed competent when he or she is able to intentionally take relevant action in the face of the unexpected. One may speak of competence when individuals put their knowledge and skills to relevant use in order to adapt to novel situations in original and non-stereotypical ways (Scallon, 2004: 11; Rey et al., 2012: 13). Hence, being competent implies the efficient articulation of different internal and external resources within a family of situations that share some degree of similarity (Tardif, 2006: 15). Competences are distinct from skills, which are acquired through simple and repetitive practice that allows an individual to learn to reproduce a given behavior in a constant, invariable way (e.g. the ability to use a computer mouse, or to launch a web browser). In contrast, competences are related to the management of novel situations, and include a metacognitive dimension (Kruger and Dunning 1999; Westera 2001): competent individuals are able to monitor their own judgment, knowledge or actions, and evaluate their accuracy and relevance. In the context of media education, Piette (2000) argued that critical thinking appeared to be tightly linked to the consciousness one has of the process that undermines the exercise of their own thought (or that of others), thereby making metacognitive skills the keystone of the educational project of media education (i.e. to foster the autonomous critical thought of the pupils).

As for the evaluation of any kind of competence, assessing media literacy is a difficult endeavor that poses several methodological challenges. This article discusses three issues related to the reliability and usefulness of media literacy assessments. This discussion is grounded in our experience of a two-year longitudinal assessment of the competences

of students enrolled in a new master's degree in Media Literacy Education, in a French-speaking Belgian college. For each issue, we will discuss the respective merits of the solutions designed as a part of our own project and of other notable media literacy assessment studies.

As these solutions partly depend on the intended use of the assessment, it may be worth clarifying our own position within the array of such intended uses, which we will sort into three categories. A first category consists in experimental studies that seek to validate media literacy assessment instruments that were designed as part of the study (e.g. Arke & Primack, 2009) by using them in a pilot study, to establish their validity before using them in subsequent evaluations. A second category uses assessment instruments to report on the media literacy levels of a given group or population (e.g. EAVI 2009, Danish Technological Institute 2011, or in the context of digital skills: van Deursen et al. 2014; van Deursen & van Dijk 2010), and sometimes to correlate these levels with other descriptive variables (e.g. van Deursen, Courtois & van Dijk 2014; see Litt 2013 for a review on this topic). A third category uses media literacy assessment to evaluate the effects of a given media education program on the competences of the students who followed it. In most cases, the program under scrutiny is taken as a token for a type of media education pedagogy, and the results of the evaluation are interpreted as evidence of the efficiency of this specific type of media education (e.g. Hobbs & Frost, 2003, Quin & McMahon 1993). Albeit belonging to the third category, the two-year assessment we are reporting on had a slightly different focus. As part of a larger project that aimed at monitoring and evaluating the whole process of implementation of the Master's Degree in Media Literacy Education, the assessment of its students' competences was used as a means of informing the optimization of the curriculum.

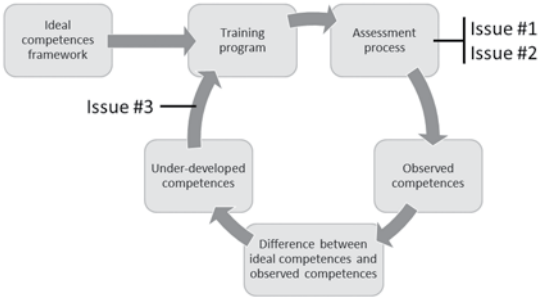

Figure 1: The integration of Media Literacy assessment into Media Education pedagogy

The role of the assessment in this process is depicted in Figure 1: the curriculum (or

training program) was based on a framework defining the ideal competences that students graduating from the Master's Degree should master. The assessment of the students' competences was designed to expose the difference between these ideal competences and the actual levels of competence of the students, leading to the identification of under-developed competences, and to the revision of the curriculum. The three issues we will be discussing in the remainder of this paper pertain to the assessment process itself (issues #1 and 2) and to the use of its results to revise the training program. These three issues we discuss are:

1. If competent individuals are able to face a variety of novel in-context problems, how can their competences be evaluated from their performance on a limited set of controlled tasks?
2. In longitudinal within-subject comparisons, how can one strike a balance between modifying the tasks between tests to avoid practice effects, and maintaining the comparability of results?
3. How can one facilitate the appropriation of the assessment results by the community of the curriculum's teachers?

The remainder of this article will first offer a description of the method we used in this two-year longitudinal assessment. Next, we will discuss the three issues presented above. Finally, we will propose some perspectives on future media literacy assessments as a short conclusion.

Context: a two-year longitudinal assessment of the competences of students in Media Literacy Education

Over the course of two academic years, we undertook a longitudinal assessment of the competences of the two first cohorts of students enrolled in a Master's program in Media Literacy Education. The assessment was independent from the evaluations included in the Master's program (i.e. exams, projects and homeworks that were part of its courses and seminars). It encompassed both their media literacy and their media education competences. Figure 2 summarizes the timeline for the assessment.

The first cohort (N = 13) of students participated in the entire evaluation. They were tested three times: at the onset of their curriculum, after a year of training, and during the finals of their second year. The second cohort (N = 18) was tested two times. Each test lasted approximately two hours, and was screen-captured on a computer, using Techsmith Morae.

The assessment included five tasks, connected by a common scenario: (1) one media education task: planning a media education class activity; and four media literacy tasks: (2)

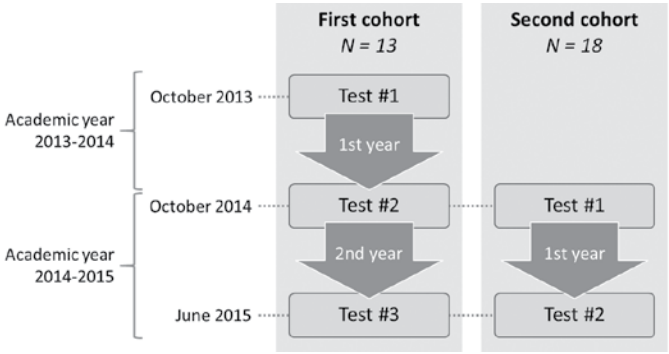

Figure 2: Timeline of the assessment

a navigation task: searching and selecting media documents from a database and from internet in preparation of the class activity ; (3) an organization task: categorizing and organizing the selected media documents; (4) a reading task: reading and comparing two media documents; and (5) a writing task: designing a new media document in the context of extending the class activity they designed.

The four media literacy tasks are based on a matrix model of media literacy, defined as a set of competences (Fastrez 2010). In this model, media literacy is defined as the competences required to perform four types of media tasks (reading, writing, navigating and organizing the media) on media taken from three different perspectives (media as informational, technical and social objects). Table 1 presents the twelve dimensions of media literacy defined by this matrix.

Media as... Tasks:	Informational objects	Technical objects	Social objects
Reading	Reading media according to its informational dimensions	Reading media according to its technical dimensions	Reading media according to its social dimensions
Writing	Writing media according to its informational dimensions	Writing media according to its technical dimensions	Writing media according to its social dimensions
Navigating	Exploring, searching, and finding media according to their informational dimensions	Exploring, searching, and finding media according to their technical dimensions	Exploring, searching, and finding media according to their social dimensions
Organizing	Organizing media according to their informational dimensions	Organizing media according to their technical dimensions	Organizing media according to their social dimensions

Table 1: A matrix definition of media literacy

For the longitudinal assessment, each of the five tasks was evaluated on several dimensions, detailed in Table 2.

<table>
<tr><th>Tasks</th><th colspan="3">Evaluated dimensions</th></tr>
<tr><td>Media education task (planning a pedagogical activity)</td><td colspan="3">Ability of the student to define, discuss or describe their planned activity in terms of...
Diagnosis of the educational problem to be addressed
Objectives of the activity
Tools for assessing the educational effects of the activity
Pedagogies used by the activity
Technical means for the activity
Human resources
Timing and scheduling
Financial implications</td></tr>
<tr><td>Media Literacy tasks</td><td colspan="3">Ability of the student to describe, compare, search for, categorize or design media in terms of their...</td></tr>
<tr><td>Navigating media
Organizing media
Reading media
Writing media</td><td><u>informational dimensions:</u>
topic of the media
languages and representations
form of the media</td><td><u>technical dimensions:</u>
techniques used to produce the media
technical qualities of the media</td><td><u>social dimensions:</u>
author(s) of the media
intentions of the author(s)
intended audience of the media
potential effects of the media</td></tr>
</table>

Table 2: Dimensions of the assessment

Each of these dimensions was rated in on three-point scale (see Table 3):

Rating	Rationale
2	The student's spontaneous answer includes the dimension, and treats it thoroughly, accurately, and consistently with the other dimensions.
1	The student's spontaneous answer includes the dimension, but treats it only partially or is inconsistent with respect to the other dimensions or the student considers the dimension only after being prompted by the interviewer.
0	The student is unable to consider the dimension even after being asked to do so.

Table 3: Rating scale for the dimensions of the assessment

The final results of this study showed that students generally presented a higher level of competence in media education and in medial literacy by the end of their curriculum. However, they did not appear to make substantial progress in two specific dimensions. First, they struggled to consider the technical and financial implications of the media education activities they planned and of the media they analyzed and produced as part of the tests. Second, they had trouble identifying the potential audiences of these media, as well as their potential effects on their audiences.

As the results of this two-year assessment were and will be published in separate papers (see e.g. Tilleul, Fastrez and De Smedt, 2015), they will not be detailed any further here. We now turn to the discussion of the first methodological issue related to the process of assessing media literacy.

Competence: documenting the unobservable

Media literacy is a very broad theoretical construct that spans multiple domains of experience with media and technology. Although it is appealing to consider it as a whole and examine how different dimensions of this construct may be related (e.g. production and evaluation abilities, informational and social abilities), it poses a serious challenge on the overall assessment of this construct. Defined as an extended set of competences (cf. supra), media literacy has a transversal character that is difficult to document either through the general survey questions or through the specific, limited tasks used in assessment instruments.

A competence, by nature, cannot be observed directly. It is a permanent, invisible potentiality of the individual, so that a competent person stays competent even when they are inactive (Scallon, 2004: 105). By contrast, performances are situated actions that are the concrete manifestation of competences, which makes them the most direct source of data from which competences can be inferred. When the observation of performance is not possible, an alternative consists in relying on the individual's discourse on their level of competence. Performance observation and discourse solicitation correspond to two different approaches towards the assessment of the media literacy levels of a given population or sample: tests presenting subjects with media-related tasks, and survey questionnaires including self-assessment questions. Both methods are limited in the way they can account for the transversality and versatility of media literacy, defined as a set of competences. We will now briefly review the respective limits of these two approaches, starting with proxy survey measures (i.e. self-reported measures), and following with performance-in-task measures.

1. Declarative assessment: level of generality of statements vs. validity of indicators
   Self-reported measures are widely used in media literacy or digital literacy assessment surveys (e.g. Brandtweiner et al. 2010; Hargittai 2005, 2008; Hargittai & Walejko 2008; Literat 2014; Phang & Schaefer 2009; van Deursen, Helsper, Enyon 2014). They typically require respondents to mark their degree of agreement with statements pertaining to their media-related knowledge (I know what x is), skills (I am able to do x), attitudes (I feel like x), and practices (I do x frequently).
   For example, in order to establish the validity of survey measures of internet skills,

Hargittai (2005, 2008) asked internet users (1) to state whether they knew how to perform different internet-related tasks, (2) to rate their own familiarity with internet-related terms, and (4) to give an overall rating of their internet skills. As part of a larger survey on digital skills, van Deursen, Helsper, Enyon (2014) asked respondents to rate the degree to which different statements related to these skills were accurate when thinking about they used the internet. In their assessment of the media literacy of Singaporean students, Phang and Schaefer (2009) asked respondents to report on their awareness of and/or command of the skills required to evaluate and create video content.

Self-assessment surveys (i.e. asking the respondent how competent they feel in a given domain) have at least two advantages. On the one hand, they are lightweight in terms of time and resources needed to implement and them and process the data they yield. On the other hand, they are not limited to one specific instance of competence use, as respondents can be prompted to reflect on their overall competence level with media across many situations.

However, the relationship between self-assessed and actual competence levels has proven to be prone to distortions. In particular, less skilled individuals seem to overestimate their abilities in many social and intellectual domains, as their lack of competence deprives them from the kind of metacognitive skills that would allow them to recognize their own incompetence (Kruger & Dunning, 1999). Moreover, the quality of the assessment appears to decrease as the generality of the statement increases (Ackerman et al. 2002). In other words, self-assessments seem to be more accurate when they focus on specific tasks and abilities. For example, Hargittai (2005) found that an overall rating of self-perceived internet skill level (i.e. responding to 'In terms of your Internet skills, do you consider yourself to be... not at all skilled / not very skilled / fairly skilled / very skilled / expert') was a much weaker predictor of actual skill level than an index constructed from seven self-reported ratings of the subject's knowledge of specific internet-related terms.

Essentially, this means that while the ability to assess media literacy through generic, transversal indicators can be seen as an advantage of self-assessments, it is largely mitigated by the increasing bias the instrument introduces in the assessment as the indicators become more and more general.

2. Observing performance in tests: assessing an abstract concept through situated performances

Compared to self-reported measures of media literacy, assessments that rely on the observation of performance in test tasks are less biased, but more demanding. In

addition, this type of assessment faces a paradox that comes from the combination of two facts, that were partly stated earlier: (1) competences are transversal: by nature, competent individuals can deal with a variety of different problems and situations, (2) competences can only be reliably established through in-situation performances. Hence, confronting the individual with a one test task puts them in one situation, and successful performance in that situation does not guarantee 'similarly successful performance in a new, but substantially different, complex situation. (Westera, 2001: 82). In other words, test results only ever provide partial evidence of one's abilities in a limited context, which cannot be considered a proof of competence.

Still, if competence cannot be definitely proven, the next best solution seems to document how individuals are able to face a variety of tasks, situations and contexts. Different studies have taken different approaches towards introducing variation and diversity in the assessment of media literacy. In a non-comprehensive attempt to review existing media literacy assessment studies involving tests and performances, we identified four criteria along which tasks vary in the reviewed studies. A first criterion is the type of tasks or activities individuals are subjected to. For example, test tasks can focus on reception (e.g. viewing video segments, reading print articles, listening to radio reports) or production activities (e.g. video editing, website or blog development, document writing and composition). A second criterion is related to the type of media being involved in the tasks: print, audio and video records, electronic documents and platforms, etc. A third criterion has to do with the perspectives on media the task requires, from media as sources of information, to media as technical apparatus, to media as supporting social interaction, etc. Finally, tasks can be defined to match different levels of difficulty or complexity, from simple operations to open-ended problems.

The studies we reviewed can be placed into three families, according to how they combine the four criteria defined above. Table 4 summarizes this distinction.

Study Family	Tasks	Media	Perspectives	Levels
Mass-media analysis	constant	variable	variable	constant
Search and navigation	constant	constant	constant	variable or constant
Overall ML / New ML	variable	variable	variable	constant

Table 4: Families of media literacy assessment studies

The first family corresponds to studies seeking to assess media literacy in terms of mass-media analysis competences. Studies in this group all involve several similar tasks, which require students to read, view or listen to various selected media, and analyze them from different perspectives. The complexity of the tasks is constant across media. Typical

examples of selected media include news reports (Hobbs & Frost 2003; Arke & Primack 2009), advertisements (Hobbs & Frost 1999; Quin & McMahon 1993), or fictions (Quin & McMahon 1993) in print, audio or video format. The perspectives applied to these media mostly relate to their semiotic construction and social entailments. For example, Hobbs and Frost (2003) asked students (among other things) to identify the purpose and target audience of three news reports, as well as the construction techniques they used, the values and points of view they presented, and the information they omitted. Arke & Primack (2009) used most of these categories, as well as the attitudes and feelings the media under scrutiny left the subject with. The tests used by Quin & McMahon (1993) required students to analyze ads in terms of the languages they used, and a sitcom segment in terms of its narrative construction.

The second family contains studies pertaining to information search and internet navigation competences. These studies limit themselves to a specific type of task (searching for information and evaluating found information). The media of choice is also constant: the World Wide Web (search engines and websites), and so is the perspective on this media: internet as an informational resource, used by subjects to answer selected questions. Some studies in this category use tasks of constant complexity. For example, in one of the first experiment of this kind, Hargittai assessed fifty-four internet users' 'ability to efficiently and effectively find information on the Web' (Hargittai 2002: 2) by asking them to find information about local cultural events, music they could listen to online, a comparison of different US presidential candidates' views on abortion, tax forms, and art by kids. In other cases, tasks that vary in difficulty are presented to the subjects. This is the case for studies published by van Deursen & van Dijk (2010) and Gui & Argentin (2011), in which tasks presented to internet users range from simple computer operations and basic interactions with hypermedia to assignments involving information evaluation, comparison and synthesis in open-ended contexts such as informing oneself before voting or planning a vacation.

The third family of studies takes a broader approach towards the assessment of media literacy. Such studies use different types of tasks on different types of media, considered from different perspectives, in an attempt to cover the variety of media literacy competences. A first example is the study by Eshet-Alkali and Amichai-Hamburger (2004), who subjected sixty media users to five independent tests related to different aspects of media literacy: photo-visual literacy (task: creating a theatre stage using a computer program), reproduction literacy (task: creating new meaning by rearranging the sentences in a text passage), information literacy (task: analyzing a news event reported in seven sources); branching literacy (task: plan a trip to Spain by searching the web), and socio-

emotional literacy (task: having a conversation in an online chat room). The design of these five tests does not allow for the exploration of the relationship between the different types of media literacy they seek to assess.

Our second example in this category is our own work on the media literacy competences of future media educators. As stated earlier, our assessment instrument included five different tasks, four of which referred to one of the dimensions of media literacy in our theoretical model (reading, navigating, organizing and writing). These four tasks confronted students with different types of media (from news reports to ads, to games, to blog posts, etc.), and required them to consider them under different (informational, technical and social) perspectives.

Compared to studies from the first two categories, which focus on one type of activity (mass-media analysis, or internet search), these two last examples have the advantage of providing a broader picture of the participants' media literacy, but the disadvantage of relying on coarse indicators.

Another way of comparing the studies from the three families is to consider the degree to which they consider the relationship between tasks, media, perspectives, and levels of complexity. From this point of view, the reviewed studies can be ordered on a continuum. At one end are tests that are analytical and additive by nature. Such tests either focus on one specific task, media and perspective combination (cf. second family), or use a collection of such combinations that are independent from one another (Eshet-Alkali & Amichai-Hamburger 2004). Halfway through the continuum are tests that apply the same set of perspectives on media across different media (cf. first family). Such tests can document how individuals master these perspectives in multiple reception contexts. Still, there seems to be room for the development of such types of tests outside the scope of critical reception tasks. Finally, at the other end of the continuum are synthetic tests that, like our own instrument, logically articulate different tasks, media, and perspectives on media, providing a multidimensional assessment. Although it only provides one problem-situation for each type of task, this method has the advantage of allowing for multiple comparisons between tasks (e.g. how is the social dimension of media mastered by student when they read, write, navigate or organize media) and between perspectives (e.g. how do informational, technical and social aspects of navigation relate?) across multiple media. These comparison appears to us as a relevant way of assessing the transversality of media literacy. However, this approach does not integrate the relative assessment of tasks with different levels of complexity.

Longitudinal assessments: comparability and practice effects

The second methodological issue related to media literacy assessment that we wish to discuss pertains to longitudinal assessments, which imply that participants be subjected to tests at different successive times. When the assessment takes place in the context of the evaluation of a media education program, each test takes place at a different stage of the participants' training. This constraint poses the question of finding the balance between a conceptual approach of competence assessment, which would require to design the successive tests so that they confront students with substantially different novel situations (Scallon, 2004; Rey et al., 2012) and an operational approach requiring that students be subjected to significantly similar tests to ensure the comparability of results and the possibility of identifying changes in competence levels.

In addition, managing the degree of similarity between successive tests implies dealing with the possibility of two types of opposite effects, both of which can jeopardize the validity of the assessment results. The first type corresponds to practice effects, when students perform better the second or third time because they know the task from having completed it before (and not because they are more competent). The second type correspond to the fact that differences in the situations, problems or materials used in each test may cause variations in the students' performances between tests.

The dominant strategy we found in the literature for avoiding these effects is to control the observational or experimental protocol. Numerous studies reporting on the evaluation of a given media education program combine identical pretest and posttest instruments, and include a comparison between a treatment and a control group (Hobbs and Frost, 2003; Gonzales et al., 2004; Pinkleton et al., 2006 ; Scharrer, 2006; Pinkham et al. 2008; Primack et al, 2014). Control groups are used to isolate program-related changes from other contextual effects. While identical pretests and post-tests control for effects of the specificities of the tasks, they introduce the risk of practice effects.

Among the studies cited above, only one (Pinkleton et al., 2006) controlled for practice effects by using additional posttest-only conditions: both the treatment and the control groups were divided in two conditions: one that underwent both the pretest and the post-test, and the other that only underwent the posttest. This type of protocol, called the Solomon four-group method, allows researchers to detect practice effects: if both treatment conditions (with or without pretest) perform better than the control conditions, one can infer 'that the treatment has an effect, regardless of any possible sensitization effects attributable to the pretest' (Pinkleton et al., 2006: 28). This method is rarely used in the media literacy literature, probably because it is costly.

While controlled approaches towards the assessment of media literacy such as the

ones described above offer valuable guarantees in terms of the validity of their measures, working with identical tests, control groups and post-test-only conditions is not always either possible, nor desirable. In the context of our own longitudinal assessment, the absence of any other similar degree in the French Community of Belgium, as well as the small number of enrolled students, precluded us from working either with a control group or with post-test-only conditions. Moreover, the need to account for the breadth of our students' media literacy (cf. issue #1) imposed the necessity to place them in different situations test after test.

As described earlier, students took part in three tests. Each test required student to perform the five same tasks (plan a media education lesson, read and compare two documents, navigate and select media through a database, organize this database and design a media) as part of a different integrated scenario each time. Each scenario introduced a novel situation and specific materials. Each time, students were evaluated on the same (informational, technical and social) dimensions. The pretest scenario related the five tasks to the theme of avatars on social networks, with an intended audience of 15-year-old students. For the halftime test, the scenario focused on the issue of advertising and the consumption of unhealthy products, with an intended audience of 12-year-old pupils. Finally, the post-test scenario dealt with conspiracy theories concerning the Charlie Hebdo attack in Paris, with a class of 18-year-old students.

In this context, we were indeed able to observe practice effects anecdotally, not from test to test, but from task to task within the same test. By the time they had reached the fourth task of their test (reading task), some students remembered the list of dimensions their answer had to cover (technical dimension, potential effects, audience, author, etc.), because the experimenter asked them to do so for previous tasks. Hence, they discussed these items spontaneously. Nevertheless, these spontaneous answers were seldom correct (or only partly correct), and failed to affect the students' results to their test.

We were also able to identify effects of the different scenarios on the students' results. For instance, students were better able to identify the author's intentions when the test scenario focused on advertising (a lot of the media documents involved were ads) rather than when it concerned conspiracy theories. In addition, some students were more comfortable with one situation than with another. One of the students, for example, was specifically working on the analysis of conspiracy theories for his honors thesis. He was thus more knowledgeable on the third scenario than on the two first ones.

These observations lead us to the important issue of how to choose the situations, scenarios and materials that compose tasks so that they provide students with constant (across test), but also fair and equal (between students) chances of performing and

showcasing their level of competence. If consistency in the complexity of situations that tasks use as a context for performance can be assured through a thorough analysis of the situations themselves, the equality of students in the face of these situations is difficult to establish, as not only does it depend on their knowledge and skills, but also on their motivation and beliefs regarding the situation, or their own abilities. On this specific issue, Scallon (2004) contends that factors such as motivations, beliefs and attitudes have either to be measured as part of the assessment, or have to be discussed with the students before the assessment, so as to identify and make use of situations for which students have similar motivations, beliefs or attitudes.

Conclusions: supporting a useful appropriation of assessment results by teachers

There are numerous papers reporting on the evaluation of (pilot) media literacy programs in schools. The efficiency of such programs is measured in various ways, including their effects on the participant students' comprehension and media analysis skills (Quin & McMahon 1993; Hobbs & Frost 1999, 2003), on their internet search and website evaluation skills (Kafaï & Bates 1997; Pinkham, Wintle & Silvernail 2008), on their critical thinking regarding media violence (Scharrer 2005, 2006), on their behavior and attitudes towards smoking (Gonzales et al. 2004; Pinkleton et al. 2007; Primack et al. 2014), and on their civic engagement (Hobbs et al. 2013; Kahne et al. 2012; Martens & Hobbs 2015). Although these assessments are often conducted in collaboration with the teachers involved in the media education programs, none of these papers discuss how the results of the assessment were appropriated by teachers.

One obvious explanation lies in the objective of these reports: to demonstrate the effectiveness of Media Education in general, rather than to focus on the optimization of a particular program. However, the scarcity of such discussions on this particular point also hints at the difficulty of translating assessment results into useful guidelines for pedagogical improvement. This is an important issue both for the researchers in charge of the assessment and for the media educators whose program is assessed: for the former, achieving that translation means valorizing their research, and increasing its social relevance; for the latter, it means reaping the benefits of their partnership with the researchers.

To conclude, we distinguish between three scenarios, in ascending order of degree of involvement of the media educators, in which the analysis of the assessment (in terms of observed levels of competences) can inform the correction of the training modules included in the program (courses, seminars, homeworks, projects) in the hope of improving

underdeveloped competences. In the first scenario (individual teacher coaching), the research team in charge of the assessment identifies the possible changes of contents and pedagogies for a given teaching module, and suggests that the teacher implements these changes. In the second scenario (curriculum management), the authority in charge of designing and managing the training program modifies it by removing existing modules or introducing new ones, based on the assessment results, and on the program's competence framework. Ideally, this authority should include the research team.

In the last scenario (teacher empowerment), the teachers themselves design and implement changes to the contents of their classes and to the teaching methods they use. This means bringing the teachers to shift from the traditional intuitive question 'what pedagogical activity am I trying to carry out?' to the question of finality: 'what competences am I trying to get my students to develop?', either by involving them into the process of the longitudinal evaluation (the ideal method, as it requires teachers to adopt a research perspective), or by supporting the teachers' appropriation of the assessment results. This second option is the solution we were able to implement as part of the work reported in this article. It consisted in providing teachers with an executive assessment report, written especially for them in an accessible language, and suggesting amendments for each teaching activity. The report was distributed and presented to the teachers at a dedicated meeting, where it was opened to the collective discussion.

When the objectives of media literacy education are defined in terms of competences, bridging the conceptual and operational gap that separates real life competences from the skills developed through formal school-based learning represents a significant challenge. We believe the last two scenarios (curriculum management and teacher empowerment) to support the autonomy of media educators (in the position of curriculum managers or teachers) in bridging this gap, thereby reducing the distance between media literacy research and media education teaching practices.

References

Ackerman, P. L., Beier M.E., and Bowen K.R., 2002. 'What We Really Know about Our Abilities and Our Knowledge'. In: *Personality and Individual Differences* 33 (4): 587–605. doi:10.1016/S0191-8869(01)00174-X.

Arke, E.T., and Primack, B.A., 2009. 'Quantifying Media Literacy: Development, Reliability, and Validity of a New Measure'. In: *Educational Media International*, 46 (1), 53–65.

Berry, A. M., and Wintle S. E., 2009. 'Using Laptops to Facilitate Middle School Science Learning: The Results of Hard Fun'. Center for Education Policy, Applied Research, and Evaluation, University of Southern Maine.

Brandtweiner, R., Donat E., and Kerschbaum J., 2010. 'How to Become a Sophisticated User: A Two-Dimensional Approach to E-Literacy'. In: *New Media & Society* 12 (5): 813–33. doi:10.1177/1461444809349577.

Conseil Supérieur de l'Education aux Médias, 2013. *Les compétences en éducation aux médias : un enjeu éducatif majeur*. Cadre général. Bruxelles: Fédération Wallonie-Bruxelles.

Danish Technological Institute, and European Association for Viewers' Interests, 2011. 'Testing and Refining Criteria to Assess Media Literacy Levels in Europe.' SMART 2010/0006 O.J. 5/12/2009 2009/S235-3355589. Aarhus / Brussels.

Directive 2010/13/EU of the European Parliament and of the Council of 10 March 2010 on the coordination of certain provisions laid down by law, regulation or administrative action in Member States concerning the provision of audiovisual media services (Audiovisual Media Services Directive)

European Association for Viewers' Interests (ed), 2009. Study on Assessment Criteria for Media Literacy Levels. *A Comprehensive View of the Concept of Media Literacy and an Understanding of How Media Literacy Levels in Europe Should Be Assessed*. Brussels: EAVI.

Eshet-Alkali, Y. and Amichai-Hamburger Y., 2004. 'Experiments in Digital Literacy.' In: *CyberPsychology & Behavior* 7 (4): 421–29. doi:10.1089/cpb.2004.7.421.

Fastrez, P. 2010. 'Quelles Compétences Le Concept de Littératie Médiatique Englobe-T-Il? Une Proposition de Définition Matricielle.' In: *Recherches En Communication* 33: 35–52.

French Ministry of Education. 2013. Arrêté du 24 juillet 2013 modifiant l'arrêté du 14 juin 2006 relatif aux référentiels de connaissances et capacités exigibles pour le brevet informatique et internet (B2i), Available at: https://www.legifrance.gouv.fr/eli/arrete/2013/7/24/MENE1319698A/jo/texte

Gonzales, R., Glik D., Davoudi M., and Ang A., 2004. 'Media Literacy and Public Health' In: *American Behavioral Scientist*, 48 (2), 189-201

Gui, M., and Argentin G., 2011. 'Digital Skills of Internet Natives: Different Forms of Digital Literacy in a Random Sample of Northern Italian High School Students.' In: *New Media & Society* 13 (6): 963–80. doi:10.1177/1461444810389751.

Hargittai, E., 2005. 'Survey Measures of Web-Oriented Digital Literacy.' In: *Social Science Computer Review* 23 (3): 371–79. doi:10.1177/0894439305275911.

Hargittai, E., 2008. 'An Update on Survey Measures of Web-Oriented Digital Literacy.' In: *Social Science Computer Review* 27 (1): 130–37. doi:10.1177/0894439308318213.

Hargittai, E., and Walejko G., 2008. 'The Participation Divide: Content Creation and Sharing in the Digital Age.' In: *Information, Communication and Society* 11 (2): 239–56.

Hobbs, R., and Frost R., 1999. 'The Acquisition of Media Literacy Skills Among Australian Adolescents.' In: *Media Literacy Review*.

Hobbs, R., and Frost R., 2003. 'Measuring the Acquisition of Media-Literacy Skills.' In: *Reading Research Quarterly* 38 (3): 330–55.

Hobbs, R., Donnelly K., Friesem J., and Moen M., 2013. 'Learning to Engage: How Positive Attitudes about the News, Media Literacy, and Video Production Contribute to Adolescent Civic Engagement.' In: *Educational Media International* 50 (4): 231–46. doi:10.1080/09523987.2013.862364.

Jolls, T. and Wilson, C., 2014 'The Core Concepts: Fundamental to Media Literacy Yesterday, Today and Tomorrow' In: *Journal of Media Literacy Education*, 6(2): 68 -78.

Kahne, J., Lee N.-J., and Feezell J.T., 2012. 'Digital Media Literacy Education and Online Civic and Political Participation.' In: *International Journal of Communication* 6: 1-24.

Kafaï, Y., and Bates M.J., 1997. 'Internet Web-Searching Instruction in the Elementary Classroom: Building a Foundation for Information Literacy.' In: *School Library Media Quaterly* 25 (2): 103–11.

Kellner, D., and Share, J. 2007. Critical media literacy, democracy, and the reconstruction of education. In Donaldo Macedo & Shirley R. Steinberg, eds, *Media literacy: A reader*. New York: Peter Lang Publishing: 3-23.

Kruger, J., and Dunning D., 1999. 'Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments.' In: *Journal of Personality and Social Psychology* 77 (6): 1121–34. doi:10.1037/0022-3514.77.6.1121.

Literat, I., 2014. 'Measuring New Media Literacies: Towards the Development of a Comprehensive Assessment Tool.' In: *Journal of Media Literacy Education* 6 (1): 15–27.

Litt, E., 2013. 'Measuring Users' Internet Skills: A Review of Past Assessments and a Look toward the Future.' In: *New Media & Society* 15 (4): 612–30. doi:10.1177/1461444813475424.

Martens, H., & Hobbs, R., 2015. How Media Literacy Supports Civic Engagement in a Digital Age. In: *Atlantic Journal of Communication*, 23 (2): 120-137. http://doi.org/10.1080/15456870.2014.961636

Masterman, L. 1985. *Teaching the Media*. London: Comedia Publishing Group.

Mihailidis, P., & Thevenin, B. (2013). Media Literacy as a Core Competency for Engaged Citizenship in Participatory Democracy. In: *American Behavioral Scientist*, XX(X): 1-12. http://doi.org/10.1177/0002764213489015

Phang, A., and Schaefer D. J., 2009. 'Is Ignorance Bliss? Assessing Singaporean Media Literacy Awareness in the Era of Globalization.' In: *Journalism & Mass Communication Educator* 64 (2): 156–72.

Piette, J., 2000. *Education aux médias et fonction critique*. Paris: L'Harmattan.

Pinkham, C., Wintle S.E., and Silvernail D.L., 2008. '21st Century Teaching and Learning: An Assessment of Student Website Evaluation Skills.' Center for Education Policy, Applied Research, and Evaluation, University of Southern Maine.

Pinkleton, B., Austin E., Cohen M., Miller A., and Fitzgerald E., 2007. 'A Statewide Evaluation of the Effectiveness of Media Literacy Training to Prevent Tobacco Use Among Adolescents.' In: *Health Communication*, 21 (1), 23–34.

Primack, B.A., Douglas E.L., Land S.R., Miller E., and Fine M.J., 2014. 'Comparison of Media Literacy and Usual Education to Prevent Tobacco Use: A Cluster-Randomized Trial.' In: *Journal of School Health* 84 (2): 106–15. doi:10.1111/josh.12130.

Quin, R., and McMahon B., 1993. 'Evaluating Standards in Media Education.' In: *Canadian Journal of Educational Communication* 22 (1): 15–25.

Rey, B., Carette V., Defrance A., and Kahn S., 2012. *Les Compétences À L'école: Apprentissage et Évaluation*. Bruxelles: De Boeck.

Scallon, G., 2004. *L'évaluation Des Apprentissages Dans Une Approche Par Compétences*. Bruxelles: De Boeck Université.

Scharrer, E., 2005. 'Sixth Graders Take on Television: Media Literacy and Critical Attitudes of Television Violence.' In: *Communication Research Reports* 22 (4): 325–33. doi:10.1080/00036810500317714.

Scharrer, E., 2006. 'I Noticed More Violence:' The Effects of a Media Literacy Program on Critical Attitudes Toward Media Violence.' In: *Journal of Mass Media Ethics* 21 (1): 69–86. doi:10.1207/s15327728jmme2101_5.

Tardif, J., 2006. *L'évaluation Des Compétences: Documenter Le Parcours de Développement*. Montréal: Chenelière éducation.

Tilleul, C., Fastrez, P., and De Smedt, T.. 2015. 'Evaluating Media Literacy and Media Education Competences of Future Media Educators.' In: Sirkku Kotilainen and Reijo Kupiainen, eds, *Reflections on Media Education Futures*. Contributions to the Conference Media Education Futures in Tampere, Finland 2014. Götenborg: Nordicom, 75-89.

Tyner, K. 1998. *Literacy in a Digital World*. Mahwah: Erlbaum.

Van Deursen, A., & van Dijk, J., 2010. 'Internet skills and the digital divide'. In: *New Media & Society*, 13(6): 893–911. http://doi.org/10.1177/1461444810386774

Van Deursen, A., Helsper E.J. and Enyon R., 2014. *Measuring Digital Skills*. From Digital Skills to Tangible Outcomes (Project Report). London School of Economics and Political Science / University of Twente / Oxford Internet Institute. Retrieved from http://www.oii.ox.ac.uk/research/projects/?id=112

Van Deursen, A., Courtois C., and van Dijk J., 2014. 'Internet Skills, Sources of Support, and Benefiting From Internet Use.' In: *International Journal of Human-Computer Interaction* 30 (4): 278–90. doi:10.1080/10447318.2013.858458.

Westera, W., 2001. 'Competences in Education: A Confusion of Tongues.' In: *Journal of Curriculum Studies* 33 (1): 75–88. doi:10.1080/00220270120625.