

Dr. phil. Anke Steckelberg

Christian Hülfenhaus

Projektbericht

Projektlaufzeit: 1.7.2005-31.10.2006



ebm @ school

Gesundheitsbildung für Schülerinnen + Schüler

Projektbericht über die Entwicklung und Pilotierung eines EBM Curriculums für Schülerinnen und Schüler Allgemeinbildender Schulen und die Entwicklung und Validierung eines Instruments zur Messung der Kompetenz *critical health literacy*

Inhaltsverzeichnis

Zusammenfassung

Einleitung

1. Entwicklung des Curriculums *ebm@school* und Ergebnisse der Pilotkurse

- 1.1 Bedingungsanalyse
- 1.2 Zielsetzung der Schulung
- 1.3 Inhalte
- 1.4 Evaluation
- 1.5 Pilotkurse
 - 1.5.1 Planung und Durchführung
 - 1.5.2 Ergebnisse
 - 1.5.2.1. Pilotkurs Neumünster – qualitative Ergebnisse
 - 1.5.2.2 Pilotkurs Harburg – qualitative Ergebnisse
 - 1.5.2.3 Ergebnisse quantitativ
- 1.6 Aufnahme der Rahmenkriterien und Methoden in das Curriculum

2. Entwicklung eines Testinstruments zur Messung der Kompetenz *critical health literacy*

- 2.1 Testentwicklung Phase I: Entwicklung der Testhefte
- 2.2 Testentwicklung Phase II: 1. Feldtest
 - 2.2.1 Stichprobe und Setting
 - 2.2.2 Ergebnisse
 - 2.2.3 Itemrevison
- 2.3 Testentwicklung Phase III: 2. Feldtest
 - 2.3.1 Stichprobe und Setting
 - 2.3.2 Ergebnisse
- 2.4 Zur Handhabung des fertigen Testinstruments: *critical health literacy*

3. Interaktive Lernelemente

- 3.1 Entwicklung
- 3.2 Evaluation
- 3.3 Ergebnisse
- 3.4 Ausblick

4. Ausblick

Literatur

Anlagen zum Projektbericht

Zusammenfassung

Patienten und Verbraucher benötigen Gesundheitskompetenzen als Voraussetzung für die Partizipation in gesundheits- und krankheitsbezogenen Entscheidungen.

Bisherige Präventionsprogramme knüpfen an das paternalistische Konzept an und zielen ausschließlich auf *Compliance* ab. Für eine kritische Auseinandersetzung einschließlich einer individuellen Bewertung im Sinne der Evidenz-basierten Medizin werden die Nutzer von Gesundheitsinformationen nicht geschult. Kompetenzkurse wurden bisher nur für Ärzte angeboten. Aus der Arbeitsgruppe Mühlhauser gibt es erste Trainingskurse in wissenschaftlichen Kompetenzen auch für Diabetesberater, Lehrer, sowie Patienten- und Verbrauchervertreter.

In einem 1-Jahres Pilotprojekt wurde als Novum ein Kompetenztraining für Schüler Allgemeinbildender Schulen entwickelt und erfolgreich erprobt. Zudem wurde in Kooperation mit Prof. Rost vom Institut für die Pädagogik der Naturwissenschaften in Kiel ein Testinstrument zur Messung der zu fördernden Kompetenz *critical health literacy* entwickelt und steht nunmehr sowohl zur Curriculumevaluation als auch zur Kompetenzmessung in anderen Settings zur Verfügung.

Einleitung

Patienten und Verbraucher übernehmen zunehmend Verantwortung bei diagnostischen und therapeutischen Entscheidungen. Die dafür erforderlichen Kompetenzen fehlen oft (1). So kann die kommerzielle Vermarktung von erfundenen Krankheiten (*disease mongering*) Ängste schüren und die Verunsicherung der Bevölkerung ausnutzen, da diese auf eine kritische Auseinandersetzung und Bewertung diagnostischer und therapeutischer Verfahren nicht vorbereitet ist (2, 3). Eine Kompetenzsteigerung bei Verbrauchern und Patienten, die auf Autonomie und informierte Entscheidung abzielt, sollte möglichst früh ein kritisches Verständnis von Wissenschaft und im Besonderen von Gesundheitswissenschaft und Medizin fördern. Auch für die von uns adressierten Jugendlichen haben Gesundheitsthemen bereits eine hohe Relevanz. Sie nutzen Beratungsstellen, nehmen selbständig Arztbesuche wahr und werden zudem gezielt von Werbekampagnen adressiert. Die Kompetenz, in diesen Situationen informierte Entscheidungen in gesundheitsrelevanten Fragen zu treffen, fehlt (4). Vorarbeiten haben gezeigt, dass Schüler Risikoinformationen nicht verstehen (5). Neben diesem individuellen Nutzen lässt die Kompetenzsteigerung im Bereich der *critical health literacy* auch einen kollektiven Nutzen erwarten, da zum einen eine Verbesserung der Behandlungsqualität erreicht werden kann und zum anderen die Effizienz der Versorgung erhöht wird und es damit insgesamt zu einer Verbesserung des Gesundheitssystems kommen kann. Die *National Academies* in den USA fordern die Einbeziehung von *health literacy* in die Curricula der Schulen und Kindergärten (6). Auch international wird die Verbesserung der *health literacy* als Herausforderung für das 21. Jahrhundert dargestellt (7). Dennoch sind bisherige Projekte, die im Rahmen der WHO Kampagne *health promoting schools* europaweit in Schulen eingeführt wurden (8, 9), in ihrem Ansatz paternalistisch ausgerichtet und zielen nicht auf eine kritische Gesundheitsbildung ab.

Der Erwerb von Kompetenzen, die das Maß an Selbstbestimmung über die Gesundheit erhöhen, setzt eine Verknüpfung von *health literacy* mit dem Konzept der Evidenz-basierten Medizin (EBM) voraus. Kurse, die beide Aspekte berücksichtigen, sind lediglich im Bereich der Aus- und Weiterbildung für Ärzte etabliert. Bisher gibt es nur vereinzelte Projekte in den USA und Großbritannien, die versuchen, Patienten und Verbrauchervertreter in *critical health literacy* zu schulen (10, 11). Das *ebm@school* Curriculum ist international ohne Vorbild und adressiert erstmalig

Schülerinnen und Schüler Allgemeinbildender Schulen. In Vorstudien konnten wir zum einen zeigen, dass Schüler Interesse an Gesundheitsthemen haben und zum anderen konnten wir Interessensbereiche identifizieren (12). Die Resonanz der Schulen war sehr positiv. Bereits in der Planungsphase dieses Pilotprojektes haben wir Kooperationszusagen für die geplante Implementierung unter kontrollierten Bedingungen erhalten. Die Schulen, die am Feldtest unseres Testinstruments teilgenommen haben, zeigten ebenfalls großes Interesse. Neben den von uns geplanten Kursen, gibt es aus den Schulen konkrete Vorschläge für die Implementierung der Inhalte in den Schulalltag.

Über die Kontakte zu den Eltern der Schüler, die am Pretest teilgenommen haben, wurde deutlich, dass auch dort ein ausgeprägtes Interesse vorhanden ist. Die Projektidee wurde zum Teil auch in die Elternratsgremien hineingetragen.

Dieses Projekt hat mit der Entwicklung des Curriculums und des Testinstruments sowie der Überprüfung in Pilotkursen die notwendigen Voraussetzungen geschaffen, das Curriculum nunmehr in einem zur Wirksamkeitsüberprüfung geeigneten Design zu überprüfen.

Im Folgenden werden zunächst das Konzept des Curriculums und die Ergebnisse der Pilotkurse präsentiert. Anschließend wird die Entwicklung und Validierung des Testinstruments vorgestellt.

1. Entwicklung des Curriculums *ebm@school* und Ergebnisse der Pilotkurse

1.1 Bedingungsanalyse

Die Schülerinnen und Schüler der adressierten Zielgruppe sind zwischen 16 und 18 Jahre alt und besuchen die 11. Klasse eines Gymnasiums, eines Fachgymnasiums oder einer Gesamtschule. Der Anteil der Schüler mit Migrationshintergrund liegt in Hamburg im Mittel bei 16% mit einer Spanne von 1%–91% zwischen den Schulen (13). Die Zielgruppe übernimmt zunehmend Verantwortung für die, die Gesundheit betreffenden Angelegenheiten. Zudem werden Sie als Zielgruppe von Marketingkampagnen adressiert, ohne dass Sie in der Lage wären, sich evidenzbasiert damit auseinanderzusetzen.

Die Schulen können das Curriculum entweder in Projektwochen oder aber über einen längeren Zeitraum angeboten bekommen. Zu berücksichtigen sind zum einen eine sehr langfristige Planung, da in den Schulen in der Regel für ein Schuljahr geplant wird sowie die sehr unterschiedliche Ausstattung der Schulen, insbesondere der Zugang zu Computern und zum Internet.

1.2 Zielsetzung der Schulung

Das Ziel dieses Curriculums ist die Vermittlung von *critical health literacy*. Zusammenfassend kann als Ziel formuliert werden, dass Schülerinnen und Schüler den Nutzen der eigenständigen Informationsbeschaffung und Bewertung erkennen und in der Lage sind, diese selbständig durchzuführen. Diese Kompetenz ermöglicht es ihnen, vielfältige Fragestellungen zu bearbeiten. Sie können die Möglichkeiten und Grenzen gesundheitswissenschaftlicher und medizinischer Erkenntnisse reflektieren und zeigen eine kritische Haltung gegenüber Gesundheitsinformationen. Die Kompetenzen werden anhand von relevanten Fallbeispielen erworben, die exemplarischen Charakter haben (problemorientierter Ansatz)

1.3 Inhalte

Aus dieser Zielsetzung heraus werden für das Curriculum *ebm@school* 6 Module, die insgesamt 22 Unterrichtsstunden umfassen, konzipiert. Die Module knüpfen inhaltlich an die Themen an, die in einer Vorstudie im September / Oktober 2005 in einer Befragung von 138 Schülerinnen und Schülern als Interessengebiete der Schüler identifiziert worden waren: Ernährung/Lebensmittel, Therapien und

Diagnostik, Sport und Fitness, Unfälle/Sportverletzungen oder aber als Gesundheitsbedrohungen genannt worden waren: Rauchen, Ernährung, Umwelteinflüsse (12).

Ergänzt wird das Curriculum um 10-12 Stunden Projektarbeit, welche den Schülern ermöglicht, parallel zum Unterricht an selbst gewählten Fragestellungen zu arbeiten, das Erlernte zu wiederholen und zu vertiefen.

Modul 1

: Trugschlüsse und Fehleinschätzungen in der Medizin; verwirrende Informationsdarstellungen; Beobachtungsstudie versus randomisiert-kontrollierte Studie am Beispiel der Vitaminsubstitution zur Prävention von Lungenkrebs.

Modul 2

: Beurteilung von randomisiert-kontrollierten Studien: Erarbeitung von Beurteilungskriterien an der in Modul 1 eingeführten Studie

Modul 3

: Der Diagnostische Test: Erarbeitung von Testgütekriterien am Beispiel des Depressionsscreenings von Jugendlichen

Modul 4

: Systematische Übersichtsarbeiten: Erarbeiten der Methodik und Bedeutung anhand einer systematischen Übersichtsarbeit zur Diagnostik von Kreuzbandrupturen.

Modul 5

: Internet- und Datenbankrecherche: Formulieren einer suchtauglichen Fragestellung, Einführung in systematische Recherchen, Übungen in MedPilot.

Modul 6

: Bewertung von Patienteninformation: Diskussion von Bewertungskriterien und Analysen von ausgewählten Patienteninformationen.

1.4 Evaluation

Zum Zeitpunkt der Evaluation befand sich das Testinstrument noch in der Entwicklungsphase. Für einen vorher – nachher Vergleich wäre ein Testaufwand von 180 Minuten notwendig gewesen. Dieses erschien im Hinblick auf die zur Verfügung stehende Zeit für den Unterricht unangemessen hoch. Deshalb wurden zum Vergleich andere 11. Klassen zum gleichen Zeitpunkt getestet und fungierten als Vergleichsgruppe.

Anknüpfend an die Methodik der internationalen Bildungsstudien PISA (14) und TIMMS (15) wurde im Rahmen des Projektes *ebm@school* erstmalig ein Testinstrument entwickelt, welches die Kompetenz der *critical health literacy* misst

(siehe Kapitel 2) und sowohl zur Überprüfung der Wirksamkeit des Curriculums als auch unabhängig vom Curriculum eingesetzt werden kann.

A priori wurde für den Prozess der Entwicklung und der Auswertung die probabilistische Testtheorie zu Grunde gelegt. Das Rasch-Modell definiert die Wahrscheinlichkeit einer Antwort eines Probanden auf eine Frage. Mit dem Modell ist es möglich, sowohl die Kompetenz, als auch die einzelnen Fragen auf einer Skala abzubilden (16). Dieses Verfahren bietet im Vergleich zur klassischen Testtheorie 3 Vorteile:

- Untersuchungsökonomisch: Für die Kompetenzmessung reichen wenige Fragen aus. Da ein jedes Item eine bestimmte Position auf der Raschskala hat, kann die Fähigkeit einer Person – anders als in der klassischen Testtheorie – mit ganz wenigen Items zuverlässig geschätzt werden.
- Dieselbe Eigenschaft der Items ermöglicht den Einsatz von unterschiedlichen Items derselben Skala zur Wiederholungsmessung in Studien mit mehreren Messzeitpunkten.
- Probabilistisch konstruierte Tests erlauben die Zusammenstellung verschiedener Teilinstrumente zur Anwendung auf verschiedenen Leistungslevels, womit Frustration der Schüler verringert werden kann.

1.5 Pilotkurse

1.5.1 Planung und Durchführung

Bei der Auswahl der Schulen wurde berücksichtigt, dass die beiden Schulen im Hinblick auf die Zusammensetzung der Schülerinnen und Schüler, deren Lernvoraussetzungen und Leistungsniveaus sehr unterschiedlich sein sollten.

Anhand dieser Kriterien wurden eine 11. Klasse des Fachgymnasiums Neumünster und eine 11. Klasse der Gesamtschule Harburg ausgewählt.

Für die erste Erprobung des Curriculums (feasibility) wurden 2 Pilotkurse durchgeführt. Im 1. Pilotkurs wurden insbesondere Daten zu den Unterrichtsmethoden und zum Umfang und Schweregrad der Unterrichtsmaterialien erhoben. Feedbacks wurden schriftlich und mündlich erhoben. Zudem wurde der Unterricht von einem der beiden Forscher (CH oder AS) systematisch beobachtet im Hinblick auf Lehrmethoden und Schülerreaktionen auf Arbeitsmaterialien (Reader, Folien etc.) und Fragestellungen. Das Material wurde daraufhin für den 2. Pilotkurs überarbeitet.

Im 2. Pilotkurs wurden Lehrmethoden und Schülerreaktionen analog zum 1. Pilotkurs systematisch beobachtet. Zusätzlich wurde der von den Schülern empfundene Nutzen anhand von Meta Plan Karten erhoben. Zudem wurden mit dem CHC Test die Kompetenzen gemessen.

1.5.2 Ergebnisse

1.5.2.1 Pilotkurs Neumünster – qualitative Ergebnisse

Der 1. Pilotkurs wurde im Fachgymnasium Neumünster in einer Klasse mit 17 Schülerinnen und 3 Schülern durchgeführt. Das Alter der Schüler lag zwischen 16 und 19 Jahren. Da der Jahrgang in dieser Woche eine Projektwoche durchführte, wurde der Pilotkurs mit einer Projektgruppe kombiniert. In Übereinkunft mit der Schulleitung wurde die Projektarbeit von den Dozenten des Projektes *ebm@school* betreut.

Der Pilotkurs konnte ohne größere Probleme durchgeführt werden, da das Konzept insgesamt zum einen auf Interesse bei den Schülern stieß und zum anderen bzgl. des Leistungsniveaus der Klasse angemessen war. Somit erwies sich der 1. Pilotkurs insgesamt als praktikabel. Zudem lieferten die Rückmeldungen der Schüler wichtige Informationen zur Modifikation und damit Optimierung des Curriculums insgesamt und der Arbeitsmaterialien im Besonderen. Die Ausweitung des Curriculums um die Phasen der offenen Projektarbeit erwies sich als sehr sinnvolle Ergänzung im Hinblick darauf, dass:

- die unmittelbare Anwendung des neu Erlernten auf eigene gesundheitsrelevante Fragestellungen und damit verbunden eine weitere Auseinandersetzung und Reflexion der individuellen Überzeugungen und Einstellungen ermöglicht wird
- die Wiederholung und Vertiefung und eine damit verbundene Konsolidierung des neu Erlernten stattfindet

1.5.2.2 Pilotkurs Harburg – qualitative Ergebnisse

Der 2. Pilotkurs wurde in einer 11. Klasse der Gesamtschule Harburg durchgeführt. In der Klasse waren 26 Schüler im Alter von 17-19 Jahren. 13 Schüler und 7 Schülerinnen nahmen an der Evaluation teil.

Die Rahmenbedingungen des Pilotkurses waren schwierig, da die Parallelklassen kurz vor den Ferien kaum noch Unterricht hatten und für die Projektwoche keine Noten vergeben werden sollten. Schüler, die wussten, dass sie nicht versetzt werden würden, nahmen kaum noch am Unterricht teil. Zudem wurden einzelne Unterrichtssequenzen von anderen Lehrern unterbrochen, um z.B. Umfragen durchzuführen, Klassenarbeiten etc. zurückzugeben. In der Klasse waren überwiegend Jungen, wodurch das Arbeitsklima sehr geprägt wurde, denn es waren durchweg die Jungen, die Unterrichtsstörungen hervorriefen und Probleme hatten, sich zu konzentrieren.

Zudem waren in der Klasse verschiedene Voraussetzungen nicht gegeben. Die Schüler hatten kaum Erfahrung mit Gruppenarbeiten und im Besonderen auch nicht mit Projektarbeit. Das Leistungsniveau war sehr niedrig, obgleich die Spanne innerhalb der Klasse recht groß war. Das heißt, dass ca. 1/3 der Klasse mit dem Lesen von deutschen Texten, die mehr als eine Seite umfassten, überfordert waren, andere dagegen auch ein englisches Abstract bearbeiten konnten. Das bedeutet, dass es für besonders leistungsschwache Schüler alternativ zu den Arbeitsmaterialien didaktisch weiter reduziertes Material geben sollte, um in solch extrem heterogenen Klassen auch diese leistungsschwachen Schüler ihren Fähigkeiten entsprechend zu schulen. Nachdem z.B. Schüler in einem Feedback angemerkt hatten, dass ihnen manche Dinge zu schwer seien, wurde das Niveau für das Thema Meta-Analyse gesenkt. Die wenigen leistungsstarken Schüler bekamen die Originalversion. Die leistungsschwachen Schüler erhielten eine Kurzfassung des deutschen Abstracts. Nach der modifizierten Version des Moduls 4 meldeten die Schüler zurück, dass Sie es jetzt besser verstehen könnten. Abschließend wurden die Schüler gefragt: „Was nehmt ihr aus der Woche für euch persönlich mit?“ (Antworten siehe Box 1)

Einzelne Schüler nahmen aufgrund der Anwesenheitspflicht am Unterricht teil, waren aber zur Mitarbeit nicht mehr bereit. Diese Schwierigkeiten gingen über übliche motivationale Probleme hinaus und lassen sich im Rahmen eines Pilotkurses nicht lösen. Die hohen Fehlzeiten in der Klasse und im gesamten Jahrgang führten zu der Entscheidung, für die geplante Implementierungsstudie die Fehlzeiten der Schüler zu dokumentieren.

Box 1: Schülerantworten des 2. Pilotkurses auf die Frage: „Was nehmt ihr aus der Woche für euch persönlich mit“?

1. *Den Unterschied zwischen relativer und absoluter Risikoreduktion.*
2. *Relative und absolute Risikoreduktion.*
3. *Aufklärung der irreführenden Prozentzahlen in der Presse*
4. *Ich habe viel dazu gelernt und kann jetzt genauere Fragen an meinen Arzt stellen, um so bessere Antworten zu bekommen. Die könnte bei der Auswahl von Medikamenten helfen.*
5. *Neues Wissen über Risiken und Forschungsmethoden in der Medizin.*
6. *Schärferes Denken über Statistiken; Fragen, die ich dem Arzt stellen kann; Wissen über einige Fachbegriffe*
7. *Ich nehme eine Vielzahl an neuen Dingen mit, die mich für eine bevorstehende Konfrontation mit einem Arzt oder mit medizinischen Studien wappnen. Vielen Dank dafür!*
8. *Den kritischen Blick auf medizinische Berichte; das genaue Ansehen von Statistiken.*
9. *Wenn ich zum Arzt nächstes Mal gehe, werde ich aufpassen, was er sagt und werde auch Fragen stellen, was ich nicht verstehe.*
10. *medpilot.de; den Ordner*
11. *Kritische Herangehensweise bei Statistiken und Werten; Ordner nehme ich mit.*
12. *Ich werde meinen Arzt besser hinterfragen und mich besser über Medikamente informieren, bevor ich sie einnehme oder nicht. Wo ich mich informieren kann weiß ich jetzt durch diesen Kurs.*
13. *Ich werde besser die Packungsbeilagen lesen, wenn ich ein Medikament einnehmen will: Ich habe viel Neues in dieser Woche gelernt.*
14. *Besseres Verständnis für den med. Sektor, vor allem im Bereich von Forschung und Studien.*
15. *Studien kann ich besser deuten, Prozentangaben anzweifeln. Wie Studien durchgeführt werden, fand ich sehr interessant. Packungsbeilagen verstehe ich nun bei den Nebenwirkungen besser.*
16. *Ich habe viel aus dieser Woche gelernt. Gut fand ich, dass es viel Anschauungsmaterial gab. Die Mappe ist sehr gut gestaltet, sodass wir auch zuhause alles nachlesen können.*
17. *Ich hasse es, mit Computern zu arbeiten.*
18. *Nichts.*
19. *Nichts.*

1.5.2.3 Ergebnisse quantitativ

Als Grundlage für die Abschätzung der Effekte des *ebm@school* Curriculums lagen aus der Pilotstudie und der Feldstudie Kompetenzdaten von 37 trainierten Schülern und 218 untrainierten Schülern vor (Tabelle 1).

Tabelle 1. Basisdaten

| | mit Training (<i>n</i> =37) | ohne Training (<i>n</i> =218) |
|---|------------------------------|--------------------------------|
| Mittleres Alter in Jahren (±SEM) | 17,6 (± 0,1) | 17,4 (± 0,1) |
| Frauen (%) | 20 (56) | 138 (64) |
| Muttersprache deutsch (%) | 28 (78) | 182 (84) |

Die Aussagekraft der Datenbasis für den Gruppenvergleich ist mit verschiedenen Limitierungen verbunden. Das Design ermöglicht keine Aussage über die Wirksamkeit des Unterrichts. Die Durchführung zweier Pilotkurse in ausgewählten Schulen war mit einer klaren Fragestellung verbunden und zielte darauf ab, Aspekte zu identifizieren, die den Unterricht optimieren können. Diese Ziele brachten es mit sich, dass Teile der zur Verfügung stehenden Zeit für die qualitative Evaluation mit den Schülern zusammen genutzt wurden. Zudem war die Entwicklung des Testinstruments noch nicht abgeschlossen. Die Daten können daher lediglich eine Tendenz andeuten.

Alle Vergleiche auf Grundlage der Rasch-Personenparameter zeigten statistisch bedeutsame Unterschiede ($p < 0,01$) zwischen *ebm@school*-trainierten und untrainierten Testabsolventen.

Im Folgenden werden die Personenparameter für Schüler mit und ohne *ebm@school* Training dargestellt (Tabelle 2).

Tabelle 2. Personenparameter für Schüler mit und ohne *ebm@school* Kurs

| | Personenparameter Mittelwert ± SD |
|---|--------------------------------------|
| Schüler mit <i>ebm@school</i> Kurs (<i>n</i>=37) | 597 ± 79 |
| Schüler ohne <i>ebm@school</i> Kurs (<i>n</i>=218) | 483 ± 94 |

Für die einzelnen Szenarien sind die Ergebnisse ebenfalls signifikant ($p < 0,001$).

Diese Ergebnisse weisen auf die Wirksamkeit der Schulungsmaßnahme hin und decken sich mit den unmittelbaren Beobachtungen der Dozenten. Sie sind jedoch aus designtechnischen Gründen in ihrer Aussagekraft limitiert. Um einen Kohorteneffekt ausschließen zu können, muss die Maßnahme im randomisiert-kontrollierten Design mit einer Vorher–Nachher Messung untersucht werden.

1.6 Aufnahme der Rahmenkriterien und Methoden in das Curriculum

Abschließend werden zusammenfassend die in den Pilotkursen identifizierten Rahmenkriterien dargestellt. Diese bildeten die Grundlage für die Überarbeitung des Curriculums.

1. Phasen der offenen Projektarbeit werden integriert
2. Das Modul Internet und Datenbankrecherche wird innerhalb des Curriculums nach vorn verlegt. Nicht vorhandene Ressourcen bzgl. Internetrecherche müssen durch Laptops mit W-LAN bereitgestellt werden.
3. der Umfang der einzelnen Module wird der Zielgruppe angepasst
4. der Schwierigkeitsgrad der Unterrichtsmaterialien wird der Zielgruppe angepasst
5. extrem leistungsschwache Schüler erhalten didaktisch reduziertes Material
6. Notenvergabe wird aufgenommen, da nicht benotete Projekte bei den Schülern teilweise nicht als Unterricht angesehen werden und damit geringer geschätzt werden.
7. Integration der Lehrer erfolgt ausschließlich in die offene Projektarbeit
8. Fehlzeiten werden dokumentiert, um Schüler mit hohen Fehlzeiten in den Analysen gesondert betrachten zu können.

2. Entwicklung eines Testinstruments zur Messung der Kompetenz *critical health literacy* (CHC Test)

Health literacy wurde 1974 erstmalig in den USA definiert als: “*The degree, to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions*” (17).

Die erweiterte Definition, die auch der Initiative der World Health Organization (WHO) *healthy schools* zugrunde liegt, beschreibt *health literacy* als: “*Health literacy represents the cognitive and social skills, which determine the motivation and ability of individuals to gain access to understand and use information in ways which promote and maintain good health*” (17). Diese Initiativen entstammen dennoch dem paternalistischen Paradigma. Sie zielen auf *compliance* und nicht auf informierte Entscheidungen.

Der in dieser Arbeit verwandte Begriff der *critical health literacy* erweitert die vorhandenen *health literacy* Ansätze um das Konzept der evidenzbasierten Medizin und zielt darauf ab, informierte Entscheidungen zu ermöglichen.

Bisherige Instrumente wurden vor allem für die Evaluation von Curricula für verschiedene Professionen im Gesundheitsbereich entwickelt. Diese sind weder für die Zielgruppe der Schüler geeignet, noch können sie zur Kompetenzmessung außerhalb von Trainings eingesetzt werden (18).

Die Entwicklung des CHC Test erfolgte in 3 Phasen.

Phase 1: Entwicklung einer ersten Version des Fragebogens und Pretests mit 8 Schülern.

Phase 2: Nach der Revision der 1. Version erfolgte der 1. Feldtest. Ziel des Feldtests war, erste Daten zum *model fit* des Rasch Modells und des Facettendesigns zu erheben. Zudem wurden anhand dieser Daten die Antwortformate, Distraktoren Itemschwierigkeiten und diskriminierenden Eigenschaften analysiert.

Phase 3: In einem 2. Feldtest wurde das Ergebnis der Revision überprüft. Zudem wurde erneut die Passung des Rasch Modells geprüft. Der zweite Feldtest zeigte die Rasch Skalierbarkeit aller Items. Damit können Personenparameter zur Messung der *critical health literacy* entweder mit dem kompletten Instrument oder aber einer verkürzten Version (einzelne Szenarien) ermittelt werden.

2.1 Testentwicklung Phase I: Entwicklung der Testhefte

Facettenstrukturen waren handlungsleitend im Prozess der Itemgenerierung. Die erste Facette bestand aus 4 Themenbereichen, die zuvor als Interessenbereiche in der Zielgruppe identifiziert worden waren: 1) Echinacin bei Erkältungen 2) Magnet Resonanz Tomographie bei Knieverletzungen; 3) Behandlung der Akne vulgaris; (Box 1) 4) Brustkrebsfrüherkennung (12). Die zweite Facette bestand aus 4 Kategorien, die Kompetenzbereiche unabhängig von der Modulstruktur des Curriculums umfassten: A) Medizinische Konzepte; B) Literatursuche; C) Statistik und D) Studiendesign und Stichprobe. Die sich aus den Facetten heraus ergebenden Zellen sollten mit mindestens je 2 Items gefüllt werden.

Der erste Itempool wurde von AS, CH und AF generiert und umfasste 250 Items. Dieser Pool wurde hinsichtlich Vollständigkeit, Doppelungen, Verständlichkeit und Lesbarkeit überprüft. Das Ergebnis war ein Itempool mit 72 Items. Die hohe *drop out* Rate ergab sich aufgrund von Doppelungen.

Die Items wurden in vier verschiedene Szenarien eingebunden. Box 2 zeigt exemplarisch das Szenario 3. Die Szenarien wurden nochmals einem Reviewing unterzogen (AS, CH, JK, JR). Überprüft wurden die Antwortformate, die Eindeutigkeit der Distraktoren und der Zeitaufwand, da die Anwendung aller Szenarien ein Maximum von 2 Unterrichtsstunden (90 Minuten) nicht überschreiten sollte.

Box 2: Deckblatt Testheft Szenario 3

Szenario 3

Case Nr.:

Wer kennt nicht den angstvoll erschreckten Blick in den Spiegel?



Einer Ihrer Freunde leidet allerdings unter hartnäckiger Akne, die auch nach verschiedensten Versuchen mit unterschiedlichsten herkömmlichen und kosmetischen Produkten nicht weichen will. Nun haben Sie von einer neuen Behandlungsmethode gehört, die viel Erfolg verspricht. Das neue Mittel heißt Tazarotene. Da Ihr Freund aber schon recht resigniert ist und Sie ihm keine weitere Enttäuschung zumuten möchten, wollen Sie sich erst einmal umfassend über diese Methode informieren.

Bitte beantworten Sie nun die folgenden Fragen. Bei den Multiple Choice Fragen kreuzen Sie die richtige(n) Antwort(en) bitte an. Es können jeweils eine oder mehrere Antworten richtig sein.

Sie möchten für Ihren Freund nach Informationen zu der oben genannten Behandlung seiner Akne suchen. Vor einer Recherche in medizinischen Datenbanken ist es sinnvoll, eine suchtaugliche Frage zu erstellen, die Ihr Interesse auf den Punkt bringt.

1. Tragen Sie hier bitte in jedes Feld des folgenden Schemas ein oder mehrere Suchwörter ein.

Patient/ Problem: _____

Behandlung: _____

Vergleichsbehandlung: _____

Ergebnisparameter: _____

2. Formulieren Sie nun anhand der von Ihnen in diesem Schema gesammelten Suchwörter eine suchtaugliche Fragestellung.

Für die Auswertung wurden Kodieranweisungen erstellt. Die Rater erhielten eine Schulung zur Anwendung dieser Kodieranweisungen. Diese wurden fortgeführt bis die Interrater-Reliabilität 95% - 100% erreicht hatte

Die Testhefte wurden in einem Vortest in der Zielgruppe (n=8) überprüft. Sechs Schüler wurden unter Anwendung der *think aloud* Methode aufgefordert, ihr Verständnis der einzelnen Items und ihre Antwortwahl zu jedem einzelnen Distraktor zu berichten. Die so erhobenen Daten waren Grundlage für die Revision. Des Weiteren wurden zwei Schüler gebeten, die Testhefte bezüglich der Einschätzung des Zeitaufwandes zu bearbeiten.

2.2 Testentwicklung Phase II: 1. Feldtest

2.2.1 Stichprobe und Setting

Im Feldtest erfolgte die empirische Überprüfung der Testhefte. Die Stichprobe umfasste 322 Schülerinnen und Schüler der 10. und 11. Klassen der Gymnasien, Gesamtschulen, Fachgymnasien, Erstsemester der Gesundheit und Studierende, die das Wissenschaftsliteraturseminar der Fachwissenschaft Gesundheit besucht hatten. Die soziodemographischen Daten sind in Tabelle 1 dargestellt.

Die Schüler / Studenten bearbeiteten innerhalb von 90 Minuten in variierender Reihenfolge vier Testhefte. Außer dem Taschenrechner waren keine weiteren Hilfsmittel erlaubt. Die Erhebung der Daten erfolgte anonymisiert.

2.2.2 Ergebnisse

Tabelle 3 zeigt die Basisdaten der Stichprobe.

Tabelle 3. Soziodemografische Daten (1. Feldtest)

| | Schüler | | Studenten |
|------------------------------|------------------------|--------------------------|--------------|
| | mit Training (n=37) | ohne Training (n=218) | (n=67) |
| Alter in Jahren* MW (SEM) | 17,61 (0,13) | 17,37 (0,07) | 27,16 (0,62) |
| Frauen | 20 (55,6%) | 138 (63,6%) | 46 (75,4%) |
| Muttersprache deutsch | 28 (77,8%) | 182 (84,3%) | 59 (76,7%) |

Die Auswertung prüfte die Rasch-Skalierbarkeit des gesamten Datensatzes. Diese Analyse vollzieht sich in mehreren prädefinierten Schritten. Im Detail wird geprüft, ob sich die Rasch-Skalierbarkeit für Gruppen von Fragen oder für Personengruppen darstellen lässt. Schließlich erbrachte eine Hybrid Analyse das „best fitting model“, da der Wert des Bayes Information Criteria (BIC) am niedrigsten war (Tabelle 4). Diese Lösung ergab sich in allen 4 Szenarien. Das Modell teilt die Stichprobe in 2 latente Klassen, von denen die leistungsstärkere Rasch-skalierbar ist.

Tabelle 4: 1. Feldtest: BIC Werte

| Analyse | Szenario 1 | Szenario 2 | Szenario 3 | Szenario 4 | Test gesamt |
|-----------------------------------|------------|------------|------------|------------|-------------|
| Rasch (1 Klasse) | 4959,56 | 6729,50 | 4178,55 | 7411,00 | 22779,37 |
| MiRa (2 Klassen) | 4915,55 | 6768,78 | 4201,51 | 7459,15 | 22725,65 |
| Hybrid Model 2 Klassen | 4875,04 | 6671,85 | 4147,00 | 7378,75 | 22486,80 |

MiRa = Mixed Rasch Modell

Die Reliabilität (WINMIRA ANOVA) des Tests insgesamt betrug 0.88 und für die einzelnen Szenarien 0.74 (Szenario 1), 0.68 (Szenario 2); 0.69 (Szenario 3) und 0.74 (Szenario 4).

Um die Effekte des Kompetenztrainings darzustellen, werden in Tabelle 5 die Rasch Personenparameter für die verschiedenen Gruppen dargestellt.

Tabelle 5: Personen Parameter der Schüler und Studenten

| | Personen Parameter* | ± SD |
|--|---------------------|---------|
| Untrainierte Schüler (n=218) | 451,75 | ± 65,29 |
| Trainierte Schüler (n=37) | 530,98 | ± 54,69 |
| Untrainierte Studenten (n=17) | 606,77 | ± 73,79 |
| Trainierte Studenten (n=50) | 651,16 | ± 65,57 |

Die Effektgröße (Cohens d) für die Schüler war 1,32 (95% CI 0,95-1,68, welche auf einen großen Effekt hinweist. Für Studenten betrug die Effektgröße (Cohens d) 0,64 (95% CI 0,08-1,2), ein mittlerer Effekt. Der Effekt des Vergleiches untrainierter Schüler mit trainierten Studenten war 3,05 (95% CI 2,65-3,45). Obwohl der Effekt der

Bildungseinrichtung (Schule oder Universität) höher erscheint als der des Trainings, wurde die Interaktion beider Faktoren nicht statistisch signifikant. Dieses Ergebnis stützt die Hypothese, dass das Kompetenztraining in dieser Zielgruppe wirksam ist.

2.2.3 Itemrevision

Das Ziel der Itemrevision war die Rasch Skalierbarkeit aller Testpersonen. Die Kriterien der Itemrevision wurden aus den Rasch Analysen abgeleitet:

- 1) Alle Item Z_q Werte wurden überprüft. Die Werte geben Informationen über die Modellpassung. Stark positive Werte indizieren niedrige Trennschärfe. Die Entscheidung, ein Item zu entfernen sollte dennoch unter der zusätzlichen Berücksichtigung des Inhaltes erfolgen.
- 2) Die Itemscores wurden auf die Lösungswahrscheinlichkeiten hin überprüft. Items mit einer Lösungswahrscheinlichkeit < 15% wurden als revisionsbedürftig angesehen. Angestrebt wurde eine Lösungswahrscheinlichkeit von 40% über alle Items berechnet, um Bodeneffekte zu eliminieren.
- 3) Distraktoren und Attraktoren wurden im Hinblick auf die Häufigkeit erfolgter Kreuzungen untersucht. Richtige Antworten, die selten gekreuzt waren, wurden auf Plausibilität überprüft. Distraktoren, die häufig gekreuzt waren und darüber falsche Antworten produzierten, wurden ebenso überprüft.

Die Revision erfolgte durch die Veränderung der Antwortformate, durch den Austausch von Distraktoren, die Veränderung der Kodieranweisung oder durch das Hinzufügen zusätzlicher Informationen. Zwei Items wurden entfernt, da sowohl die Lösungswahrscheinlichkeit als auch die Trennschärfe sehr niedrig waren. Insgesamt wurden 51 Items überarbeitet und 6 neue hinzugefügt.

Im Anschluss an die Itemrevision wurden die Kodieranweisungen überarbeitet.

2.3 Testentwicklung Phase III: 2. Feldtest

2.3.1 Stichprobe und Setting

Im 2. Feldtest erfolgte die abschließende Validierung der Testhefte nach der Revision. Die Stichprobe umfasste 107 Schülerinnen und Schüler der 11. Klassen einer Gesamtschule sowie Studierende der Fachwissenschaft Gesundheit, die das Wissenschaftsliteraturseminar bereits absolviert hatten.

Die Schüler / Studenten bearbeiteten analog zum 1. Feldtest innerhalb von 90 Minuten vier Testhefte, die in unterschiedlicher Reihenfolge zur Verfügung gestellt

wurden. Auch hier waren außer dem Taschenrechner keine weiteren Hilfsmittel erlaubt. Die Erhebung der Daten erfolgte ebenfalls anonymisiert.

2.3.2 Ergebnisse

Die soziodemographischen Daten sind in Tabelle 6 dargestellt.

Tabelle 6: Soziodemografische Daten (2. Feldtest)

| | Schüler (n=94) | Studenten (n=13) |
|------------------------------|----------------|------------------|
| Alter in Jahren | | |
| Mittelwerte(± SEM) | 16.5 (±0.072) | 30.62 (±1.412) |
| Frauen | 59 (62.8%) | 13 (100%) |
| Muttersprache deutsch | 84 (89.4%) | 13 (100%) |

Die Gruppe der Schüler im 2. Feldtest ist vergleichbar mit der Gruppe im 2. Feldtest. Der Altersunterschied zwischen den Studenten im 1. und 2. Feldtest ist auf 2 Studierende zurückzuführen, die 41 und 42 Jahre alt sind.

Zur Beurteilung der Objektivität der Dateneingabe wurde die Interrater Reliabilität für jeden einzelnen Distraktor (n=313) für 30 Datensätze berechnet, die von 2 Ratern unabhängig voneinander bewertet waren. Cohen`s Kappa war sehr gut für 277 Ratings, (k=0.9-1.0); moderat oder gut für 31 Ratings (k=0.7-0.89) und niedrig für 5 Ratings (k<0.7). Aus diesem Ergebnis resultierte die Ergänzung zusätzlicher Informationen in den Kodieranweisungen.

Mit den im 2. Feldtest erhobenen Daten sollte überprüft werden, ob die Revision des Instruments erfolgreich war, d.h., die intendierte Rasch Skalierbarkeit erreicht werden konnte. Die *model fit* Analysen zeigten, dass das Rasch Modell das *best fitting model* ist. Diese Lösung ergab sich in allen 4 Szenarien. Damit wurde das Ziel, eine Rasch-Skalierbarkeit für alle Leistungsniveaus herzustellen, erreicht.

Tabelle 7 zeigt die Bayes Information Criteria (BIC) der einzelnen Modelltests.

| Tabelle 7: 2. Feldtest: BIC Werte | | | | | |
|--|-------------------|-------------------|-------------------|-------------------|---------------------|
| Analyse | Scenario 1 | Scenario 2 | Scenario 3 | Scenario 4 | Overall test |
| Rasch (1 Klasse) | 2249,73 | 2451,22 | 1779,18 | 2662,31 | 8986,01 |
| MiRa (2 Klassen) | 2317,16 | 2556,69 | 1865,92 | 2742,71 | 9437,05 |
| Hybrid Model (2 Klassen) | 2272,97 | 2484,94 | 1837,40 | 2697,90 | 9162,78 |

NOTE: MiRa = Mixed Rasch Modell

Die Lösungswahrscheinlichkeit erreichte 0,34 verglichen mit 0,27 im 1. Feldtest. Die Lösungswahrscheinlichkeiten der einzelnen Szenarien veränderten sich wie folgt: Szenario 1: 0,27→0,38; Szenario 2: 0,27→0,34; Szenario 3: 0,27→0,30 und Szenario 4: 0,28→0,34.

Die Items des 1. Szenarios sind am leichtesten. Dieses zeigt sich daran, dass die mittleren Personen Parameter (Min.- Max) diese Szenarios am niedrigsten sind: 394,90 (199,10 – 675,71) verglichen mit den Personen Parametern des 2. Szenarios: 496,67 (133,07-763,04); des 3. Szenarios: 635,04 (372,73-973,34) und des 4. Szenarios: 473,39 (153,65-776,61). Das bedeutet, dass weniger Kompetenz erforderlich ist, um das 1. Szenario zu lösen. Daher kann mit dem leichtesten begonnen werden, wenn aufgrund mehrerer Messzeitpunkte jeweils einzelne Szenarien zur Kompetenzmessung eingesetzt werden. Um die verschiedenen Messungen miteinander vergleichen zu können, werden Korrekturterme ergänzt. Diese Terme ergeben sich aus dem *Weighted Likelihood Estimate*, einem Gewichtungsfaktor, der in der Normstichprobe für jedes Szenario in Relation zum Gesamttest ermittelt wurde (Tabelle 8).

Tabelle 8: WLE und Korrekturterme

| | WLE* | Korrekturterme |
|-------------------|--------|----------------|
| Scenario 1 | 605,10 | -105,10 |
| Scenario 2 | 503,33 | -3,33 |
| Scenario 3 | 364,96 | 135,04 |
| Scenario 4 | 526,61 | -26,61 |

*angegeben sind Mittelwerte

Die Personen Parameter werden getrennt für Schüler und Studenten dargestellt (Tabelle 9). Die Effektgröße (Cohens d) war 4,33 (95% CI 3,51-5,16). Studenten wurden als trainiert angesehen. Der Unterschied zwischen den Gruppen ist zurückzuführen auf die Unterschiede bzgl. Absolvierter Schuljahre und Training an der Universität.

Table 9: Personen Parameter des 2. Feldtests

| | Personen Parameter | ± SD |
|--|--------------------|---------|
| Trainierte Studenten (n=13) | 716,14 | ± 53,74 |
| Untrainierte Schüler (n=94) | 470,11 | ± 59,63 |

Der zweite Feldtest zeigte eine Verbesserung der Reliabilität (WINMIRA ANOVA) für das Instrument insgesamt: 0.91 und auch für die einzelnen Szenarien: 0.71 (Szenario 1); 0.78 (Szenario 2); 0.75 (Szenario 3) und 0.80 (Szenario 4). Diese Werte entsprechen im Verhältnis zur Reliabilität des Gesamttests der Messgenauigkeit, die unter Maßgabe der Spearman Brown –Formel aus der kleineren Itemanzahl der Teilttests erwartet werden kann.

2.4 Zur Handhabung des fertigen Testinstruments: *critical health literacy*

Der Stand der Validierung des Testinstruments entspricht der Zielsetzung und hat die folgenden Implikationen für seine Anwendung in Forschung und Praxis:

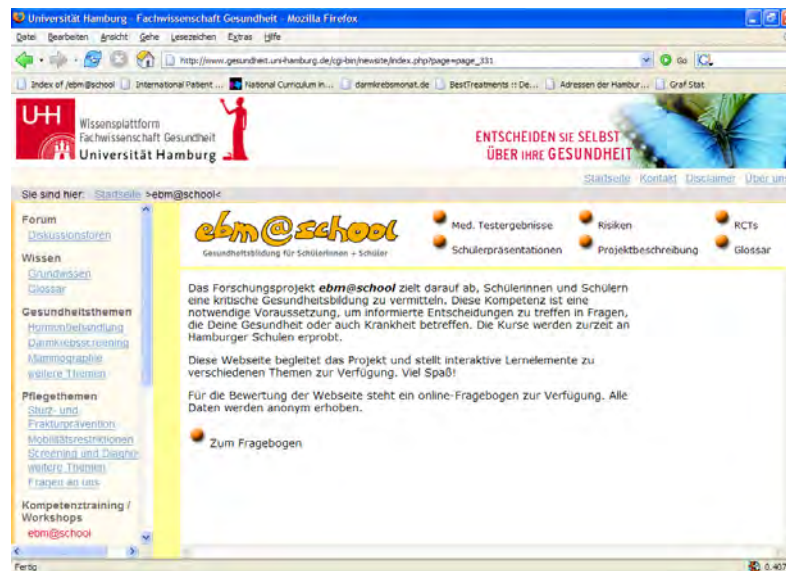
- Zur Bestimmung der *critical health literacy* einer Person genügt die Bearbeitung eines der vier Szenarien.
- Die Szenarien sollten unter Berücksichtigung ihrer unterschiedlichen Schwierigkeiten für die Testung ausgewählt werden. Z.B. sollten sie bei wiederholten Testungen im Rahmen eines Curriculums mit wachsenden Schwierigkeiten gestaffelt sein.
- Die Dateneingabe erfolgt eingabeobjektiv über die Kodieranweisung in die bereitgestellte SPSS-Eingabemaske (16).
- Die Testauswertung erfolgt mit der SPSS Syntax (16).
- Die Berechnung der Personen-Parameter erfolgt mit WINMIRA an dichotomisierten Datenfiles.
- Die Ermittlung des gültigen Rasch-Personen Parameters ergibt sich aus dem jeweiligen Personen-Parameter einer Person durch die Bearbeitung eines bestimmten Szenarios und des Korrekturterms.

3. Interaktive Lernelemente

3.1 Entwicklung

Im Projekt war die Entwicklung von interaktiven Lernelementen geplant, die inhaltlich eng am Curriculum orientiert sein sollten. Aus der Fülle an Möglichkeiten entschieden wir uns zunächst die Themen Risiko und diagnostischer Test interaktiv aufzuarbeiten. Die Internetseite *ebm@school* wurde in die bestehende Wissensplattform der Fachwissenschaft Gesundheit integriert. Zu den ausgewählten Themen sind dort auch die Bereiche: Informationen zum Projekt, Schülerpräsentationen (aus den Pilotkursen) und ein Glossar zu finden.

Screenshot der Einstiegsseite



3.2 Evaluation

Im Rahmen eines Pilotkurses hat eine erste Evaluation der interaktiven Lernelemente stattgefunden. In einem ersten Schritt sollten die Nutzung dieses neuen Mediums allgemein sowie die Handhabung der Lernelemente im Besonderen exploriert werden.

3.3 Ergebnisse

An der Evaluation haben 12 Schüler teilgenommen. Die Schüler schätzten ihre Computerkompetenz auf einer 5-Punkte Skala von sehr schlecht (0) bis sehr gut (4) im Mittel mit 2,83 ein. Neun von zwölf Schülern gaben an, gute oder sehr gute Computerkenntnisse zu haben. Die Frage, wie häufig Lernprogramme genutzt werden, beantworteten neun von zwölf mit nie oder selten, drei mit manchmal, und

kein Schüler gab an, Lernprogramme oft oder eher oft zu nutzen. Die Schüler, die Lernprogramme nutzen, tun dies im Besonderen in den Fächern Biologie (n=6), Chemie (n=6) und Mathematik (n=4).

Die Frage, ob die Schüler bei der Bearbeitung der Themen Risiko und Test Probleme hatten, beantworteten für den Bereich Risiko zehn Schüler mit nein und zwei mit ja. Für den Bereich Test gaben neun Schüler an, keine Probleme bemerkt zu haben, und zwei gaben an, dass Probleme auftraten. In einer ersten Befragung konnten keine gravierenden Probleme im Umgang mit den Lernbausteinen aufgezeigt werden. Diese erste Befragung erfüllt von ihrem Ansatz her die Ansprüche an einen Pretest. Insbesondere scheint das Niveau der Lernelemente didaktisch ausreichend reduziert zu sein, da diese Rückmeldungen von eher leistungsschwachen Schülern gekommen sind.

3.4 Ausblick

Im Rahmen eines studentischen Projektes wird zurzeit eine umfangreiche Evaluation der Lernelemente geplant und im Wintersemester durchgeführt werden. Anknüpfend an unsere Vorstudie werden in der Evaluation die Lerneffekte überprüft werden. Zudem wird in den nächsten Monaten das Thema RCT dazukommen. Im Zuge der geplanten randomisiert kontrollierten Studie zur Wirksamkeit des ebm@school Curriculums stehen diese Lernelemente den Schülern für die Wiederholung der Inhalte zur Verfügung.

4. Ausblick

Mit Abschluss des Pilotprojektes stehen sowohl ein Curriculum zur Vermittlung von *critical health literacy* zur Verfügung, als auch das CHC Instrument zur Kompetenzmessung. Die Ergebnisse der Pilotphase stützen die Hypothese, dass das Kompetenztraining wirksam ist.

Nunmehr könnte das Curriculum in die Praxis eingeführt werden. Diese sollte unter kontrollierten Bedingungen erfolgen und anhand des CHC Instrumentes evaluiert werden. Für die Durchführung der Implementierung unter kontrollierten Bedingungen liegt bereits ein Studienprotokoll vor. Gleichzeitig ist die vorhandene hohe Akzeptanz in den Schulen von enormer Bedeutung. Es stehen bereits Kooperationspartner zur Verfügung.

Literatur

1. Taking health literacy seriously (Editorial). *Lancet* 2005; 366:95.
2. Moynihan R, Heath I, Henry D. Selling sickness: the pharmaceutical industry and disease mongering. *BMJ* 2002; 324:886-891.
3. Verband der Angestellten-Krankenkassen e.V. (VdAK) (2004): MDS und VdAK/AEV warnen vor der kommerziellen Vermarktung von erfundenen Krankheiten, Versicherte müssen vor überzogenen Angeboten und Erwartungen geschützt werden, <http://www.vdak.de/pe/spik/presseallg206.htm>, Zugriff am 15.08.2005.
4. Gray NJ, Klein JD, Noyce PR et al. The internet: a window on adolescent health literacy. *J Adolesc Health* 2005; 37(3):243.e1-243.e7.
5. Hülfenhaus C, Steckelberg A, Mühlhauser I: Niederschlagswahrscheinlichkeit: Wie Hamburger Schülerinnen und Schüler Wahrscheinlichkeitsangaben verstehen. Poster. 7. Jahrestagung des DNEbM e.V. März 2006, Bochum.
6. The National Academies (2005): Board on Science Education, <http://www7.nationalacademies.org/bose/index.html>, Zugriff am 15.08.2005.
7. Nutbeam D. Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century. *Health Promot Int* 2000; 15:259-267.
8. European Network of Health Promoting Schools (2004): The European Network of Health Promoting Schools - the alliance of education and health, <http://www.euro.who.int/document/e62361.pdf>, Zugriff am 20.08.2005.
9. Lynagh M, Perkins J, Schofield M. An evidence-based approach to health promoting schools. *J Sch Health* 2002; 72:300-302.
10. Dickersin K, Braun L, Mead M, Millikan R, Wu AM, Pietenpol J, Troyan S, Anderson B, Visco F. Development and implementation of a science training course for breast cancer activists: Project LEAD (leadership, education and advocacy development). *Health Expect* 2001; 4:213-220.
11. Milne R, Oliver S. Evidence-based consumer health information: developing teaching in critical appraisal skills. *Int J Qual Health Care* 1996; 8:439-445.
12. Steckelberg A, Hülfenhaus C, Mühlhauser I: Ebm@school: Ein Curriculum zur Kompetenzentwicklung con critical health literacy. Welche Interessen haben Schülerinnen und Schüler Allgemeinbildender Schulen? Poster. 7. Jahrestagung des DNEbM e.V. März 2006, Bochum.
13. Freie und Hansestadt Hamburg, Behörde für Bildung und Sport <http://fhh.hamburg.de/stadt/Aktuell/behoerden/bildung-sport/service/statistik/susi/start-susi.html>. Accessed 16.03.2007.
14. Deutsches PISA-Konsortium (Hrsg.) PISA 2000. Basiskompetenzen von Schüler-innen und Schülern im internationalen Vergleich. Opladen, Leske und Budrich 2001.
15. Max Planck Institut für Bildungsforschung (2000): Third International Mathematics and Science Study – TIMSS. <http://www.timss.mpg.de/>, Zugriff am 10. 01. 2005
16. Rost, J.: Lehrbuch Testtheorie - Testkonstruktion. Bern, Huber 2004 (2. Aufl.).
17. Simonds SK. Health education today: issues and challenges. *J Sch Health*. 1977; 47(10):584-593.
18. European Network of Health Promoting Schools (2004): The European Network of Health Promoting Schools - the alliance of education and health, <http://www.euro.who.int/document/e62361.pdf>, Zugriff am 20.08.2005.

19. Shaneyfelt T, Baum KD, Bell D, et al. Instruments for evaluating education in evidence-based practice: a systematic review. *JAMA*. 2006; 6;296(9):1116-1127.

Anlagen zum Projektbericht:

1. 2 Reader
2. 2 Testhefte
3. 2 Kodieranweisungen
4. DVD: Reader Revision April 2007 PDF Dateien, QXP Dateien, Bilder 1, Bilder 2, Anlagen
5. DVD: Testinstrument Nov 2006
6. DVD: Interaktive Lernelemente