
Medienkompetenz messen: Die Entwicklung des Medienprofis-Tests in der Schweiz

Martin Hermida, Michael Hielscher und Dominik Petko

Zusammenfassung

Trotz intensiver theoretischer Diskussionen über den Begriff Medienkompetenz liegen bis anhin kaum empirisch validierte Messinstrumente zur empirischen Erfassung dieser Kompetenzen bei Kindern und Jugendlichen vor. Die bestehenden Angebote lassen sich in drei Typen gliedern: 1) Kommerzielle Prüfungen im Kontext von Weiterbildungsangeboten, 2) wissenschaftliche Tests im Kontext empirischer Studien sowie 3) spielerische Tests und Quizzes im Kontext von Sensibilisierungskampagnen. Ausgewählte Angebote werden vorgestellt und in Bezug auf ihre Brauchbarkeit für praktische Medienarbeit diskutiert. Als Ergänzung und Erweiterung des bestehenden Angebots wird der schweizerische Medienprofis-Test vorgestellt. Dabei handelt es sich um einen spielerischen Test, der frei zugänglich und empirisch validiert ist. Der Test umfasst fünf Themenbereiche: 1. Digitalisierung und Computerisierung, 2. Informationsfreiheit und Glaubwürdigkeit, 3. Sex, Gewalt und Jugendschutz, 4. Neue Geschäftsmodelle und Kommerz sowie 5. Personalisierung und Datenschutz. Die abgefragten Kompetenzen orientieren sich darüber hinaus am Modullehrplan «Medien & Informatik» des neuen deutschschweizerischen Lehrplans 21. Die von Experten validierten Items wurden anhand einer Stichprobe von Schülerinnen und Schülern der 3. bis 8. Klassen erprobt und mittels Raschmodell skaliert. Lehrpersonen können den Test im Klassenverband durchführen und erhalten dabei neben aggregierten Ergebnissen auch Empfehlungen für passende Unterrichtsmaterialien. Der Medienprofis-Test ist ein erster Schritt, ein standardisiertes und validiertes Messinstrument zur Verfügung zu stellen und die Diskussion um die Ausprägungen der Medienkompetenz mit empirisch Daten zu bereichern.

Measuring media literacy: The Development of the Medienprofis-Test in Switzerland

Abstract

Despite intensive theoretical discussions regarding the concept of media literacy, so far no empirically validated measuring instruments have been developed to determine this competence among children and adolescents. Existing offers can be categorized into three types: 1) commercial tests in the context of further education and training, 2) scientific tests in the context of empirical studies, and 3) playful tests and quizzes in the context of awareness campaigns. We present selected products and discuss them with reference

to their suitability for practical media work. The Swiss Medienprofis-Test is introduced as an addition and extension to existing products. This is a playful test, which is freely available and empirically validated. It encompasses five subject areas: 1. digitalization and computerisation, 2. freedom of information and credibility, 3. sex, violence and youth protection, 4. new business models and commerce and 5. personalisation and privacy protection. The tested literacies are structured parallel to the school subject «media & computer science» from the new Swiss-German curriculum, Lehrplan 21. The test items which were validated by experts, were tested using a random sample of students from 3rd to 8th grade and were scaled according to the Rasch model. Teachers can conduct the test in school classes and receive aggregated feedback as well as suggestions for suitable teaching materials. The Medienprofis-Test is a first step towards providing a standardized and validated measuring instrument and enhancing the discussion regarding the definition of media literacy with empirical data.

Grundprobleme der Messung von Medienkompetenz

Mehr als zehn Jahre ist es her, seit Gapski (2006) den Sammelband «Medienkompetenz messen?» vorgelegt hat. Sein Hauptfazit zur Messung von Medienkompetenz lautete damals: «Die scheinbare Griffigkeit dieser Wortmarke steht im umgekehrten Verhältnis zur Leichtigkeit ihrer Operationalisierung und damit auch Messbarkeit» (ebd., 15). Nichtsdestotrotz blieb Medienkompetenz eine kaum bestrittene und zugleich schillernde Schlüsselkompetenz der Informationsgesellschaft (Rychen und Salganik 2003; Schelhowe et al. 2009), die unterdessen mit zahlreichen komplementären Begriffen angereichert und differenziert wurde, u. a. «digitale Kompetenz» oder «Medienbildung» (Jenkins et al. 2006; Marotzki und Jörissen 2010; Tulodziecki 2011). Dass solche Begriffe trotz teilweise intensiv geführten theoretischen Debatten in den letzten Jahren nie ganz greifbar geworden sind, kann auch daran liegen, dass immer noch verhältnismässig selten der Versuch gemacht wurde, diese Kompetenzen zu messen. Während die Frage nach der Messbarkeit von Kompetenzen in Bildungsstandards und internationale Schulleistungsstudien für Lerninhalte wie Deutsch, Mathematik und Naturwissenschaften letztlich in konkreten Messinstrumenten mündete (Klieme et al. 2003; Oelkers und Reusser 2008), steht die Messung von medienbezogenen Kompetenzen – vor allem im deutschsprachigen Raum – noch relativ am Anfang. Moser (2012) wies darauf hin, dass es eigentlich im Interesse der Medienpädagogik sein müsste, die Entwicklung von Bildungsstandards und empirisch validierter Kompetenzmodelle voranzutreiben, um nicht mittelfristig in Konkurrenz mit besser messbaren Fächern an Bedeutung zu verlieren. Diese Entwicklung in Richtung messbarer Output-Orientierung kann natürlich auch kritisch gesehen werden. So besteht die berechtigte Sorge, dass durch die Formulierung von Bildungsstandards, die sich per Definition nur auf den Kern eines Faches beziehen, eine Dynamik

in Gang gesetzt wird, die weiterführende Inhalte zunehmend verdrängt und einem Phänomen des «Teaching to the Test» Vorschub leistet. Die empirische Überprüfung von Bildungsstandards sollte auch nicht darüber hinwegtäuschen, dass auch sie ursprünglich auf normativen Entscheidungen basieren und keineswegs rein objektive Massstäbe sind. Wenn solche Aspekte jedoch mitbedacht werden, dann liesse sich mit einer empirischen Messung von medienbezogenen Kompetenzen in grossen Studien nicht nur bildungspolitisch einiges gewinnen, sondern dies könnte auch in kleinen Studien praktische Fragen der Vermittlung von Medienkompetenz neu anstossen (Petko 2011). Wenn geeignete Testverfahren vorliegen würden, dann liesse sich die Wirksamkeit unterschiedlicher medienpädagogischer Massnahmen und Projekte besser überprüfen. Bei der Messung von Medienkompetenz ergeben sich jedoch drei fundamentale Probleme, 1) bezüglich des Medienbegriffs, 2) des Kompetenzbegriffs und 3) des Medienkompetenzbegriffs.

1. Die Medienlandschaft unterliegt rasanter Veränderung und dementsprechend unübersichtlich ist der Begriff der *Medien*. War vor 40 Jahren die Einteilung in Individualmedien (Brief, Telefon) und Massenmedien (Presse, Radio, Fernsehen) noch relativ überschaubar, sind mit den Computertechnologien und dem Internet viele neue und medienkonvergente Kanäle dazugekommen, und ständig entstehen neue Medienanwendungen mit neuen Funktionalitäten und immer grösseren Algorithmen. Angesichts der Komplexitätsexplosion des Medienbegriffs, können diesbezügliche Tests bestenfalls einen Ausschnitt abbilden.
2. Ebenso schwierig fassbar ist der Begriff der *Kompetenz*. Der Begriff wurde ursprünglich von Noam Chomsky (1965) in der Linguistik eingeführt, um die Fähigkeit zum kreativen Formulieren zu beschreiben. Seither wurde der Begriff in der Erziehungswissenschaft auf viele andere Bereiche erweitert und differenziert (Klieme und Hartig 2008). Als potenzielle Fähigkeit ist Kompetenz nur mittelbar, über bestimmte Performanzleistungen, zu erschliessen. Dabei stellt sich die Frage, wie konkret oder abstrakt eine bestimmte Performanz für die Messung einer Kompetenz bestimmt werden soll. Definitionen von Kompetenzen die stark vom Konkreten abstrahieren führen zu theoretisch überzeugenden Konstrukten, die sich jedoch typischerweise kaum konkret in einer Testumgebung umsetzen lassen. Definitionen von Kompetenzen, die wenig abstrahieren, sind unmittelbar anwendbar, theoretisch jedoch häufig nicht befriedigend (Weinert 2001).
3. Allein aus der Kombination der Teilprobleme ist es nicht verwunderlich, dass auch der Begriff der *Medienkompetenz* hinsichtlich seiner Dimensionalisierung nie von einem grossen Konsens geprägt war. Ursprünglich von Baacke (1997) in den deutschen Diskurs eingeführt, entstanden in der Folge schnell eine Vielzahl weiterer Ausdifferenzierungen. Bereits um die Jahrtausendwende konnte Gapski (2001) zeigen, dass in den meisten Definitionen der Medienkompetenz 4 bis 5 der folgenden 6 Dimensionen genannt werden: Medienkunde/Wissen, Kritik/

Reflexion, Handhabung/Anwendung, Kreativität/Gestaltung, soziale Verantwortung/Ethik und Emotionen/Affekte. Eine darüber hinausgehende Konkretisierung des Gegenstandes konnte in der Medienpädagogik jedoch bis jetzt nicht abschließend erreicht werden, so dass häufig vor allem das Nebeneinander unterschiedlicher Definitionen mit unterschiedlich dimensionalisierten Differenzierung als Endstand konstatiert wird (Tulodziecki 2011). Dies wird noch komplexer, wenn Begriffsdefinitionen aus dem englischsprachigen Raum einbezogen werden (Mills 2010; Bawden 2001).

Trotz dieser Grundprobleme sind in den letzten Jahren einige Projekte entstanden, die Medienkompetenz in der einen oder anderen Weise versucht haben zu messen. Exemplarische Angebote werden im folgenden Kapitel beschrieben und mit Bezug auf gängige wissenschaftliche Qualitätskriterien diskutiert.

Bestehende Angebote zur Prüfung und Messung von Medienkompetenzen

Die bestehenden Angebote lassen sich grob in drei Kategorien einteilen: 1) Kommerzielle Prüfungen im Kontext von Weiterbildungsangeboten, 2) wissenschaftliche Tests im Kontext empirischer Studien und 3) spielerische Quizzes im Kontext von Sensibilisierungskampagnen. Die Typen unterscheiden sich in ihrem Gegenstandsbereich, im Grad ihrer theoretischen Anschlussfähigkeit, in den Verfahren ihrer empirischen Überprüfung, ihrer Dokumentation sowie ihrer praktischen Verfügbarkeit. Die Beurteilung solch unterschiedlicher Angebote fällt nicht leicht. Exemplarische Angebote werden im folgenden Kapitel kurz skizziert und im Zusammenhang mit den wichtigsten Qualitätskriterien dargestellt.

Kommerzielle Prüfungen im Kontext von Weiterbildungsangeboten

Kommerzielle Prüfungen zur Individualdiagnostik beschränken sich in der Regel auf das Anwendungswissen bestimmter Software. Sie bestehen typischerweise aus nachgebauten Benutzeroberflächen, auf denen vorgegebene Aktionen ausgeführt werden müssen oder unter verschiedenen Optionen die Beschreibung der richtigen Aktion ausgewählt werden kann. Die bekannteste dieser Prüfungen ist vermutlich die «European Computer Driving Licence» (www.ecdl.org), die sich an Sekundarschüler und Erwachsene richtet und die «praktische Fertigkeiten in den gebräuchlichsten Computeranwendungen» zertifiziert (ECDL 2016). Die Testinhalte folgen keiner dokumentierten theoretischen Konzeption, sondern beziehen sich direkt auf die gebräuchlichsten Programmoberflächen – wobei Programme von Microsoft typischerweise im Vordergrund stehen – sowie auf das Thema (Daten-)Sicherheit. Über eine empirische Testskalierung und Eichung ist nichts bekannt. Als Vorbereitung für den Test stehen kostenpflichtige elektronische Selbstlernkurse zur Verfügung und die

Tests können gegen Bezahlung in Testzentren durchgeführt werden. Ähnliche und ebenfalls kommerzialisierte Tests im englischsprachigen Raum sind der Certiport IC3 (www.certiport.com/IC3) und der iSkills Test (www.ets.org/iskills). Letzterer wurde teilweise auch wissenschaftlich überprüft, aber auch dies ist nur unzureichend öffentlich dokumentiert (Katz 2007; Somerville et al. 2008). Ein Beispiel für einen nicht kommerzialisierten Test ähnlicher Machart bietet hingegen das Northstar Digital Literacy Project (www.digitalliteracyassessment.org). Auch hier wird Anwendungswissen in verschiedenen Programmtypen abgefragt, wobei die Probanden in der Regel nicht nur die richtigen Antworten auswählen müssen, sondern die richtigen Handlungen in originalgetreuen, simulierten Interfaces ausführen müssen. Wie bei den anderen Prüfungen, ist auch hier nichts über die theoretischen Grundlagen und die empirischen Kennwerte zur Testgüte bekannt.

Wissenschaftliche Tests im Kontext empirischer Studien

Einen anderen Ansatz verfolgen wissenschaftlich überprüfte Tests zur Populationsdiagnostik. Als sozialwissenschaftliche Routineverfahren zur Erfassung empirisch messbarer Personenmerkmale unterliegen sie einer ganzen Reihe standardisierter Qualitätskriterien (Bühner 2011; Rost 2004). Ein guter wissenschaftlicher Test muss in Bezug auf das zu messende Merkmal einerseits theoretisch fundiert und andererseits empirisch überprüft sein. Die wichtigen empirisch zu bestimmenden Gütekriterien für einen Test sind seine Objektivität, Validität und Reliabilität. Heute existieren mit der klassischen und der probabilistischen Testtheorie unterschiedliche Ansätze zur Testkonstruktion und Testinterpretation, die sich auch in unterschiedlichen statistischen Verfahren und kriterialen Koeffizienten niederschlagen. Wenn ein Test zur Individualdiagnostik eingesetzt werden soll, ist ausserdem eine Eichung des Testinstruments an einer Referenzstichprobe nötig. Während Tests im vordigitalen Zeitalter typischerweise in Form von ausgedruckten Materialien mit offenen oder standardisierten Antwortformaten vorlagen, sind heute auch digitale Varianten möglich, die auch interaktive und multimediale Aufgabenformate enthalten können, z. B. über die Verwendung simulierter Interfaces, und überdies adaptive Vorgehensweisen ermöglichen. Es gibt hier verschiedene Instrumente, die in Deutscher Sprache verfügbar sind.

Ein aktuelles Beispiel ist der Test der International Computer- and Information Literacy Study (ICILS). Bei ICILS handelt es sich um eine internationale large-scale Schulleistungsstudie nach dem Vorbild von PISA, TIMSS oder PIRLS, in der die computer- und informationsbezogenen Kompetenzen von Achtklässlern im Zentrum stehen (Fraillon et al. 2013; Eickelmann et al. 2014). Die Studie misst Computer- und Informationskompetenz in zwei Teilbereichen: 1) Informationen sammeln und organisieren und 2) Informationen erzeugen und austauschen. Der erste Teilbereich wird in

drei Unterbereiche differenziert: a) Wissen zur Nutzung von Computern, b) auf Informationen zugreifen und Informationen bewerten und c) Informationen verarbeiten und organisieren. Der zweite Teilbereich gliedert sich in die vier Bereiche Informationen a) umwandeln, b) erzeugen, c) kommunizieren und austauschen und d) sicher nutzen. Bei ICILS existieren unterschiedliche Aufgabenformate. Dazu gehören offene Textantworten und standardisierte Multiple-Choice und Drag&Drop Formate, lineare und nichtlineare Aufgaben in simulierten Interfaces sowie die Bearbeitung grösserer Aufgaben am Computer, wobei dann das Ergebnis standardisiert von geschulten Beurteilern geprüft wurde. Der Test umfasst insgesamt 62 Items, die auf Basis der probabilistischen Testtheorie überprüft wurden und auf deren Basis für jeden Schüler und jede Schülerin ein plausibler Fähigkeitswert statistisch geschätzt wird. Zur leichteren Interpretation wird aufgrund dieser Punktzahl zudem eine Zuordnung jedes Schülers und jeder Schülerin zu einer von fünf Kompetenzstufen vorgenommen. Die Verfahren der Testskalierung sind transparent beschrieben und folgen üblichen Standards (Fraillon et al. 2015). Allerdings wurde – wie in large scale Studien leider üblich, um Probleme des «Teaching-to-the-Test» zu vermeiden und Testitems in späteren Studien wiederverwenden zu können – die genaue Formulierung der Items nur exemplarisch und nicht vollständig publiziert. Die Kombination dieser Faktoren führt letztlich dazu, dass der Test ausserhalb der ICILS Studie nicht nutzbar ist.

Dass dies für wissenschaftliche Tests ein gängiges Problem ist, zeigt auch der für Deutschland im Rahmen des Nationalen Bildungspanel (NEPS) entwickelte Technologische und informationsbezogener Literacy-Test (TILT), der für unterschiedliche Alterskohorten skaliert wurde, von Schülerinnen und Schülern der 5. bis 9. Stufe (11-15 Jahre) bis zu Erwachsenen (Senkbeil et al. 2013a; Senkbeil et al. 2013b). TILT testet insbesondere den Bereich der Handhabung von gängiger Anwendungssoftware in den Bereichen Betriebssystem/Textverarbeitung, Tabellenkalkulation, Präsentation, E-Mail/Kommunikation sowie Suchmaschinen. In Kombination mit den vier Aktivitätstypen «Access», «Create», «Manage», «Evaluate» ergibt sich eine komplexe Typologie von Testbereichen. Die 30 standardisierten Items beruhen primär auf Screenshots, auf die sich Fragen und Multiple-Choice-Antwortoptionen beziehen. Auch hier sind zwar die statistischen Verfahren der probabilistischen Testkonstruktion transparent dokumentiert, nicht jedoch die vollständigen Testitems (Senkbeil 2013a, 680).

Ein drittes Beispiel bietet der «Medienkritikfähigkeit»-Test (MKF) von Sowka et al. (2015). Der Test prüft anhand von 14 Fragen, bei denen mediale Stimuli wie Videos beurteilt werden müssen, die Ausprägung der Medienkritikfähigkeit bei Jugendlichen zwischen 15 und 17 Jahren. Die Medienkritikfähigkeit wird dabei als Kompetenz verstanden, «rezipierte Medienangebote kritisch hinsichtlich ihrer Aussagen, Qualität, anzunehmenden Produktionsumständen und/oder gesellschaftlicher oder normativer Implikationen zu bewerten» (ebd., 65). Der Test bestimmt die Ausprägung der Kritikfähigkeit für vier Arten von Medieninhalten: Informationsangebote,

Unterhaltungsangebote, Werbung und Nutzerkommunikation. Angesichts der geringen Itemzahl in den verschiedenen Teilaspekten ist die Aussagekraft, insbesondere zur Individualdiagnostik, jedoch vermutlich eingeschränkt. Obwohl auch hier die Testskalierung wissenschaftlich dokumentiert ist, ist auch dieser Test bislang nicht vollständig öffentlich dokumentiert.

Daneben existieren verschiedene ältere Tests für Individualdiagnostik, die teilweise auch gut dokumentiert sind, die jedoch nicht mehr weiterbetrieben werden oder klar veraltet sind, da sie noch aus einer Zeit vor Smartphones und Social Media stammen, z. B. INCOBI und INCOBI-R (Richter et al. 2001; Richter et al. 2010), Start-C (Wagener 2008) sowie C-WIS-4 (Wagener 2003); ausserdem TestYourICTKnowledge (Moser und Keller 2004). Als Zwischenfazit bleibt festzuhalten, dass im wissenschaftlichen Bereich durchaus Tests vorliegen, die jedoch nur für wissenschaftliche Zwecke einsehbar und ausserhalb der Wissenschaft kaum praktikabel einsetzbar sind. Für einen praktischen Einsatz in Schulen und anderen Bildungskontexten stehen sie bislang nicht zur Verfügung. Dies ist einerseits auf die Geheimhaltung der Items im Rahmen der jeweiligen Studien zurückzuführen, andererseits auf die aufwändigen und statistisch anspruchsvollen Auswertungsprozeduren, die von Laien kaum zu bewältigen sind. Gleichzeitig scheinen die vorliegenden individualdiagnostischen Tests, die im praktischen Kontext handhabbar wären, tendenziell veraltet.

Spielerische Tests und Quizzes im Kontext von Sensibilisierungskampagnen

Bei den Instrumenten, die in den letzten Jahren explizit für den Einsatz in Schulen entwickelt wurden, handelt es sich mehrheitlich eher um Quizzes als um Tests. Sie wurden typischerweise im Rahmen von Präventionskampagnen entwickelt, die Kinder und Jugendliche für Teilaspekte des Medienangebots und seiner Nutzung sensibilisieren möchten. Bei Quizzes handelt es sich um eine der einfachsten Formen von Serious Games (Granic et al. 2014). Quizzes sind strukturierte Wissensspiele, bei denen typischerweise eine einfache Frage-und-Antwortmechanik im Mittelpunkt steht, die zusätzlich um kooperative oder kompetitive Spielregeln erweitert wird. Das Grundmuster besteht darin, dass sich mit richtigen Antworten auf die Quizfragen punkten lässt. Mit Tests haben sie die Gemeinsamkeit, dass sie auf einer Reihe von Wissens- oder Könnensfragen basieren, die sich auf mehr oder weniger oberflächliche oder tiefe Verständnisse eines Sachverhalts richten (Cox und Clark 1998). Im Unterschied zu Tests, haben Quizzes normalerweise nicht den Anspruch, Gütekriterien wie Objektivität, Reliabilität oder Validität zu entsprechen, jedoch unterliegen sie dem Kriterium der überprüfbaren Korrektheit der Antworten. Wenn Quizantworten nicht eindeutig richtig oder falsch sind, kann dies zu Diskussionen unter den Spielern führen und damit den Spielspass entscheidend mindern. Ob es sich bei Quizzes um Serious Games handelt, bei denen nicht nur bestehendes Wissen angewendet,

sondern auch neues Wissen erworben wird, ist von Quiz zu Quiz unterschiedlich. Typischerweise sind viele Quizzes jedoch so aufgebaut, dass ein Spassfaktor gerade darin liegt, marginales Faktenwissen abzufragen (z. B. beim bekannten *Trivial Pursuit™*). Wenn Quizzes für Lernzwecke verwendet werden sollen, dann müssten sie sich jedoch generell an allgemeinen Designprinzipien von Serious Games orientieren. Ob es möglich und sinnvoll ist, testtheoretische Kriterien auch für die Entwicklung eines Quiz anzuwenden, ist noch nicht geklärt, da Quizzes vielfältige Elemente enthalten, die in einem strengen Test verfälschend wirken können. Dazu gehört insbesondere formatives Feedback auf einzelne Antworten während der Durchführung des Quizzes aber auch grafische Elemente. Dennoch kann es Sinn machen, dass auch Quizfragen valide und weder zu schwer noch zu leicht sind. Weitere testtheoretische Kriterien wie Reliabilität und Objektivität können ergänzende Informationen liefern, um gute von weniger guten Quizfragen zu unterscheiden. Wie die nachfolgenden Beispiele zeigen, orientieren sich die bestehenden Quizzes jedoch noch nicht an solchen Kriterien.

Der *Teachtoday Medienkompetenztest*, einer Initiative der Deutschen Telekom, ist nach eigener Deklaration ein spielerischer Online-Test, der sich an 9- bis 12-Jährige richtet. Nach formuliertem Anspruch behandelt er die Aspekte der *Kommunikationsfähigkeit*, *Kooperationsfähigkeit*, *Beurteilungsvermögen*, *Schöpferische Fähigkeit*, *Entscheidungskompetenz* und *Eigenverantwortung* (http://www.teachtoday.de/218_Medienkompetenztest.htm). Jede dieser Dimensionen wird mit 4 Fragen im Test abgedeckt. Die Dimensionen wurden nicht in wissenschaftlicher Hinsicht auf testtheoretische Kriterien überprüft. Auch ohne solche Überprüfung kann jedoch diskutiert werden, wie valide diese Fragen sind. Z. B. wird bei der Frage, welchen Kanal man zur Hilfe bei der Reparatur eines Fahrrades aufsuchen würden, für die Antwortoption *YouTube* die höchste Punktzahl vergeben, während die anderen Antwortoptionen *Google* und *Chat mit Freunden* für eine tiefere Ausprägung der Medienkompetenz stehen. Solche Fragen scheinen eher Rückschlüsse auf eine *Medienpräferenz* und nicht auf eine *Medienkompetenz* zuzulassen. Bei der Auswertung des Tests werden die Ergebnisse jeweils entlang der sechs eingangs erwähnten Dimensionen als Prozentangaben wiedergegeben und mit einem kurzen Textbaustein als Feedback versehen, z. B.: «Du hast besonders gute Ergebnisse im Bereich Wissen! Du hast bereits erste Erfahrungen im Umgang mit Medien gemacht, aber so langsam solltest du dich auf die Socken machen. Noch weißt du zu wenig über die Möglichkeiten einer sicheren Mediennutzung; auch die technologischen Schritte sind dir noch nicht vertraut.» (automatisch generiertes Feedback nach Ausfüllen des Tests). Allerdings fehlen konkrete Hinweise zu Materialien, mit denen die Wissenslücken geschlossen werden könnten.

Etwas anders gestaltet ist das Schlundz-Quiz (<http://www.lpr-hessen.de/schlundz-quiz/>). Das Quiz richtet sich an Grundschul Kinder und soll ihnen dabei helfen, «sich [...] Hintergrundwissen über verschiedenen Medien anzueignen». Das Quiz

besteht aus den vier Themenbereichen Fernsehen, Werbung, Internet und Handy. Das Wissen zu jedem Thema wird mit 10 Fragen abgefragt. Die Probanden erhalten unmittelbar Feedback darüber, ob sie die richtige Antwort gegeben haben, indem nach jeder Frage ein kurzer Info-Text folgt. Auf Wunsch steht auch noch ein weiterführender Info-Text zur Verfügung. Zur Veranschaulichung ein Beispiel einer Frage aus dem Themenbereich Internet: «Wie nennt man Briefe oder Mitteilungen, die du über das Internet verschicken kannst?» Antwortmöglichkeiten: a) A-Mail, b) E-Mail, c) X-Mail. Darauf folgen die Info-Texte: «E-Mail steht für «electronic mail» und bedeutet wörtlich übersetzt «elektronische Post». Wenn du eine E-Mail verschickst, ist sie innerhalb von Sekunden beim Empfänger – egal ob bei den Nachbarn oder einer Familie in Australien. Neben dem Text kannst du auch Bilder und Dateien in einer E-Mail verschicken.» Der spielerische Charakter des Quiz wird unterstrichen durch Hintergrundmusik, Soundeffekte bei richtigen und falschen Antworten und eine animierte Benutzeroberfläche. Die Fragen, Antworten und Erklärungstexte bestehen jedoch ausschließlich aus Text.

Noch deutlicher als Game kommt der Medienquiz der Thüringer Landesmedienanstalt daher (www.medienquiz.de). Hier ist die Quizoberfläche klar an gängige Fernsehformate angelehnt. Punkte erhalten Spielende nicht nur für korrekte Antworten, sondern auch für die Antwortzeit. Das Quiz mischt in unterhaltsamer Weise spielerische Frageformate, wie «Infofrage», «Speedfrage» oder «Recherchefrage», bei der im Internet nachgeschlagen werden soll. Abgefragt wird allerdings – ganz im Sinne typischer Unterhaltungsquizzes – primär marginales und deklaratives Wissen, z. B. ob «Bug» ein Ausdruck für einen Programmierfehler ist. Daneben existieren weitere Angebote mit Quizzes, die in interaktive Geschichten verpackt sind (z. B. www.internet-hero.at oder www.internauten.de).

In englischer Sprache finden sich eine Vielzahl ähnlicher Quizzes, u. a. die australische *eSmart Digital Licence* der *Alannah and Madeleine Foundation*. Dieses Quiz deckt ein breiteres Themenspektrum ab: Digital Devices, Protecting Privacy, Searching & Researching, Creating & Sharing, Social Networking & Gaming, Communicating Safely Online, Friends & Strangers, and Managing Money & Online Credits (The Alannah and Madeline Foundation 2016). Nach jeder Frage erhalten die Probanden ein Feedback dazu, ob und warum ihre Antwort richtig oder falsch war, wobei nicht nur sachliche Korrektheit beurteilt wird, sondern teilweise auch normative und moralische Kriterien angelegt werden.

Im Hinblick auf Quizzes lässt sich zusammenfassend der Eindruck festhalten, dass die verschiedenen Angebote sich primär als sensibilisierende Spiele verstehen und nicht den hohen Ansprüchen von Tests genügen können und wollen. Dass dennoch immer wieder Fragen in diesen Quizzes auftauchen, die nicht klar beantwortet werden können, kann nicht wünschenswert sein. Auch in Bezug auf spielbezogene

Kriterien könnten solche Quizzes noch verbessert werden, zumal sie nur beschränkt Gebrauch von multimedialen Möglichkeiten machen und in erster Linie Text-Quizzes mit multimedialer Rahmung darstellen.

Der Ansatz des Schweizerischen Medienprofis-Tests

Nach Sichtung der skizzierten Projekte ist es augenfällig, dass es bislang immer noch kein Angebot gibt, das die Potenziale von Tests und Quizzes überzeugend verbindet. Gefragt wären insbesondere Angebote, die theoretisch fundiert, empirisch geprüft, transparent dokumentiert, öffentlich und möglichst kostenfrei verfügbar sind und die ausserdem die spielerischen und interaktiven Potenziale von Quizzes – die Erhöhung der Motivation und des Engagements der Lernenden (Lee und Hammer 2011) – sinnvoll nutzen. Ein Medienkompetenz-Test, der sowohl die wissenschaftlichen Ansprüche an Tests erfüllt (objektiv, valide und reliabel) und die motivierenden Potentiale von Quizzes aufweist, ermöglicht eine präzise Messung der Medienkompetenz bei gleichzeitig engagierter Teilnahme der Schülerinnen und Schüler. Um einen ersten Schritt in diese Richtung zu machen, wurde in der Schweiz versucht, unter der Schirmherrschaft der Stiftung *Pro Juventute* und mit finanzieller Unterstützung von *Google Schweiz*, ein Angebot zu entwickeln, das testtheoretischen Kriterien genügt, spielerische Elemente integriert und das frei verfügbar ist. Seit November 2016 ist dieses Angebot unter www.medienprofis-test.ch online.

Inhaltliche Struktur

Der Test gliedert sich aktuell in fünf Themenbereiche (1. Digitalisierung und Computerisierung, 2. Informationsfreiheit und Glaubwürdigkeit, 3. Sex, Gewalt und Jugendschutz, 4. Neue Geschäftsmodelle und Kommerz, 5. Personalisierung und Datenschutz, vgl. Abb. 1).

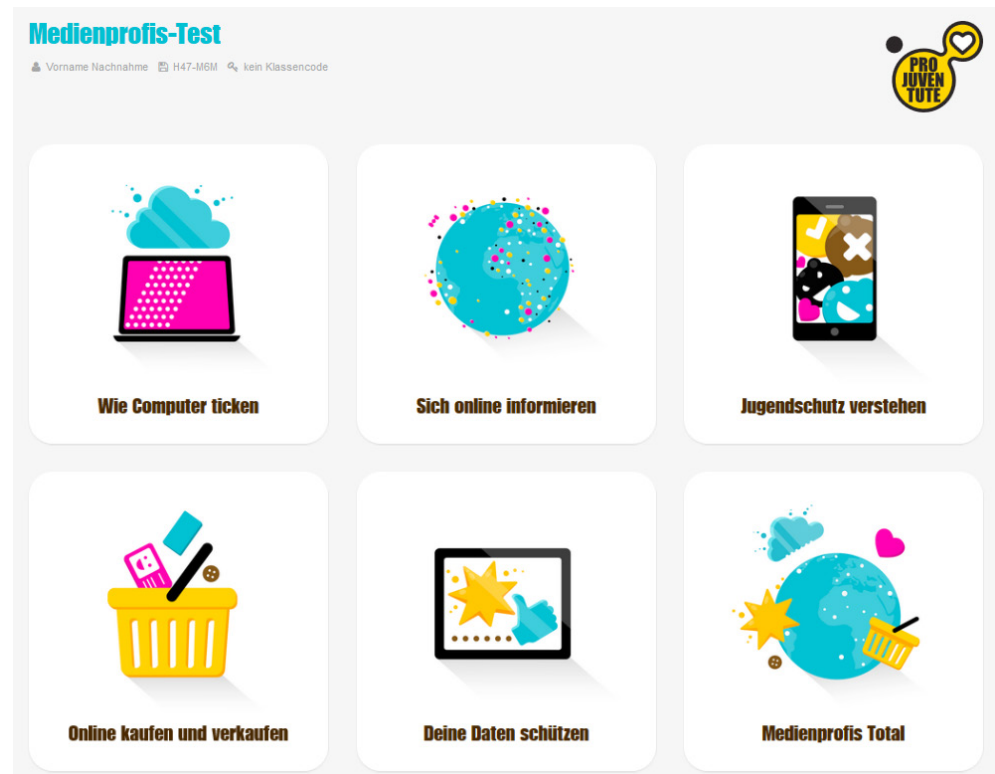


Abb. 1: Übersicht über die verfügbaren Testmodule von www.medienprofis-test.ch.

Jeder dieser Themenbereiche umfasst sowohl positive wie kritische Aspekte. Diese Themen bilden freilich nicht das ganze mögliche Kompetenzspektrum im Bereich der Medienkompetenzen ab, was sich allein schon daran zeigt, dass sich der Test weitgehend auf digitale Medien konzentriert. Diese Themen spiegeln teilweise die Weiterbildungsangebote von *Pro Juventute*, damit der Test auch als Vor- und Nachbereitung dieser Kurse dient, die von Schulen über die Webseite der *Pro Juventute* in der Schweiz buchbar sind (www.medienprofis.ch). Das Themenspektrum wurde ausserdem in einer Expertengruppe diskutiert, in der neben bekannten Hochschulvertreterinnen und -vertretern der Schweizer Medienbildung später auch Präventionsstellen der Polizei und weitere Beratungsstellen vertreten waren. Übergeordnetes Ziel war es, alltägliche Nutzungssituationen von Kindern und Jugendlichen zu berücksichtigen. Jedes Thema ist ausserdem in vier Kompetenzbereiche untergliedert (A. Grundlagenwissen, B. rezeptive Mediennutzung, C. produktive Mediennutzung, D. kommunikative Mediennutzung). Diese theoretische Dimensionierung der Medienkompetenz in vier Kompetenzbereiche wurde aus dem Modullehrplan «Medien und Informatik» des neuen Schweizerischen Lehrplan 21 übernommen, um die Anschlussfähigkeit an die schulischen Lernziele zu gewährleisten. Laut dessen Zielsetzungen sollen die Schülerinnen und Schüler die Bedeutung der Medien für

Individuen und Gesellschaft verstehen und diese kritisch, kompetent und verantwortungsvoll nutzen. Dabei sollen sie die mit der Nutzung verbundenen Chancen und Risiken einschätzen können und ein sicheres und sozialverantwortliches Verhalten in und mit Medien zeigen (D-EDK 2017).

	A. Grundlagenwissen	B. Rezeptive Mediennutzung	C. Produktive Mediennutzung	D. Kommunikative Mediennutzung
1. Digitalisierung und Computerisierung: Möglichkeiten/ Grenzen	<p>Computertechnik:</p> <ul style="list-style-type: none"> • Welche Medien brauchen Strom <p>Internettechnik:</p> <ul style="list-style-type: none"> • Mit welchen Geräten kann man ins Internet • Wie sind Computer vernetzt • Protokolle/Sprachen/Codes <p>Digitalisierung:</p> <ul style="list-style-type: none"> • Neue und alte Medien • Was leisten Computer <p>Mediengeschichte:</p> <ul style="list-style-type: none"> • Welche Medien gab es schon vor 50 Jahren <p>Informationsgesellschaft:</p> <ul style="list-style-type: none"> • Abhängigkeit von Computertechnik • Akzeleration • Neue Berufe 	<p>Dateiformate:</p> <ul style="list-style-type: none"> • Was kann als Datei gespeichert werden • Dateiformate <p>Medienkonvergenz:</p> <ul style="list-style-type: none"> • Was kann ein Smartphone • Für was gibt es Apps • Smart TVs <p>Medialitätsbewusstsein:</p> <ul style="list-style-type: none"> • Glaubwürdigkeit im Internet 	<p>Medienproduktion:</p> <ul style="list-style-type: none"> • Welche Apparate ersetzt der Computer • Peripheriegeräte <p>Textverarbeitung:</p> <ul style="list-style-type: none"> • Eigenschaften von digitalen Texten • Texteingabe • Text Formatieren <p>Videobearbeitung:</p> <ul style="list-style-type: none"> • Komprimierung • Spezialeffekte 	<p>Kommunikationskanäle:</p> <ul style="list-style-type: none"> • Eigenschaften • One-to-one vs. one-to-many <p>Adressaten:</p> <ul style="list-style-type: none"> • Wann SMS/Chat <p>Netiquette:</p> <ul style="list-style-type: none"> • Grundsätze • Passende Anreden <p>Netspeak:</p> <ul style="list-style-type: none"> • Gamer-Slang
2. Informationsfreiheit / Glaubwürdigkeit	<p>Grundelemente:</p> <ul style="list-style-type: none"> • Was sind Internetadressen • Links <p>Regulierung:</p> <ul style="list-style-type: none"> • Wer darf Webseiten erstellen • Wie ist das Internet organisiert • Was ist im Internet strafbar 	<p>Suchmaschinen:</p> <ul style="list-style-type: none"> • Anbieter Suchmaschinen • Funktionsweise <p>Glaubwürdigkeit:</p> <ul style="list-style-type: none"> • Glaubwürdigkeit von Quellen • Informationen verifizieren • Kriterien zur Beurteilung 	<p>Informationsverarbeitung:</p> <ul style="list-style-type: none"> • Was kann man mit Bildern machen • Geräte synchronisieren <p>Webseitenerstellung:</p> <ul style="list-style-type: none"> • Elemente von Webseiten • Technische Grundlagen 	<p>Foren:</p> <ul style="list-style-type: none"> • Aussagen über Foren <p>Soziale Medien:</p> <ul style="list-style-type: none"> • Was leisten soziale Medien • Rahmenbedingungen • Was bedeuten viele Likes

<p>2. Informationsfreiheit/ Glaubwürdigkeit</p>	<p>Zugang:</p> <ul style="list-style-type: none"> • Wie passiert die Datenübertragung • Diffusion in versch. Ländern <p>Meinungsfreiheit:</p> <ul style="list-style-type: none"> • Grundsätze • Nationale Kontexte 	<p>Bild und Video:</p> <ul style="list-style-type: none"> • Gefälschte Bilder erkennen <p>Personalisierte Informationen:</p> <ul style="list-style-type: none"> • Logins • Personalisierte Suchresultate • Personalisierung surfen 		
<p>3. Sex und Gewalt/Jugendschutz</p>	<p>Liberalisierung:</p> <ul style="list-style-type: none"> • Inhalte für Erwachsene vs. für Kinder • Effekte gewalthaltiger/schlimmer Dinge • Zensur (Schweiz vs. andere Länder) <p>Zugang:</p> <ul style="list-style-type: none"> • Warum sind verbotene Inhalte im Netz • Wo kann man Jugendschutzfilter aktivieren <p>Recht:</p> <ul style="list-style-type: none"> • Für den Konsum welcher Inhalte kann man bestraft werden • Für die Herstellung/Verbreitung welcher Inhalte kann man bestraft werden 	<p>Gewalt:</p> <ul style="list-style-type: none"> • Toleranzgrenze Gewalt • Warum reale Gewalt nicht ansehen? <p>Altersempfehlungen:</p> <ul style="list-style-type: none"> • Was sagt FSK aus • Was bedeuten die Pegi-Symbole • Warum Altersempfehlungen <p>Pornografie:</p> <ul style="list-style-type: none"> • Was ist Pornografie? • Versteckte Werbung für Pornoseiten • Effekte von Pornografie 	<p>Produktion:</p> <ul style="list-style-type: none"> • Wann Fotos machen und zeigen • Was zu filmen ist strafbar • Reichweite der Gesetze <p>Veröffentlichungen:</p> <ul style="list-style-type: none"> • Was ist in Eigenproduktionen OK • Folgen von Sexting <p>Verbreiten:</p> <ul style="list-style-type: none"> • Illegale Inhalte liken und teilen 	<p>Cybermobbing:</p> <ul style="list-style-type: none"> • Gesetzliche Grundlagen • Was ist Mobbing • Eigenschaften von Cybermobbing <p>Sexting/Grooming:</p> <ul style="list-style-type: none"> • Auf welche Fragen nicht antworten • Gesetzliche Grundlagen pornografische Inhalte <p>Reporting:</p> <ul style="list-style-type: none"> • Was bewirkt blockieren von Nutzern • Wann blockieren, wann melden • Rechtsräume und Site-Policies <p>Unterstützung:</p> <ul style="list-style-type: none"> • Wo Unterstützung holen? • Wie gegen Cybermobbing wehren • Wo Antworten zum Thema Sex finden
<p>4. Neue Geschäftsmodelle/ Kommerz</p>	<p>Online einkaufen:</p> <ul style="list-style-type: none"> • Was erhält man alles online • Vor- und Nachteile Online-Einkäufe • Probleme beim Online-Kauf 	<p>Bezahlen:</p> <ul style="list-style-type: none"> • Wie kann man online bezahlen • Wie funktionieren Guthabekarten • Wie funktionieren Zahlungsmittel 	<p>Copyright:</p> <ul style="list-style-type: none"> • Copyright Musik • Copyright Konzertaufnahmen • Copyright, Creative Commons und Public Domain 	<p>Monetarisierung Social Media:</p> <ul style="list-style-type: none"> • Werbung auf Youtube • Monetarisierungsmöglichkeiten von Youtube-Stars • Monetär bedingte Verzerrung von Meinung

<p style="writing-mode: vertical-rl; transform: rotate(180deg);">4. Neue Geschäftsmodelle/ Kommerz</p>	<p>Verkaufen:</p> <ul style="list-style-type: none"> • Verschiedene Möglichkeiten • Verpflichtungen als Verkäufer • Der eigene Online-Handel <p>Werbung:</p> <ul style="list-style-type: none"> • Stimmt Werbung immer? • Verschiedene Arten von Werbung • Online-Werbemodelle <p>Kommerzielle Absichten:</p> <ul style="list-style-type: none"> • Kommerzielle Absichten erkennen • Wie Webseiten Geldverdienen • Native Advertising/ Product Placement 	<p>Sicheres Einkaufen:</p> <ul style="list-style-type: none"> • Regeln für Kinder • Käufe rückgängig machen • Hinweise für sichere Shops <p>In-App-Käufe:</p> <ul style="list-style-type: none"> • Kauf-Buttons erkennen • Grundlagen • Monetarisierung von Apps <p>Nutzungsrechte:</p> <ul style="list-style-type: none"> • Kopien zum Privatgebrauch • Kauf vs. Streaming • Filesharing 	<p>Autorenrechte:</p> <ul style="list-style-type: none"> • Was dürfen andere mit meinen Bildern machen • Ein Video verkaufen • Einnahmen- und Ausgabenquellen bei selbstproduzierten Games <p>Businessmodelle:</p> <ul style="list-style-type: none"> • Ablauf Online-Auktion • Ablauf Crowdfunding • Businessmodelle 	<p>Online-Betrug:</p> <ul style="list-style-type: none"> • Warum Betrüger schwer zu fassen sind • Betrügerische Mails • Wie auf Abfallen reagieren <p>Möglichkeiten und Grenzen der Werbung:</p> <ul style="list-style-type: none"> • Unerlaubte Werbung • Werblocker • Virale Werbung
<p style="writing-mode: vertical-rl; transform: rotate(180deg);">5. Personalisierung/Datenschutz</p>	<p>Datensammlung:</p> <ul style="list-style-type: none"> • Was sind persönliche Daten • Welche Daten speichern Service-Provider • Was ist Vorratsdatenspeicherung <p>Daten schützen:</p> <ul style="list-style-type: none"> • Wie Daten schützen? • Gute Passwörter • Wann Passwörter ändern <p>Verwendung von Daten:</p> <ul style="list-style-type: none"> • Daten als Hinweise auf Merkmale der Person • Auswertungen von Handydaten • Datenschutzgesetz 	<p>Personalisierung:</p> <ul style="list-style-type: none"> • Personalisierung im Alltag • Personalisierung bei Online-Werbung • Filter-Bubble <p>Login/Tracking/ Mining:</p> <ul style="list-style-type: none"> • Funktion Logins • Ablauf Tracking • Welche Daten werden ausgewertet <p>Grenzen der Anonymität:</p> <ul style="list-style-type: none"> • Informationen in Nicknamen • Inkognito-Modus • Datenspuren 	<p>Selbstdarstellung:</p> <ul style="list-style-type: none"> • Welche Informationen kommen ins Profil? • Fotos online posten • Was posten und was nicht? <p>Interagieren:</p> <ul style="list-style-type: none"> • Welche Freundschaftsanfragen beantworten? • Funktionen in sozialen Netzwerken • Fotos von und mit anderen posten und markieren <p>Datensicherung:</p> <ul style="list-style-type: none"> • Daten passend benennen • Häufigkeit und Ort der Sicherung • Eigenschaften Cloud 	<p>Datenschutzrichtlinien:</p> <ul style="list-style-type: none"> • Warum Webseiten persönliche Daten wollen • Folgen des exzessiven Teilens von Daten • Komplizierte Formulierungen in AGB. <p>Phishing:</p> <ul style="list-style-type: none"> • Wann ist es Ok, Daten weiterzugeben • Phishing Mails erkennen • Social Engineering <p>Verschlüsselung:</p> <ul style="list-style-type: none"> • Was kann verschlüsselt werden • Schlüssel vs. Nachricht • Asynchrone Verschlüsselung

Tab. 1: Themenmatrix von www.medienprofis-test.ch (Stand 14.03.2017).

Entwicklung der Testfragen

Die Fragen des Medienprofis-Tests wurden in einem ersten Schritt von einem Team des *Instituts für Medien und Schule* der *Pädagogischen Hochschule Schwyz* entlang des Themenrasters in Tabelle 1 entworfen. Zu jedem der aufgeführten Unterthemen wurde ein Fragenblock aus drei bis zwölf Items entwickelt, so dass der gesamte Test bei einer anvisierten Länge von 10 Fragen typischerweise mehr als fünfzig Items umfasst. Insgesamt wurden somit rund 250 Fragen entwickelt. Dabei gab es fünf unterschiedliche Fragetypen: Entscheidungsfragen mit dichotomen Optionen (z. B. stimmt/stimmt nicht), Einschätzfragen mit mehrfachen Optionen (z. B. stimmt/stimmt teilweise/stimmt nicht), Paarzuordnungen («Ziehe die passenden Begriffe/Abbildungen zusammen»), Bildung von Rangreihen («Ordne die Elemente nach ihrer Dateigröße») sowie Einordnungen auf einem Zahlen- oder Zeitstrahl («Wann wurden folgende Medien erfunden?»). Grafisch wurden die Fragen in einer Lernkartenmetapher gestaltet, wobei sich die Karten per Drag und Drop auf dem Bildschirm an den korrekten Ort verschieben lassen, was insbesondere auf Touchscreens intuitiv ist (Abb. 2).



Abb. 2: Auswertung einer dichotomen Entscheidungsfrage: Woran kann es liegen, dass Suchanfragen in einer Suchmaschine bei verschiedenen Personen zu unterschiedlichen Ergebnissen führen?

In einem zweiten Schritt wurden die Fragen wiederum in der Expertengruppe begutachtet. Grundlegender Anspruch der Testfragen war es, dass Expertinnen und Experten die Fragen nicht nur als relevant und zielstufenadäquat beurteilen, sondern sie

auch zweifelsfrei richtig beantworten können. Dies war eine Grundbedingung zur Sicherung der Expertenvalidität des Tests. Zu jeden Teilthema wurden zunächst drei Fragenblöcke entwickelt, einer für Schülerinnen und Schüler der 3. bis 4. Klasse, eine für solche der 5. bis 6. Klasse und eine für solche der 7. bis 8. Klasse. Dabei war vorgesehen, dass Schülerinnen und Schüler bei entsprechendem Wissensstand auch Testfragen der angrenzenden Altersstufe zum Lösen erhalten. Anschliessend wurden alle Fragen in der Testumgebung implementiert und für die Dauer des Pilottests mit zusätzlichen Feedbackmöglichkeiten für die Schülerinnen und Schüler ausgestattet (siehe Abb. 3).

Noch nicht ganz!

Einfache und alte Geräte haben keine Computerchips.
Komplizierte und neue Geräte beinhalten aber häufig Computerchips.

Sag uns bitte deine Meinung. Diese Frage war ...

gut OK schlecht weil ...

Ich habe Wörter nicht verstanden.
 Ich wusste nicht, was ich machen soll.
 Ich verstehe nicht, was ich falsch gemacht habe.
 Etwas anderes war schlecht.

... sag uns was schlecht war ...

Abb. 3: Feedbackformular für Schülerinnen und Schüler während des Pilottests.

Empirische Überprüfung

In einem nächsten Schritt wurden die Fragen von einem Convenience-Sample von mehr als 2000 Schülerinnen und Schülern überprüft. Hierfür wurden insbesondere die Kanäle der Stiftung Pro Juventute aktiviert. Aufgrund der hohen Anzahl von Fragen wurde jeder Schülerin und jedem Schüler zunächst eine zufällige Auswahl von 10 Frageblöcken passend zum gewählten Thema und zur Schulstufe zugewiesen. Nachdem jeder Frageblock von mehr als 200 Schülerinnen und Schülern beantwortet wurde, konnte eine erste empirische Skalierung, zunächst innerhalb der Frageblöcke, durchgeführt werden. Dafür wurde ein dichotomes Raschmodell zugrundegelegt und die Passung jedes Items innerhalb einer Itemgruppe in Bezug Item- und

Personenparameter überprüft. In Bezug auf den Itemfit wurde eine Schwierigkeit zwischen .2 und .8 anvisiert, wobei angesichts der hohen Ratewahrscheinlich dichotomer Items auch leichtere Items zugelassen wurden. In Bezug auf den Itemfit und Personenfit wurden Infit- und Outfit-Werte zwischen $.05 < \text{MNSQ} < 1.5$ als akzeptabel betrachtet. Items, die diesen Kriterien nicht entsprachen, wurden ausgeschlossen. Frageblöcke, bei denen sich durch diese Reduktion der Itemzahl keine durchgängig akzeptablen Fit-Werte ergaben, wurden gesamthaft ausgeschlossen. Dadurch ergab sich eine Reduktion der Fragenblöcke um annähernd ein Drittel. Insgesamt wurden in der Pilotphase ausserdem rund 18000 Feedbackbewertungen von den Schülerinnen und Schüler zu den einzelnen Frageblöcken abgegeben. Bei ca. 400 von insgesamt rund 2300 negativen Bewertungen wurde zudem eine schriftliche Begründung abgegeben, warum sie die jeweilige Frage schlecht fanden. Das Schülerfeedback wurde bei der Überarbeitung und Reduktion der Frageblöcke einbezogen. Die Themenmatrix in Tabelle 1 spiegelt für die Testmodule 1 bis 3 bereits die reduzierte Fragenzahl während diese Reduktion für die Module 4 bis 5 noch aussteht. Die Skalierung des Tests ist aktuell noch nicht abgeschlossen und es werden insbesondere frageblockübergreifende Fit-Werte zu berechnen sein, was im laufenden Betrieb erfolgt. Mittlerweile haben mehr als 5000 Schülerinnen und Schüler Module des Tests absolviert.

Technische Umsetzung und Begleitmaterialien

Die technische Implementierung des Tests erfolgte in Form einer Webapplikation, die lediglich einen aktuellen Webbrowser für die Teilnahme voraussetzt. Die Applikation wird auf der Infrastruktur der Pro Juventute betrieben und fortlaufend durch die Pädagogische Hochschule Schwyz evaluiert. Die einzelnen Frageblöcke wurden auf der Plattform LearningApps.org implementiert und werden von dort unmittelbar im Test eingebunden.

Im Wissen um die grosse Heterogenität des Kenntnisstandes der Schülerinnen und Schüler wurde der Test von Anfang an als adaptiver Test konzipiert, d.h. die Testumgebung errechnet nach jedem Fragenblock einen vorläufigen Schätzwert B für die Kompetenz des Teilnehmers und wählt auf dieser Grundlage den nächsten Frageblock so aus, sodass dessen Schwierigkeitsgrad für die Person angemessen ist. Für alle Einzelfragen innerhalb der Frageblöcke wurde ein Schwierigkeitswert D_i mit Hilfe des Raschmodells aus der Pilotierungsphase bestimmt. Für den adaptiven Test wird der Algorithmus von Wright (1988) verwendet. Das Verfahren wurde für die Verwendung von Frageblöcken mit Teilfragen und Teilschwierigkeitswerten geringfügig angepasst (siehe Tab. 2).

Verfahren nach Wright (1988)	angepasstes Verfahren für Frageblöcke
$L_{k+1} = L_k + 1$ $R_{k+1} = R_k + c$ $B_{k+1} = \frac{H_{k+1}}{L_{k+1}} + \ln\left(\frac{R_{k+1}}{W_{k+1}}\right)$	$L_{k+1} = L_k + n$ $R_{k+1} = R_k + \sum_{i=0}^n c_i$ $W_{k+1} = L_{k+1} - R_{k+1}$ $B_{k+1} = \frac{H_{k+1}}{L_{k+1}} + \ln\left(\frac{R_{k+1}}{W_{k+1}}\right)$
<p><i>L</i> ... Gesamtanzahl gestellter Fragen <i>H</i> ... Gesamtschwierigkeit <i>D</i> ... Schwierigkeit der Frage <i>R</i> ... Gesamtanzahl richtiger Antworten <i>W</i> ... Gesamtanzahl falscher Antworten <i>c</i> ... Antworten auf Frage (0=falsch, 1=richtig)</p>	<p><i>n</i> ... Anzahl Teilfragen des Frageblocks <i>D_i</i> ... Schwierigkeit der Teilfrage <i>c_i</i> ... Antwort auf Teilfrage (0=falsch, 1=richtig) <i>B</i> ... Fähigkeitseinschätzung</p>

Tab. 2: Neuberechnung der Variablen des Algorithmus nach Beantwortung der -ten Frage bzw. Frageblocks (ohne Berücksichtigung der Sonderfälle für $R = 0$ bzw. $W = 0$).

Bei diesem Verfahren spielt die Reihenfolge der Fragen für die Bestimmung des Fähigkeitsschätzwerts B keine Rolle. Dieser ergibt sich immer aus dem bisherigen Gesamtschwierigkeitsgrad H und der Gesamtanzahl gestellter Fragen L bezogen auf die Anzahl richtiger (R) zu falschen Antworten (W). Für die Auswahl der Frageblöcke beginnen die Teilnehmenden mit einem mittleren Startschwierigkeitswert. Wie in Linacre (2000) beschrieben ist es sinnvoll, diesen im unteren Drittel des Wertebereichs vorzugeben, dadurch sollen Teilnehmenden nicht frühzeitig durch zu schwierige Fragen demotiviert werden. Bei der Auswahl der nächsten Testfrage wird gegenüber dem klassischen Modell zudem ein geringfügiger Zufallswert einbezogen, um die Wiederspielbarkeit des Tests zu erhöhen. Dabei ist jedoch anzumerken, dass die Anzahl der Testfragen in diesem Test sich für adaptive Verfahren an der untersten Grenze bewegt. Ausserdem entsprechen adaptive Verfahren nicht unbedingt den Anforderungen und Wünschen der Lehrpersonen. Für Lehrpersonen wird es beispielsweise schwieriger einzelne Fragen zu thematisieren, wenn nicht alle Schülerinnen und Schüler den gleichen Fragesatz lesen. Da es ausserdem schwierig sein kann, wenn nicht alle Schülerinnen und Schüler gleich viele Fragen vorgelegt bekommen und sich damit die Testdauer stark unterscheidet, wurde im Gegensatz zu traditionellen adaptiven Tests deshalb nicht eine Fehlergrenze als Abbruchbedingung definiert, sondern eine fixe Anzahl Frageblöcke vorgegeben. Das Testergebnis ergibt sich unmittelbar aus der zuletzt berechneten Fähigkeitsschätzung B , die auf einen Wertebereich von 0 bis 100 Punkte skaliert wird. Um Lehrpersonen eine bessere Planbarkeit zu ermöglichen wurde deshalb auch eine nicht-adaptive Alternative implementiert, bei dem alle Fragen des jeweiligen Themenbereichs gestellt werden. Das Testergebnis wird auch in

diesem Modus auf die gleiche Weise bestimmt, womit schwierige Items eine höhere Gewichtung als einfache Items im Gesamtergebnis aufweisen.

Die Testplattform bietet Lehrpersonen zudem die Möglichkeit Klassen zu erstellen, um Testergebnisse zu aggregieren und damit das mittlere Kompetenzlevel einer gesamten Klasse bestimmen zu können. Zusätzlich zu diesen Auswertungen liefert der Test eine konkrete Auswahl an möglichen Lehrmitteln und Unterrichtsmaterialien, mit denen die Kompetenzen in den verschiedenen Themenbereichen verbessert werden können (vgl. Abb. 4).

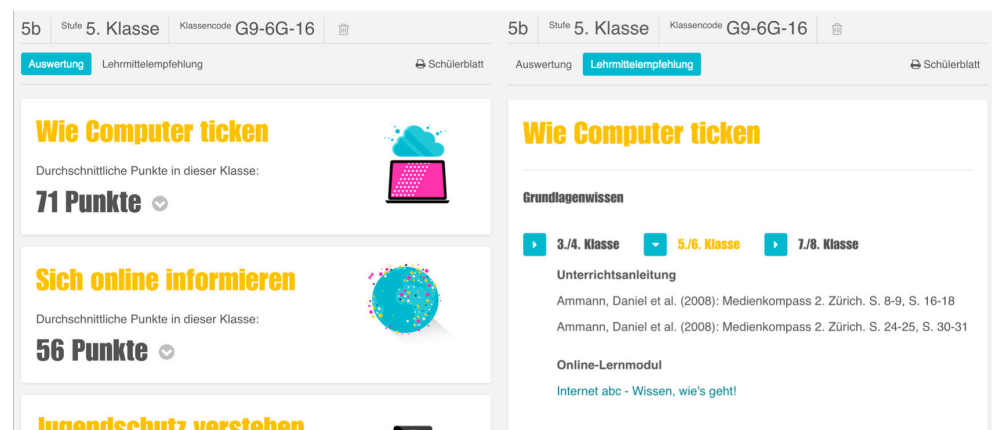


Abb. 4: Auswertungsscreen und weiterführende Hinweise.

Diskussion

Die Messung von Medienkompetenzen erfolgt heute in drei verschiedenen Angebotsformen: in kommerziellen Prüfungen im Kontext von Weiterbildungsangeboten, in wissenschaftlichen Tests im Kontext empirischer Studien und in spielerischen Tests und Quizzes im Kontext von Sensibilisierungskampagnen. Für den praktischen Einsatz in Schulen und anderen pädagogischen Kontexten sind die bisher bestehenden Angebote aus unterschiedlichen Gründen jedoch unbefriedigend. Während die kommerziellen Prüfungen normalerweise marken- und produktgebundenes Anwendungswissen abfragen, sind wissenschaftliche Tests weder öffentlich verfügbar noch für statistische Laien auswertbar. Die bestehenden Quizzes wiederum sind zwar teilweise attraktiv aufgemacht, lassen bei der Begutachtung ihrer Fragen jedoch Zweifel an ihrer Relevanz und Validität aufkommen. Der hier vorgestellte «Medienprofis-Test» versucht, hier neue Wege zu gehen und ein spielerisches Angebot zu entwickeln, das öffentlich verfügbar ist und ausgewählte Aspekte der Medienkompetenz in empirisch validierter Weise abfragt. Das Projekt zeigt, dass Tests und Quizzes dabei nicht notwendigerweise einen Widerspruch darstellen. Es ist möglich, Quizzes zu entwickeln, die theoretisch fundiert sind und bei denen die Auswahl der Fragen

aus konzeptionellen Überlegungen abgeleitet werden, die empirisch überprüft sind und wissenschaftlichen Kriterien standhalten und die trotzdem für die Praxis handhabbar sind. Der Medienprofis-Test stellt einen ersten Versuch dar, der freilich auch noch einige Schwächen aufweist. Dazu gehört nicht zuletzt, dass er nicht die ganze Breite möglicher Themen abdeckt, dass die Testfragen noch überwiegend textbasiert und nicht multimedial sind und dass auch der spielerische Charakter noch stärker einbezogen werden könnte. Damit könnten das Engagement und die Motivation der Schülerinnen und Schüler zusätzlich erhöht werden und dem Test mehr Ähnlichkeit zu realen Anwendungssituationen verliehen werden. Nichtsdestotrotz soll mit diesem ersten Schritt versucht werden, der Medienbildung ein standardisiertes und validiertes Instrument zur Verfügung zu stellen und gleichzeitig die Diskussion um die Ausprägungen von Medienkompetenz mit empirisch Daten zu bereichern. Im Schulalltag kann der Test verwendet werden, um das Kompetenzlevel von Schülerinnen und Schülern in den verfügbaren Themenbereichen zu bestimmen. Im Sinne einer Bedarfsanalyse können die Lehrpersonen vorgängig feststellen, in welchen Bereichen Defizite vorhanden sind und im Unterricht entsprechende Prioritäten setzen. Mögliche Lehrmittel zur Bearbeitung einzelner Themenbereiche werden dabei bei der Testauswertung ebenfalls mitgeliefert und erleichtern so die Unterrichtsplanung und -gestaltung.

Ab Juni 2017 liegt das Instrument vollständig skaliert vor und ist damit in seiner Entwicklung vorläufig abgeschlossen.

Literatur

- Baacke, Dieter. 1997. *Medienpädagogik. Grundlagen der Medienkommunikation*. Tübingen: Niemeyer.
- Bawden, David. 2001. «Information and digital literacies: a review of concepts». *Journal of documentation*, 57, 2, 218–59.
- Bühner, Markus. 2011. *Einführung in die Test- und Fragebogenkonstruktion*. 3. Aufl. München: Pearson Studium.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge: MIT Press.
- Cox, Kevin, und David Clark. 1998. «The use of formative quizzes for deep learning». *Computers and Education*, 30, 3, 157–68.
- D-EDK. 2017. *Medien und Informatik. Bedeutung und Zielsetzungen*. <http://v-ef.lehrplan.ch/index.php?code=e|10|2>.
- Eickelmann, Birgit, Wilfried Bos, Julia Gerick, und Julia Kahnert. 2014. «Anlage, Durchführung und Instrumentierung von ICILS 2013». In *ICILS 2013. Computer- und informationsbezogene Kompetenzen von Schülerinnen und Schülern in der 8. Jahrgangsstufe im internationalen Vergleich*, herausgegeben von Wilfried Bos, Birgit Eickelmann, Julia Gerick, Frank Goldhammer, Heike Schaumburg, Knut Schwippert, Martin Senkbeil, Renate Schulz-Zander, und Heike Wendt, 43–82. Münster: Waxmann.

- Fraillon, Julian, Wolfram Schulz, Tim Friedman, John Ainley, und Eveline Gebhardt. 2015. *ICILS 2013 Technical Report*. Amsterdam: International Association for the Evaluation of Educational Achievement (IEA).
- Fraillon, Julian, Wolfram Schulz, Tim Friedman, John Ainley, und Eveline Gebhardt. 2013. *Preparing for Life in a Digital Age: The IEA International Computer and Information Literacy Study*. Melbourne: Springer Open.
- Gapski, Harald. 2001. *Medienkompetenz. Eine Bestandsaufnahme und Vorüberlegungen zu einem systemtheoretischen Rahmenkonzept*. Wiesbaden: Westdeutscher Verlag.
- Gapski, Harald, Hrsg. 2006. *Medienkompetenzen messen? Verfahren und Reflexionen zur Erfassung von Schlüsselkompetenzen*. Düsseldorf/München: kopaed.
- Granic, Isabela, Adam Lobel, und Rutger C. M. E. Engels. 2014. «The benefits of playing video games». *American Psychologist*, 69, 1, 66–78.
- Jenkins, Henry, Katie Clinton, Ravi Purushotma, Alice J. Robison, und Margaret Weigel. 2009. *Confronting the Challenges of Participatory Culture: Media Education for the 21st Century*. White Paper. Chicago: MacArthur Foundation.
- Katz, Irvin R. 2007. «Testing Information Literacy in Digital Environments: ETS's iSkills Assessment». *Information Technologies and Libraries*, 26, 3, 3–12.
- Klieme, Eckhard, Hermann Avenarius, Werner Blum, Peter Döbrich, Hans Gruber, Manfred Prenzel, Kristina Reiss, Kurt Riquarts, Jürgen Rost, Heinz-Elmar Tenorth, und Helmut J. Vollmer. *Zur Entwicklung nationaler Bildungsstandards. Eine Expertise*. <http://www.intranet.bbsii-kl.de/media/Unterricht/Lehr-%20und%20Lernforschung/Klieme%20u.%20a.%20Expertise%20Bildungsstandards.pdf>.
- Klieme, Eckhard, und Johannes Hartig. 2008. «Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs». In *Kompetenzdiagnostik: Zeitschrift für Erziehungswissenschaft. Sonderheft 8, herausgegeben von Manfred Prenzel, Ingrid Gogolin, und Heinz-Hermann Krüger*, 11–29. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Lee, Joey J., und Jessica Hammer. 2011. «Gamification in Education: What, How, Why Bother?». *Academic Exchange Quarterly*, 15, 2, 1-5.
- Linacre, John Michael. 2000. «Computer-Adaptive Testing: A Methodology Whose Time Has Come». In *Development of Computerized Middle School Achievement Test. MESA Memorandum 69*, herausgegeben von Sunhee Chae, Unson Kang, Euhwa Jeon, und John Michael Linacre. Seoul: Komesa Press. <http://rasch.org/memo69.htm>.
- Marotzki, Winfried, und Benjamin Jörissen. 2010. «Dimensionen strukturaler Medienbildung». In *Jahrbuch Medienpädagogik 8*, herausgegeben von Bardo Herzig, Dorothee M. Meister, Heinz Moser, Horst Niesyto. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Mills, Kathy Ann. 2010. «A Review of the «Digital Turn» in the New Literacy Studies». *Review of Educational Research*, 80, 2, 246–71.
- Moser, Heinz. 2012. «Bildungsstandards im Medienbereich». In *Jahrbuch Medienpädagogik 9*, herausgegeben von Renate Schulz-Zander, Birgit Eickelmann, Heinz Moser, Horst Niesyto, und Petra Grell, 249–269. Wiesbaden: Springer VS.

- Moser, Urs, und Florian Keller. 2004. *Test Your ICT. Knowledge: Ausbilden - Unterrichten - Testen*. Zürich: KBL.
- Oelkers, Jürgen, und Kurt Reusser. 2008. *Expertise: Qualität entwickeln, Standards sichern, mit Differenz umgehen*. Berlin: Bundesministerium für Bildung und Forschung, Referat Bildungsforschung.
- Petko, Dominik. 2011. «Praxisorientierte medienpädagogische Forschung: Ansätze für einen empirischen Perspektivenwechsel und eine stärkere Konvergenz von Medienpädagogik und Mediendidaktik». Herausgegeben von Heinz Moser, Petra Grell, und Horst Niesyto. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung* 20 Medienbildung im Spannungsfeld medienpädagogischer Leitbegriffe sowie Medienbildung und Medienkompetenz. Beiträge zu Schlüsselbegriffen der Medienpädagogik. München: kopaed: 245–58. doi:10.21240/mpaed/20/2011.09.22.X.
- Richter, Tobias, Johannes Naumann, und Norbert Groeben. 2001. «Das Inventar zur Computerbildung (INCOBI): Ein Instrument zur Erfassung von Computer Literacy und computerbezogenen Einstellungen bei Studierenden der Geistes- und Sozialwissenschaften». *Psychologie in Erziehung und Unterricht*, 48, 1–13.
- Richter, Tobias, Johannes Naumann, und Holger Horz. 2010. «Eine revidierte Fassung des Inventars zur Computerbildung (INCOBI-R)». *Zeitschrift für pädagogische Psychologie*, 24, 1, 23-37.
- Rost, Jürgen. 2004. *Lehrbuch Testtheorie – Testkonstruktion*. Bern: Huber.
- Rychen, Dominique S., und Laura H. Salganik. 2003. *Key competencies for a successful life and a well-functioning society*. Cambridge: Hogrefe & Huber.
- Schelhowe, Heidi, Silke Grafe, Bardo Herzig, Jochen Koubek, und Horst Niesyto. 2009. *Kompetenzen in einer digital geprägten Kultur. Medienbildung für die Persönlichkeitsentwicklung, für die gesellschaftliche Teilhabe und für die Entwicklung von Ausbildungs- und Erwerbsfähigkeit*. http://www.dlr.de/pt/en/Portaldata/45/Resources/dokumente/bildungsforschung/Expertenkommission_Maerz_2009.pdf.
- Senkbeil, Martin, Jan Marten Ihme, und Jörg Wittwer. 2013a. «Entwicklung und erste Validierung eines Tests zur Erfassung technologischer und informationsbezogener Literacy (TILT) für Jugendliche am Ende der Sekundarstufe I». *Zeitschrift für Erziehungswissenschaft*, 16, 4, 671–91.
- Senkbeil, Martin, Jan Marten Ihme, und Jörg Wittwer. 2013b. «The Test of Technological and Information Literacy (TILT) in the National Educational Panel Study: Development, empirical testing, and evidence for validity». *Journal for educational research online*, 5, 2, 139-61.
- Somerville, Marry M., Gordon W. Smith, und Alexius Smith Macklin. 2008. «The ETS iSkills™ Assessment: a digital age tool». *The Electronic Library*, 26, 2, 158–171.
- Sowka, Alexandra, Christoph Klimmt, Dorothee Hefner, Fenja Mergel, und Daniel Possler. 2015. «Die Messung von Medienkompetenz. Ein Testverfahren für die Dimension «Medienkritikfähigkeit» und die Zielgruppe «Jugendliche»». *M&K Medien & Kommunikationswissenschaft*, 63, 1, 62–82.

- Tulodziecki, Gerhard. 2011. «Zur Entstehung und Entwicklung zentraler Begriffe bei der pädagogischen Auseinandersetzung mit Medien». Herausgegeben von Heinz Moser, Petra Grell, und Horst Niesyto. *MedienPädagogik: Zeitschrift für Theorie und Praxis der Medienbildung* 20 Medienbildung im Spannungsfeld medienpädagogischer Leitbegriffe sowie Medienbildung und Medienkompetenz. Beiträge zu Schlüsselbegriffen der Medienpädagogik. München: kopaed: 11–39. doi:10.21240/mpaed/20/2011.09.11.X.
- Wagener, Dietrich. 2003. «Der Computerwissenstest CWIS-4: Befunde zur Reliabilität und Validität». *Zeitschrift für Personalpsychologie*, 2, 4, 169–181.
- Wagener, Dietrich. 2008. *START-C: Testbatterie für Berufseinsteiger – Computer*. Göttingen: Hogrefe.
- Weinert, Franz E. 2001. «Concept of Competence: A Conceptual Clarification». In *Defining and Selecting Key Competencies*, herausgegeben von Dominique S. Rychen, Laura Hersh Salganik, 45-65. Ashland: Hogrefe & Huber.
- Wright, B.D. 1988. «Practical adaptive testing». *Rasch Measurement Transactions* 2, 2, 24. <http://rasch.org/rmt/rmt22g.htm>.

Abbildungen

- Abb. 1:** Übersicht über die verfügbaren Testmodule von www.medienprofis-test.ch.
- Abb. 2:** Auswertung einer dichotomen Entscheidungsfrage: Woran kann es liegen, dass Suchanfragen in einer Suchmaschine bei verschiedenen Personen zu unterschiedlichen Ergebnissen führen?
- Abb. 3:** Feedbackformular für Schülerinnen und Schüler während des Pilottests.
- Abb. 4:** Auswertungsscreen und weiterführende Hinweise.

Tabellen

- Tab. 1:** Themenmatrix von www.medienprofis-test.ch (Stand 14.03.2017).
- Tab. 2:** Neuberechnung der Variablen des Algorithmus nach Beantwortung der -ten Frage bzw. Frageblocks (ohne Berücksichtigung der Sonderfälle für $R = 0$ bzw. $W = 0$).