Wassner, Christoph; Martignon, Laura; Biehler, Rolf

Bayesianisches Denken in der Schule

Unterrichtswissenschaft 32 (2004) 1, S. 58-96

Quellenangabe/ Reference:
Wassner, Christoph; Martignon, Laura; Biehler, Rolf: Bayesianisches Denken in der Schule - In: Unterrichtswissenschaft 32 (2004) 1, S. 58-96 - URN: urn:nbn:de:0111-opus-58087 - DOI: 10.25656/01:5808

https://nbn-resolving.org/urn:nbn:de:0111-opus-58087
https://doi.org/10.25656/01:5808

in Kooperation mit / in cooperation with:

BELTZ JUVENTA

http://www.juventa.de

Unterrichtswissenschaft

Zeitschrift für Lernforschung
32. Jahrgang / 2004 / Heft 1

Thema
Stochastisches Denken

Verantwortliche Herausgeber
Jürgen Baumert, Gerd Gigerenzer, Laura Martignon

Christoph Wassner, Laura Martignon, Rolf Biehler

Bayesianisches Denken in der Schule

Bayesian Inference in School

Repräsentation von Information kann einfach, transparent und einleuchtend sein. Sie kann aber auch schwer verständlich werden, wenn sie zu symbolisch und kryptisch ist. Die moderne Mathematik leidet leider oft daran, dass Notationen und Begriffe so elaboriert sind, dass sie für den Nicht-Mathematiker unverständlich sind. Da, wie wir glauben, Wahrscheinlichkeitstheorie, wenn nicht die wichtigste, so sicherlich die nützlichste mathematische Disziplin in der Schule ist, sollten statt formaler Mathematik geeignete Repräsentationen unterrichtet werden, die das Verständnis erleichtern. Die stochastischen Problemstellungen, mit welchen die Schüler zum Lernen motiviert werden können, sollten ebenfalls möglichst nicht modellhaft und artifiziell sein, sondern der Realität entspringen. Diese Leitideen sind zentral für ein Unterrichtsprojekt, das in diesem Artikel näher beschrieben werden soll.

Information is communicated by means of representations. These representations can be simple and transparent but also cumbersome and obscure, when their aim is the economy of symbols and the consistency in the general framework of set theory. Modern Mathematics is often presented in terms of notations and concepts which are far too sophisticated for the layman. Since, as we believe, Probability Theory is the most useful branch of Mathematics when it comes to every day applications, it should be taught by means of transparent and natural representational devices, which foster understanding. The stochastic tasks which pupils solve at school also should be real life problems rather than abstract and artificial exercises. These are the two concerns that motivated the school intervention on Bayesian inference described in this article.

Eine der Hauptschwierigkeiten in der Entwicklung der mathematischen Wahrscheinlichkeitstheorie war die Definition des Begriffs „Wahrscheinlichkeit" (Hacking, 1984). Diese sollte einerseits präzise genug für ihre mathematische Behandlung sein, aber auch gleichzeitig die Nähe zur ursprünglichen Intuition behalten. Die Suche nach einer annehmbaren Definition dauerte fast drei Jahrhunderte und war von großen Kontroversen charakterisiert (Hacking, 1984; Daston, 1988). Ein Ende der Debatten wurde im zwanzigsten Jahrhundert gesetzt, als man Wahrscheinlichkeiten auf rein axiomatischer Basis behandelte. 1933 legte der russische Mathematiker Kolmogorov die Basis der axiomatischen Theorie nieder (Kolmogorov, 1933). Seitdem wurde sein Ansatz weiterentwickelt und die axiomatische Wahrscheinlichkeitstheorie ist heute ein Teil der Maßtheorie. In ihrer Anwendung behält die Wahrscheinlichkeitstheorie jedoch einen induktiven Charakter und wird wie eine empirische Wissenschaft betrieben. Diese hybride Natur der Wahrscheinlichkeitstheorie war der Grund dafür, dass sie erst spät als Teil der Mathematik akzeptiert wurde (Daston, 1988) und dass sie fern von der Schule blieb. Auffallend ist nämlich, dass, obwohl ihre Grundstrategien oft nur einfache arithmetische Operationen benötigen, sie nicht Teil von Lehrinhalten wurde. Dabei ist es heute eine Selbstverständlichkeit, dass die Grundzüge der Wahrscheinlichkeitstheorie für das Leben eines erwachsenen Bürgers wesentliche Instrumente sind. Erst Ende der siebziger Jahren des 20. Jahrhunderts wurden wahrscheinlichkeitstheoretische Begriffe zögerlich in die deutschen Lehrpläne eingeführt und Stochastik (also die Kombination von Wahrscheinlichkeitstheorie und Statistik) ist in den Schulcurricula der Gymnasien in allen Bundesländern präsent.

In diesem Artikel beschreiben wir verschiedene Arbeiten in einem von der DFG unterstützten Projekt[1], dessen Ziel es war, eine Erneuerung des Unterrichts über bedingte Wahrscheinlichkeiten und den Satz von Bayes vorzuschlagen und die Vorteile dieser Erneuerung empirisch nachzuweisen. Eine zentrale Frage im Projekt ist die der Repräsentation von Problemstellungen. Sie wird ausführlich aus kognitionspsychologischer Sicht erläutert. Es folgt eine didaktische Betrachtung des Themengebietes „Bedingte Wahrscheinlichkeit und Satz von Bayes" mit dem speziellen Fokus von üblichen Darstellungsformen. Danach werden die Resultate von Laboruntersuchungen mit Schülern der Sekundarstufe berichtet, bei denen es um das Lösen und Verstehen von Problemen mit bedingten Wahrscheinlichkeiten und Bayesianischem Schließen ging. Zuletzt beschreiben wir erste Versuche einer Implementation einer Unterrichtsreihe in diesem Themenbereich.

1. Repräsentation von probabilistischer Information - die kognitionspsychologische Perspektive

Während der Aufklärung sah man das wahrscheinlichkeitstheoretische Denken als eine Gegebenheit der menschlichen Natur, deren numerische Verwirklichung die formalisierte Wahrscheinlichkeitstheorie war (Daston, 1988). Laplace (1812) meinte, Wahrscheinlichkeitstheorie sei der Kalkül des „sense commun". Diese Vision wurde in Frage gestellt. Die Kognitionspsychologen des 20. Jahrhunderts wollten die These der Aufklärung

---

1 Das Projekt „Entscheidungsfindung unter Unsicherheit als fächerübergreifende Kompetenz - Alltagsorientierter Stochastikunterricht am Gymnasium" wurde von L. Martignon & P. Sedlmeier im Rahmen des BIQUA-Schwerpunktes beantragt und zunächst am MPI für Bildungsforschung in Berlin durchgeführt. (Förderung durch die DFG seit 1.1.2001, Geschäftszeichen: Ma 1544 / 10). Seit September 2002 wurde es gemeinsam mit R. Biehler an der Universität Kassel weitergeführt.

empirisch testen, dass Bayesianisches (also probabilistisches) Schließen und menschliches Denken übereinstimmen, und entwickelten Experimente, bei denen die Probanden derartige Aufgaben lösen mussten. Die empirischen Resultate schienen zunächst einer ganz anderen Wirklichkeit zu entsprechen. Tversky und Kahneman (1972) kamen, nach einer Reihe von empirischen Studien zu dem Schluss, dass der Mensch bei Bayesianischem Schließen versagt.

Eine der berühmtesten Aufgaben, die z.B. untersuchen sollte, ob Menschen probabilistisch denken können, ist die folgende (Eddy, 1982):

Die Wahrscheinlichkeit, dass eine symptomfreie Frau einer gewissen Altersgruppe Brustkrebs hat, beträgt 1%. Die Wahrscheinlichkeit, dass diese Krankheit mit einer Mammografie erkannt wird, ist laut Literatur 80%. Die Wahrscheinlichkeit, dass die Krankheit mit einer Mammografie irrtümlich diagnostiziert wird, obwohl sie gar nicht vorliegt, ist 9,6%. Wenn nun eine symptomfreie Frau dieser Altersgruppe bei einer Routineuntersuchung einen positiven Mammografiebefund erhält, wie groß ist dann die Wahrscheinlichkeit, dass bei ihr tatsächlich Brustkrebs vorliegt?

Bezeichnen wir das Ereignis „Brustkrebs" mit B und einen positiven Mammografiebefund mit M_+, so würde der Stochastiker die Aufgabe mit Hilfe des Satzes von Bayes lösen, mit der Anwendung einer einfachen Formel, wie folgt:

$$P(B \mid M_+) = \frac{P(B) \cdot P(M_+ \mid B)}{P(B) \cdot P(M_+ \mid B) + P(\overline{B}) \cdot P(M_+ \mid \overline{B})}$$

$$= \frac{0{,}01 \cdot 0{,}8}{0{,}01 \cdot 0{,}8 + 0{,}99 \cdot 0{,}096} \approx 7{,}8\%$$

Eddy berichtet, dass 95 von 100 Ärzten, denen er diese Aufgabe stellte, große Probleme hatten: Die Schätzwerte dieser 95 Ärzte lagen zwischen 70% und 80% statt bei 7,8%. Auch zeigten Untersuchungen an Personen, die die Bayessche Regel gelernt hatten, dass diese eher selten zur Problemlösung herangezogen wurde und andere Intuitionen mit dem formalen Wissen koexistieren. Die naheliegende Frage war: Woran liegt es, dass Menschen diese Aufgaben nicht lösen können? Hatten Tversky und Kahneman Recht, oder sollte man vielmehr die Ursache der Schwierigkeiten „woanders" suchen? Gigerenzer und Hoffrage (1995) sahen die Quelle der Fehleinschätzungen darin, dass die kommunizierte Unsicherheit in der Aufgabenstellung in Prozentwerten oder in Wahrscheinlichkeiten ausgedrückt war. Weiterhin schlugen sie vor, eine „natürliche" Repräsentation der Information zu verwenden. Sie zeigten empirisch, dass natürliche Häufigkeiten probabilistischem Schließen Transparenz verschaffen und Probanden leichter zu korrekten Antworten kommen.

Auf welches Repräsentationsformat für Unsicherheit ist das menschliche Denken natürlicherweise eingestellt? Diese Frage stellten sich Gigerenzer und Hoffrage (1995). Offensichtlich können Wahrscheinlichkeiten und Prozentwerte in der Umwelt nicht direkt wahrgenommen werden und somit bei „natürlichen", menschlichen Denkvorgängen kaum verarbeitet werden. Doch welche weiteren Möglichkeiten gibt es, um Unsicherheit darzustellen? Für stochastische Informationen sind das die folgenden numerischen Repräsentationsformate[2]:

Numerische Repräsentationsformate von Unsicherheit	*Beispiel*
Prozente	40%
Dezimalzahlen	0,4
Brüche	$\frac{4}{10}$
Häufigkeits - Verhältnisse	4 von 10
Chancen - Verhältnisse	4 : 6

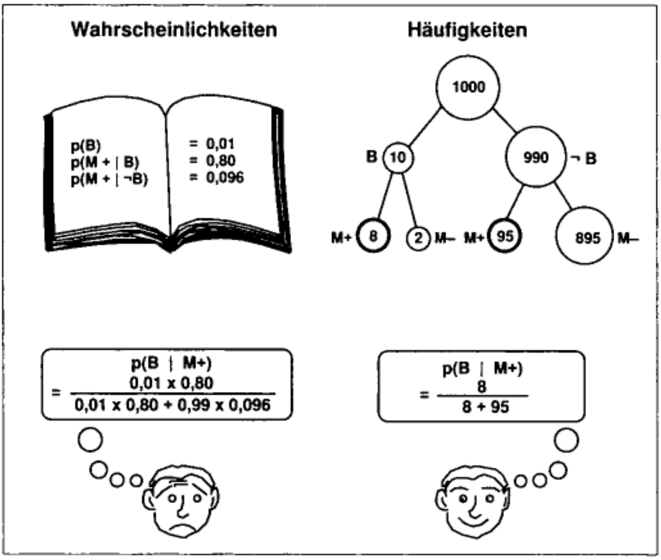
Die Frage war nun, welches dieser Repräsentationsformate am besten an die menschlichen Denkvorgänge angepasst ist. Nach psychologischen Theorien über Gedächtnis und Aufmerksamkeit gehören Häufigkeiten zu den wenigen Informationen, die automatisch registriert werden, d.h. ohne bewusste Intention und ohne Interferenz mit anderen kognitiven Prozessen (Hasher & Zacks, 1979). Demzufolge lag es nahe, zu versuchen, probabilistische Probleme in Häufigkeits-Verhältnissen zu formulieren. Das Brustkrebsproblem lautet dann (vgl. Gigerenzer & Hoffrage, 1995):

Von je 1000 Frauen einer gewissen Altersgruppe haben 10 Brustkrebs. Von diesen 10 Frauen, die Brustkrebs haben, erhalten 8 einen positiven Mammografiebefund. Von den 990 restlichen Frauen, die keinen Brustkrebs haben, erhalten dennoch 95 einen positiven Mammografiebefund. Stellen Sie sich eine Anzahl von Frauen dieser Altersgruppe vor, die einen positiven Mammografiebefund erhalten haben. Wieviele dieser Frauen sind tatsächlich an Brustkrebs erkrankt?

In empirischen Versuchen gaben nun wesentlich mehr der Ärzte (46%), die richtige Antwort: 8 von 103 (= 7,8%, vgl. Abb.1). Gigerenzer und Hoffrage nannten dieses Informationsformat *natürliche Häufigkeiten*, weil in einer natürlichen Umgebung durch das Zählen von Fällen („natural sampling") aus einer gegebenen Grundmenge, diese Art von Information gewonnen werden kann. Das Häufigkeitsformat macht dieses Bayesianische Problem auch für Menschen, die keine stochastische Ausbildung genossen haben, verständlich. Darüber hinaus ist die Veranschaulichung in einem „Häufigkeitsbaum" (siehe Abbildung 1, rechts) sehr intuitiv. Gegenübergestellt ist das übliche Wahrscheinlichkeitsformat ohne eine Repräsentation in Bäumen (links):

2 Nicht-numerische Repräsentationsformate von Unsicherheit sind zum Beispiel verbale Aussagen wie „sehr sicher" oder „eher unwahrscheinlich".

Abb. 1: Die Lösung des Brustkrebsbeispiels im Wahrscheinlichkeitsformat und im Häufigkeitsbaum (vgl. Gigerenzer & Hoffrage, 1995).



Im Gegensatz zum Wahrscheinlichkeitsformat kann man das Wesentliche der Aufgabe mit Hilfe des Häufigkeitsbaums einfach verstehen, denn man *sieht* geradezu: Obwohl die Krankheit bei relativ vielen der tatsächlich kranken Frauen mit der Mammografie auch erkannt wird (bei 8 von 10) und nur bei relativ wenigen der gesunden Frauen fälschlich diagnostiziert wird (bei 95 von 990), gibt es trotzdem viel mehr gesunde Frauen (95) als kranke (8) mit positivem Mammografiebefund. Das liegt ganz einfach daran, dass es *grundsätzlich* („a-priori") viel mehr Gesunde (990) als Kranke (10) gibt. Der Widerspruch zwischen Mathematik und Intuition wird also „repariert", wenn man die probabilistische Information in ein natürliches Informationsformat übersetzt.

Was kann man nun didaktisch durch eine Transformation von Wahrscheinlichkeitsformaten in Häufigkeitsformate gewinnen? Sollte man diesen „Trick" einer Übersetzung der Wahrscheinlichkeitinformationen im Aufgabentext in Häufigkeitsinformationen und Rückübersetzung der Häufigkeitslösung in eine Wahrscheinlichkeitsantwort als didaktisches Mittel einsetzen, um Schülern Bayesianisches Denken beizubringen?

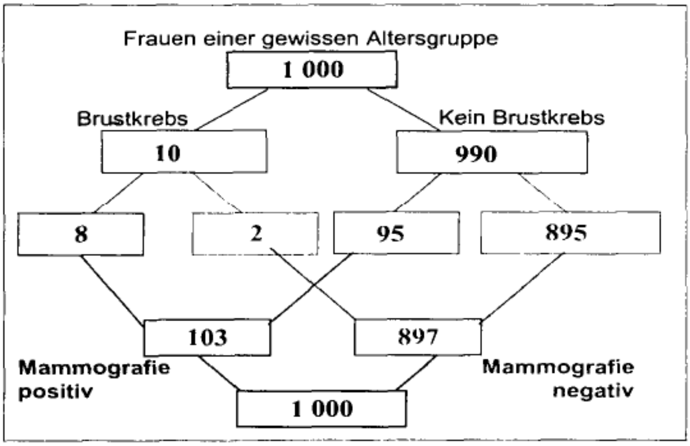
Der didaktische Vorschlag des Projektes war, auf jeden Fall mit den Schülern diesen „Umweg" der Übersetzung sorgfältig einzuüben. Aus empirischen Ergebnissen (z.B. Sedlmeier, 1999) ging bereits hervor, dass gewisse Wahrscheinlichkeitsprobleme mit Häufigkeiten leichter zu lösen sind und Lehrversuche deutlich effektiver werden.

Zur visuellen Unterstützung dieses Prozesses eignet sich hervorragend das erwähnte Baumdiagramm, das auch oft im Stochastikunterricht üblichen

Zugang zum Satz von Bayes zu finden ist. Die empfohlene grafische Darstellungsform, die auf dem Häufigkeitsbaum von Gigerenzer (z.B. 1993) basiert und kurz als „Doppelbaum" bezeichnet werden soll (vgl. Abb. 2), unterscheidet sich von den in Schulbüchern üblichen Baumdiagrammen mit sogenannten Pfadwahrscheinlichkeiten dadurch, dass alle Informationen in Häufigkeiten eingetragen sind und auch die „Umkehrung" des Baumes in der Darstellung integriert ist (vgl. Wassner, Krauss & Martignon 2002).

Man findet die Lösung bereits in der grafischen Darstellung: Man muss nur den Knoteninhalt für „Brustkrebs und Mammografie positiv" (8) durch den Inhalt des Knotens für „Mammografie positiv" (103) teilen. Im üblichen Baumdiagramm mit Wahrscheinlichkeiten gibt es keinerlei direkte Hinweis auf die Lösung. Nach Anwendung von verschiedenen Pfadregeln, die mathematischen Sätzen entsprechen, und Durchführung von Operationen in Wahrscheinlichkeitsformaten (also Rechnen mit Prozentwerten oder Dezimalzahlen, das bei Schülern erwiesenermaßen zu großen Schwierigkeiten führen kann) kann man schließlich die Lösung errechnen. Warum dieses Ergebnis allerdings die Lösung für das gegebene Problem darstellt, bleibt den meisten Schülern ziemlich rätselhaft, wie die Erfahrung zeigt. Dass auch der Einsatz von Baumdiagrammen, die mit Wahrscheinlichkeiten belegt sind, keine entscheidende Verbesserung des Lernerfolges gegenüber formaler Darstellung erbringen, wurde auch bei Sedlmeier (1999) empirisch nachgewiesen. Eine ausführliche Diskussion der Unterschiede und der Vorteile des „Doppelbaumes" gegenüber „üblichen" grafischen Darstellungen in diesem Bereich findet sich bei Wassner et al. (2002).

Abb. 2: Das Brustkrebsbeispiel in einem „Doppelbaum"



Wie im folgenden Kapitel deutlich wird, wurde bisher didaktisch kaum beachtet wurde, dass eben nicht nur die Repräsentation der algorithmisch-strukturellen Eigenschaften, sondern auch die *Repräsentation der gegebe-*

nen Information eine entscheidende Rolle bei der kognitiven Verarbeitung spielt. Das *Häufigkeitskonzept* berücksichtigt dies.

2. Tendenzen in der Schuldidaktik

Die Rolle der Bayes-Statistik für einen allgemeinbildenden Stochastikunterricht und damit auch das Bayes-Theorem findet in der deutschen fachdidaktischen Diskussion seit ca. 20 Jahren deutliche Beachtung. Riemer (1985) verweist in seinem Buch auf Verständnisschwierigkeiten der Lernenden in Stochastik. Er führt im Vorwort seines Buches aus: „Im Mittelpunkt stehen der subjektive Wahrscheinlichkeitsbegriff und das frühzeitige Abwägen zwischen verschiedenen Alternativhypothesen, nicht aber Mengenalgebra und Laplace-Wahrscheinlichkeiten" (S. 5). Die Bayessche Regel soll im Zentrum eines Unterrichts stehen, der an den Intuitionen der Lernenden anknüpft und anwendungsbezogene Ideen der Statistik in den Vordergrund stellt. Eine andere frühe Veröffentlichung ist die von Borovcnik (1984). Die Bayes-Statistik spielt auch in verschiedenen Monographien zum Stochastikunterricht eine wesentliche Rolle (Borovcnik 1992, Riemer 1991, Wickmann 1990). Sie wird nicht von allen Autoren als so zentral gesehen, beispielsweise spielt sie in den Büchern von Kütting (1981, 1994) und von Harten & Steinbring (1984), die sich mit der Sekundarstufe I - Didaktik beschäftigen, keine wesentliche Rolle.

In den Empfehlungen des AK Stochastik der Gesellschaft für Didaktik der Mathematik (AK Stochastik 2002) gehört der anwendungsbezogene Umgang mit der Bayesschen Regel zu den Mindestzielen der Sekundarstufe II, nicht aber zu denen der Sekundarstufe I.

Schaut man sich die Schulbücher der Sekundarstufe II an, so stellt man folgende Tendenz fest. In den früheren Büchern gehörte die Bayes-Regel zwar zum obligatorischen Stoff, wurde aber ohne großen Anwendungsbezug unterrichtet und war ein eher unbedeutendes Anhängsel bei der Behandlung der bedingten Wahrscheinlichkeit. Der Straffung des Curriulums zugunsten beurteilender Statistik im Sinne herkömmlichen Signifikanztestens fiel die Bayessche Regel manchmal ganz zum Opfer. Erst in neueren Lehrwerken zur Sekundarstufe II wird sie mit anwendungsbezogenen Aufgaben stärker betont.

Die Bemühungen um anwendungsbezogenen Mathematikunterricht mit authentischen Problemen war eine andere Quelle dafür, der Bayes-Formel wieder einen prominenten Platz in der Sekundarstufe II zu verschaffen. Die Arbeit von Böer (1997) aber auch die Arbeit von König (1991) ist in dieser Hinsicht exemplarisch.

2.1 Bayes-Formel in der Sekundarstufe I

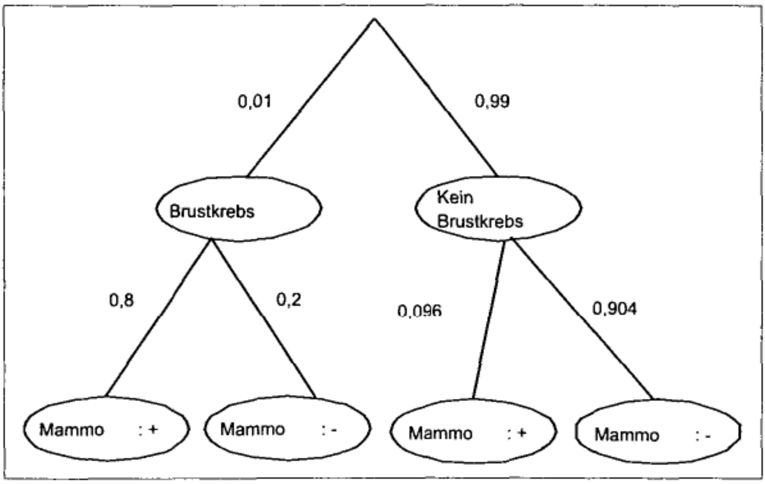
Die Lehrpläne der 16 Bundesländer, die i.d.R. differenziert nach den 3-4 Schulformen der Sekundarstufe I existieren sind besonders uneinheitlich hinsichtlich der Stochastik. Das hat auch eine Lehrplananalyse des AK Stochastik der GDM im Jahre 2000 ergeben. Das Thema Bayessche Regel kommt nur im Lehrplan Gymnasium in NRW vor.

Abb. 3: Themenkreis Bayessche Regel im Lehrplan (NRW 1993, S. 57)

> ***Alternative (a)***
>
> **Probleme aus dem Themenkreis der Bayesschen Regel werden mit Hilfe von Baumdiagrammen bearbeitet. Dabei soll die Bayessche Regel weder formal behandelt noch bewiesen werden (das ist der gymnasialen Oberstufe vorbehalten). Hier geht es um ein inhaltliches Verständnis der Zusammenhänge, wobei der Begriff der bedingten Wahrscheinlichkeit intuitiv verwendet wird. Für statistisches Denken wird die Regel relevant, wenn man studiert, wie sich die Wahrscheinlichkeiten verschiedener Alternativen auf Grund von Beobachtungen ändern.**

Man kann in NRW allerdings alternativ Bayes, Lottoprobleme oder Vertiefungen zum Gesetz der großen Zahl wählen. Die Lehrplanvorgaben haben sich in Lehrbücher nieder geschlagen, vor allem in Lambacher-Schweizer Bd. 9 (Schmid et al., 1996; Mitautor Riemer) aus dem Klett-Verlag und Cornelsen Mathematik Bd. 10 (Kuypers et al., 1999). Insofern sind in den nordrhein-westfälischen Gymnasien die optimalen Lehrplan- und Schulbuchvoraussetzungen gegeben, um das Thema zu unterrichten.

Abb. 4: Das Brustkrebsbeispiel in einem üblichen Baumdiagramm mit Wahrscheinlichkeiten



Um verständlich zu machen, welche vereinfachende Rolle die von uns favorisierten Repräsentationen haben, wollen wir kurz darstellen, wie Baumdiagramme und Häufigkeiten in einem typischen Lehrbuch der Sekundarstufe I verwendet werden.Voraussetzung ist, dass Schüler mehrstufige Zufallsexperimente modellieren können. Sie lernen in der 8. Klasse den Baum als Repräsentation mehrstufiger Zufallsexperimente kennen. Die Aufgabe, die oben vorgestellt wurde, könnten die Schüler mit ihren Mitteln wie folgt lösen. Zunächst muss erkannt werden, dass man die Situation als ein zweistufiges Zufallsexperiment modellieren kann. Beim zufälligen Ziehen einer Frau hat man in der ersten Stufe die möglichen Ergebnisse Brustkrebs/kein Brustkrebs, in der zweiten Stufe die Mammografie positiv/negativ.

Ausgerechnet werden können nun die Wahrscheinlichkeiten der 4 vorkommenden Pfade über die sog. Pfadmultiplikationsregel, die in der 8. Klasse behandelt wurde. Die Wahrscheinlichkeiten eines Pfades ergeben sich durch Multiplikation aller Wahrscheinlichkeiten längs eines Pfades. Die Wahrscheinlichkeit des linken Pfades ist dann z.B. $0{,}01 \cdot 0{,}8 = 0{,}008$.

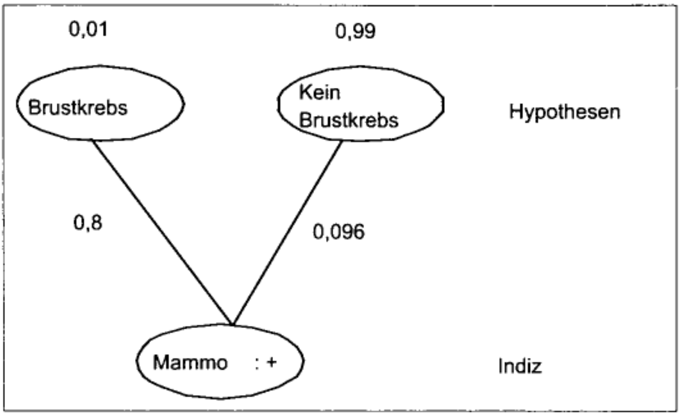
Für eine sachgerechte Interpretation müssen die Lernenden die Beschriftungen durch Wahrscheinlichkeiten Ereignissen zuordnen können. Die erste Wahrscheinlichkeit ist die für Brustkrebs, die zweite die für eine positive Mammografie, wenn Brustkrebs vorliegt und der Gesamtpfad steht für das Ergebnis „Brustkrebs und positive Mammografie". Die Formel $P(B \cap M_+) = P(B) \cdot P(M_+ \mid B)$, die „dahintersteckt" wird nicht explizit behandelt. Man belässt es bei umgangssprachlichen Formulierungen. Die Multiplikationspfadregel wurde durch Rückgriff auf natürliche Häufigkeiten begründet. Hätte man dies Beispiel zur Einführung verwendet, so hätte man im Schulbuch etwa das Argument gefunden: Stelle Dir vor, man wiederholt diesen Prozess 100000 mal, dann bekommen wir (ungefähr) 1%, also $100000 \cdot 0{,}01 = 1000$ Frauen mit Brustkrebs, davon (ungefähr) 80 % mit positiver Mammografie also $1000 \cdot 0{,}80 = 800$ Frauen mit Brustkrebs und positiver Mammographie, insgesamt also die Häufigkeit $h = 100000 \cdot 0{,}01 \cdot 0{,}80$. Die Wahrscheinlichkeit ist dann $h / 100000 = 0{,}01 \times 0{,}80$, also die Multiplikation der einzelnen Wahrscheinlichkeiten. An solchen Beispielen wird die allgemeine Regel motiviert, manchmal wird sie zusätzlich mit Variablen hergeleitet. Die Argumentation basiert also wesentlich auf den Regeln zur Bruchmultiplikation (ein Anteil von einem Anteil von ...).

Mithilfe der Häufigkeitsinterpretation wird eine Regel hergeleitet, die ab dann als Regel für Wahrscheinlichkeiten verwendet und eingeübt wird. Auf Häufigkeiten wird praktisch nicht mehr Bezug genommen. Man hat hier eine Elementarisierung erreicht, die Mengenalgebra vermeidet und den Schülern ein Modellierungsinstrument in die Hand gibt. Die Hürde, die aber beim Mammografie-Problem bei diesem Zugang entsteht, ist, wie man denn jetzt die Wahrscheinlichkeit ausrechnen könnte, dass unter der Bedingung einer positiven Mammografie auch Brustkrebs vorliegt.

Riemers Vorschlag, der im wesentlichen auch in Lambacher-Schweizer umgesetzt ist, wäre auf unser Beispiel angewendet der folgende. Man unterscheidet bei der Modellierung einerseits zwischen Hypothesen und Indizien, zum anderen wird ein verkürztes Baumdiagramm benutzt. Die Untersuchung soll helfen zwischen zwei alternativen Hypothesen abzuwägen:

Brustkrebs oder kein Brustkrebs. Als Indiz zählt ein positives Mammografieergebnis.

Abb. 5: Das Brustkrebsbeispiel im verkürzten Baumdiagramm mit Wahrscheinlichkeiten



Zum verkürzten Baumdiagramm wird wie folgt argumentiert. Die Pfade zum Indiz „Mammografie positiv“ haben die Wahrscheinlichkeiten:

Pfad über „Brustkrebs“ $0{,}01 \cdot 0{,}8 = 0{,}008$
Pfad über „kein Brustkrebs“ $0{,}99 \cdot 0{,}096 = 0{,}09504$

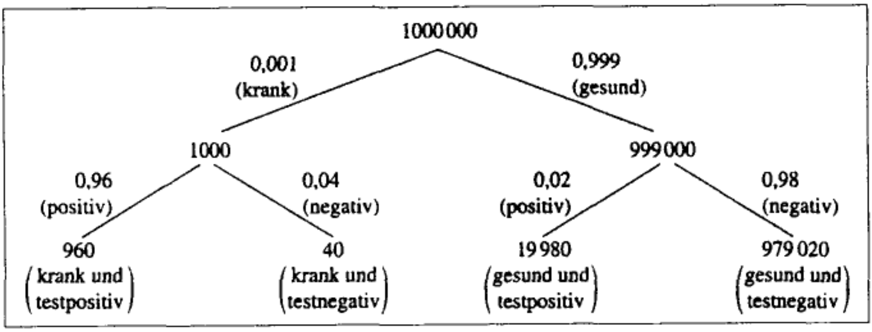
Die Wahrscheinlichkeit, das Indiz zu beobachten, ist also 0,008 + 0,09504. Der Anteil des Pfades über Brustkrebs ist

$$\frac{0{,}008}{0{,}008 + 0{,}09504} = 0{,}078.$$

Diese abstrakte Herleitung liefert den Lernenden keine Repräsentation dieser a posteriori - Wahrscheinlichkeit durch Verhältnisse von Häufigkeiten, auch keine Repräsentation durch ein vorstellbares Zufallsexperiment, in dem diese Wahrscheinlichkeiten als relative Häufigkeiten deutbar sind. Allerdings kann die Herleitung die Basis bilden für einen Kalkül, Aufgaben nach der Bayesschen Regel schematisch mit einfachen Mitteln lösen zu können, was auch eine wesentliche Intention des genannten Schulbuches ist.

Das Schulbuch von Cornelsen verwendet im Unterschied dazu für die Herleitung der Bayesschen Regel ein Baumdiagramm mit Häufigkeiten und Wahrscheinlichkeiten und auch fiktive Daten aus der medizinischen Diagnostik.

Abb. 6a: Ein Diagnosebeispiel in der Baumdarstellung (bei Kuypers et al., 1999, S. 175)



Anschaulich wird die bedingte Wahrscheinlichkeit mit Häufigkeiten berechnet, dann aber die Überlegung genutzt, um die Bayes-Regel formal zu begründen:

Abb. 6b: Formale Herleitung der Bayes-Formel (bei Kuypers et al., 1999, S. 178)

Wir haben diese Wahrscheinlichkeit oben dadurch berechnet, dass wir die Zahl der Fälle mit (H_1 und D) durch die Zahl **aller** Fälle mit D, also mit (H_1 und D) und mit (H_2 und D), dividiert haben:

$$P(H_1 | D) = \frac{H(H_1 \text{ und } D)}{H(H_1 \text{ und } D) + H(H_2 \text{ und } D)} = \frac{960}{960 + 19980}.$$

Stattdessen können wir natürlich auch die zugehörigen Wahrscheinlichkeiten durcheinander dividieren:

$$P(H_1 | D) = \frac{P(H_1 \text{ und } D)}{P(H_1 \text{ und } D) + P(H_2 \text{ und } D)}.$$

Mit den Gleichungen von Seite 177 ergibt sich daraus:

$$P(H_1 | D) = \frac{P(H_1) \cdot P(D | H_1)}{P(H_1) \cdot P(D | H_1) + P(H_2) \cdot P(D | H_2)}.$$

Fortan wird zum Modellieren und Berechnen von Bayes-Aufgaben nur die formale Darstellung in Verbindung mit einem Wahrscheinlichkeitsbaum benutzt, auf Häufigkeiten wird nicht mehr Bezug genommen. Beiden didaktischen Ansätzen ist gemeinsam, Häufigkeiten punktuell zu benutzen, um Regeln für Wahrscheinlichkeiten zu begründen, dann wird aber mit Wahrscheinlichkeitsbäumen und/oder formalen Darstellungen gearbeitet.

2.2 Überlegungen zu unserem Unterrichtsentwurf

Die Darstellungen machen deutlich, dass unser Vorschlag, konsequent mit natürlichen Häufigkeiten zu arbeiten, wesentlich über bisherige Ansätze hinausgeht. Wir haben die Ideen insofern aufgegriffen, als wir in der zweiten

Unterrichtsphase das am Beispiel erarbeitete Lösungs- und Darstellungsverfahren als allgemeine Methode reinterpretieren, um von Indizien auf Hypothesen zu schließen. Andere Kontexte als der medizinische sollten dann dazu dienen, diese Verallgemeinerung zu vollziehen.

3. Trainingsstudie zu individuellen Verarbeitungsprozessen bei unterschiedlichen Repräsentationsformen

3.1 Untersuchungsgegenstand

Im Projekt wurde zunächst eine Trainingslaborstudie mit Gymnasialschülern der Sek. I (10. Klasse; Stichprobe I) und Sek. II (Leistungskurs Mathematik; Stichprobe II) durchgeführt. Der Einfluss der Repräsentationsform auf den Lernerfolg bei Problemen mit bedingter Wahrscheinlichkeit und Bayesscher Inferenz[3] war Untersuchungsgegenstand. Es lagen bereits positive Befunde über derartige Unterschiede aus umfassenden Labortrainingsstudien von Sedlmeier (1999) bzw. Sedlmeier & Gigerenzer (2001) mit Studenten der Universität vor. Wir nahmen analog zu diesen Befunden an, dass auch bei Schülern der Sek. I und der Sek. II auf Häufigkeitsrepräsentation basierendes Training zu einem höheren Lernerfolg führt als Training basierend auf herkömmlicher Repräsentationsform mit Wahrscheinlichkeitsformaten. Eine weitere Modifikation neben der Stichprobe bei dieser Replikationsstudie lag in der Gestaltung der Erhebungsinstrumente.

Eine Erweiterung der Trainingsstudie (weitere Stichprobe: LK-Schüler und 11. Klasse) sollte die wichtige Frage untersuchen, inwieweit nach Training mit verschiedenen Repräsentationsformen die Schüler nicht nur kontextbezogen Probleme in diesem Bereich konnten, sondern fähig waren, sich vom Problemkontext zu lösen und entsprechende, abstrakt-mathematische Regeln herzuleiten und zu formulieren.[4]

Wir geben im Folgenden einen Überblick über das methodische Vorgehen und quantitative, sowie qualitative Ergebnisse (vgl. auch Wassner, Martignon & Sedlmeier, 2002).

3.2 Methodisches Vorgehen

Beschreibung der Trainingsvarianten

Es wurden computergestützte Tutorien der Trainingsstudien von Sedlmeier und Gigerenzer verwendet.[5] Die verglichenen Varianten basierten auf zwei unterschiedlichen Repräsentationsmethoden. Einerseits „Wahrscheinlich-

---

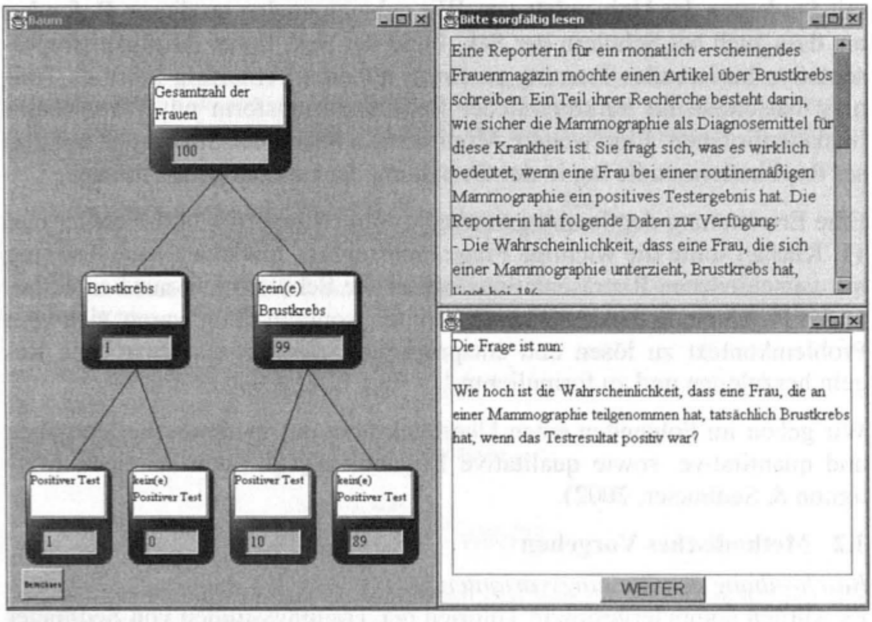
3 Im Folgenden sprechen wir kurz von „Bayes-Problemen" und meinen damit Probleme zur Anwendung des Satzes von Bayes und damit natürlich auch implizit zum Umgang mit bedingten Wahrscheinlichkeiten. Beide Begriffe sind inhaltlich eng miteinander verknüpft.

4 Diese Erweiterung der Fragestellung wird in diesem Artikel nicht dargestellt.

5 Das zugrundeliegende Computertutorium „BasicBayes" und theoretische Hintergründe werden in Sedlmeier (1997) detailliert beschrieben, Trainingsvarianten genauer in Sedlmeier & Gigerenzer (2001).

keitstraining“ (*PROB*) mit dem Inhalt, die nötigen Komponenten der konventionellen Bayes-Regel kennen zu lernen, diesen Komponenten analoge numerische Informationen in Aufgabentexten zuordnen und so konkrete Bayes-Problemstellung lösen zu können. Andererseits „Häufigkeitstraining“ (*FREQ*) mit dem Inhalt, die Lösung von Bayes-Problemen durch Transformation der numerischen Informationen in Häufigkeiten mit zusätzlicher visueller Darstellung in den gezeigten Häufigkeitsbäumen zu finden. Das Häufigkeitstraining basiert auf dem empirisch gestützten „Natural Frequency“-Konzept, das die Verbindung von kognitiven Algorithmen und Informationsformaten beachtet (Beitrag Gigerenzer in diesem Heft). In einer dritten Variante wurden beide Lösungsmethoden eingeführt und trainiert, jedoch ohne dass die zeitliche Dauer gegenüber den anderen Varianten erhöht war (*COMBI*). [6]

Abb 7: „Screenshot“ einer Aufgabe des computergestützten Häufigkeitstrainings von Sedlmeier (1997)



Während des Trainings (in allen Varianten) sehen die Versuchspersonen drei Fenster: Ein „Problemfenster“ enthält die Problemstellung mit den numerischen Informationen (in Prozentwerten), ein „Instruktionsfenster“ dient für Erklärung, Instruktion und Kommunikation mit dem Lernenden, ein

6 Diese Trainingsgruppe wurde nur bei Stichprobe I gebildet, die vor Training kein inhaltsbezogenes Wissen und demnach keine Repräsentationspräferenz besaß. Sie diente vor allem der Untersuchung von entwickelten Repräsentationspräferenzen nach Training.

„Repräsentationsfenster“ zeigt das entsprechende Darstellungsmittel und erlaubt Manipulationen an diesem durch den Lernenden.

Alle Trainingsvarianten beinhalten zwei Modi: In einem „Instruktionsmodus“ wird den Lernenden anhand zweier typischer Bayes-Probleme schrittweise erklärt, wie sie mit der jeweiligen Methode zur Lösung gelangen. Danach üben die Lernenden in einem „Praxismodus“ an analogen Aufgaben mit der erlernten Methode das Lösen der Problemstellungen. Wenn sie Schwierigkeiten bei einem gewissen Lösungsschritt haben, erhalten sie „Feedback“ bzw. Hilfestellungen durch die Software („step-by-step feedback“). In jedem Falle reicht die Hilfe aus, um zu gewährleisten, dass alle Versuchspersonen korrekte Lösungen erreichen und das Training vollständig durchlaufen. Das Training dauerte je nach persönlichem Arbeitstempo 1 bis 1,5 Stunden.

Stichproben

Die untersuchte Stichprobe I bestand aus Gymnasialschülern der 10. Klasse ($N = 80$, 35 männlich / 45 weiblich; Verteilung auf Trainingsgruppen: $n_{prob} = 27$, $n_{freq} = 27$, $n_{combi} = 26$), die vorher noch nie im Mathematikunterricht etwas über Stochastik erfahren hatten. Die Stichprobe II bestand aus Mathematik-Leistungskursschülern der 13. Klasse ($N = 47$, 31 männlich / 16 weiblich; Verteilung auf Trainingsgruppen: $n_{prob} = 23$, $n_{freq} = 24$), die das Themengebiet „Bedingte Wahrscheinlichkeit und Satz von Bayes“ unmittelbar vorher im Unterricht behandelt hatten. Die Schüler stammten aus verschiedenen Gymnasien aus den Ländern Berlin und Brandenburg und wurden den Gruppen zufällig zugewiesen. Kontrolliert wurde, dass alle Gruppen hinsichtlich geschlechtsbezogener und allgemeiner leistungsbezogener Merkmale (bisherige Schulleistung in Mathematik[7]) weitgehend parallel waren.

Erhebungsinstrumente

Leistungstest (BAYES)

Zur Erfassung von Grundkompetenzen im Zusammenhang mit Bayesschem Schließen und bedingter Wahrscheinlichkeit entwickelten wir einen Leistungstest. Die Testkonstruktion basierte auf einem Itempool, den wir auf Basis langjähriger Erfahrung mit diesem Themenbereich zusammenstellten. Während die Leistungstests bei Sedlmeier und Gigerenzer ausschließlich Aufgabenitems zu dichotomen Bayesschen Inferenzproblemen (vgl. Kapitel 1) enthielten, differenzierten wir mehrere Itemtypen mit unterschiedlichen Anforderungen (siehe unten). Die Kontexte der Aufgabenitems stammen weitgehend aus vorangegangenen Studien bzw. aus bekannten realen Anwendungsproblemen. Die Tests waren als „paper & pencil“ - Tests mit offenen Antwortformaten konzipiert waren, um Erkenntnisse über Vorgehensweisen und (Fehl-)konzepte der Schüler gewinnen zu können. Versuchsper-

7 Wir sind uns bewusst, dass die schulische Leistungsbewertung nicht als objektiv gelten kann und somit nur bedingt zur Frage der Vergleichbarkeit der Gruppen beiträgt.

sonen wurden während der Testphase mehrfach aufgefordert, alle Überlegungen bzw. Begründungen mit auf die Testbögen zu schreiben. Dadurch wurde eine detailliertere qualitative Untersuchung der Lösungs- und Verarbeitungsprozesse ermöglicht, die auch als Grundlage für die Bewertung diente.

Der Itempool enthält folgende Itemtypen (Angabe jeweils eines Beispielitems):

Itemtyp A0	Es soll angegeben werden, ob gewisse Informationen aus dem Aufgabentext für eine mathematische Lösung benutzt werden müssen (Format: multiple choice + freie Antwort). *Beispielitem (bezieht sich auf ein Item vom Typ A2):* Brauchen Sie für die Antwort eigentlich die Information A unbedingt? ❑ ja ❑ nein Begründen Sie ihre Entscheidung: ________
Itemtyp A0+	Es soll angegeben werden, ob und wie gewisse Informationen das Ergebnis beeinflussen können (Format: multiple choice + freie Antwort). *Beispielitem (bezieht sich auf ein Item vom Typ A2):* Ändert sich etwas, wenn die Prozentangabe in Information A kleiner wird (z.B. 5%)? ❑ nein ❑ ja, das Ergebnis wird größer ❑ ja, das Ergebnis wird kleiner Begründen Sie ihre Entscheidung: ________
Itemtyp A1 und A1+	Es soll angegeben werden, ob mit den Informationen im Aufgabentext ein kontextbezogenes Problem im Bereich „totale Wahrscheinlichkeit" gelöst werden kann (A1). Wenn mit „ja" geantwortet wird, soll es gelöst werden (A1+). (Format: multiple choice + numerische Antwort + offene Antwort). *Beispielitem (bezieht sich auf ein Item vom Typ A2):* Kann man mit den gegebenen Informationen beantworten, wie hoch die Wahrscheinlichkeit ist, dass ein in die Ambulanz kommender Patient an Fieber, Schüttelfrost und Hautabschürfungen leidet? ❑ ja, und zwar ____________ ❑ das kann man mit den obigen Informationen nicht eindeutig beantworten Begründen Sie ihre Entscheidung: ________
Itemtyp A2	Aus den gegebenen Informationen soll ein kontextbezogenes Bayes-Problem gelöst werden. Allein diese Anforderung war, wie oben beschrieben, expliziter Gegenstand der Trainings. (Format: numerische Antwort + offene Antwort) *Beispielitem:* Sie arbeiten in der Ambulanz eines Krankenhauses. Ein Patient mit hohem Fieber und Schüttelfrost kommt herein, bei dem Sie auch Hautabschürfungen bemerken. Folgende Informationen aus der Krankenstatistik haben Sie zur Verfügung: A Durchschnittlich 10% der in die Ambulanz kommenden Patienten haben Blutvergiftung. B Wenn ein Patient Blutvergiftung hat, besteht eine Wahrscheinlichkeit von 80%, dass er an diesen Symptomen (Fieber, Schüttelfrost, Hautabschürfungen) leidet.

	C Wenn ein Patient keine Blutvergiftung hat, besteht trotzdem eine Wahrscheinlichkeit von 10%, dass er an diesen Symptomen leidet. Die Frage lautet nun: Wie groß ist die Wahrscheinlichkeit, dass dieser Patient Blutvergiftung hat, wenn er diese Symptome aufweist? Antwort: ________ Geben Sie Ihre Überlegungen ausführlich an: ________
Itemtyp A3	Aus den gegebenen Informationen soll ein kontextbezogenes Bayes-Problem gelöst werden, das in der Struktur von der trainierten abweicht. Anstatt der üblichen dichotomen Struktur, kommen trichotome Ereignisstrukturen vor. (Format: numerische Antwort + offene Antwort) *Beispielitem:* Das Wetter in Nürnberg (Nordbayern) ist an 20% aller Tage regnerisch, an 50% wechselhaft und an 30% heiter. Es hat sich herausgestellt, dass am Vorabend eines Tages mit heiterem Wetter die Wettervorhersage für Nordbayern zu 80% gut und zu 20% schlecht lautet. Bei wechselhaftem Wetter am nächsten Tag lautete sie zu 50% schlecht und zu 50% gut. Ein Regenwettertag wird zu 60% mit schlecht angekündigt und zu 40% mit gut. Gestern abend wurde für heute gutes Wetter in Nordbayern vorhergesagt. Mit welcher Wahrscheinlichkeit ist heute heiteres Wetter in Nürnberg? Antwort: ________ Geben Sie Ihre Überlegungen ausführlich an: ________
Itemtyp A4	Aus den gegebenen Informationen soll ein kontextbezogenes Bayes-Problem gelöst werden, das im Unterschied zu Trainingsaufgaben einen komplexeren Kontext mit extrem seltenen bzw. sicheren Ereignissen enthält. (Format: numerische Antwort + offene Antwort) *Beispielitem:* Im Rahmen der Untersuchungen zu einem Mord im Raum Berlin wurde ein „DNA-Test" von potentiell Tatverdächtigen von der Polizei durchgeführt. Es wurde schließlich eine Übereinstimmung des DNA-Profils eines Untersuchten mit einer DNA-Spur am Tatort festgestellt. Der Verdächtige steht nun vor Gericht. Ein Sachverständiger erstellt ein Gutachten, das folgende Informationen enthält: *A* Ein im Rahmen des DNA-Tests untersuchter Mann ist ohne weitere Einschränkung durch Indizien zu 0,0001% der Spurverursacher. *B* Wenn der Verdächtigte wirklich der Spurverursacher ist, dann können wir davon ausgehen, daß es der durchgeführte DNA-Test mit Sicherheit (also zu 100%) anzeigt. *C* Die Wahrscheinlichkeit, daß das DNA-Profil eines Unschuldigen rein zufällig mit dem DNA-Profil der Spur übereinstimmt, beträgt 0,0001%. Sie wollen als Verteidiger nun wissen, mit welcher Wahrscheinlichkeit der Angeklagte der Verursacher der DNA-Spur ist, wenn der DNA-Test Übereinstimmung zeigte. Antwort: ________ Geben Sie Ihre Überlegungen ausführlich an: ________

Es ist zu beachten, dass allein Itemtyp A2, wie oben beschrieben, expliziter Gegenstand der Trainings war. Die Itemtypen A0, A0+ und A1/A1+ können als Teilanforderungen bei der Lösung von Bayes-Problemen charakterisiert werden. Sie basieren jeweils auf einem A2-Item und prüfen tiefergehendes Verständnis der Bayes-Problemsituation. Die Bearbeitung der Itemtypen A3 und A4 erfordert Transferleistungen bei der Modellierung. Die im Training gelernte und geübte Modellierung „einfacher“ dichotomer Bayes-Problem muss selbstständig vom Lernenden modifiziert bzw. weiterentwickelt werden.

Design
Alle Trainingsgruppen erhielten zunächst einen Vortest, um das Grundniveau in diesem Bereich festzustellen. Der Vortest setzte sich aus jeweils einem Item des Typs A0, A0+, A1/A1+ und A2 zusammen. Anschließend erfolgte das Training in der jeweiligen Trainingsvariante (siehe Versuchsplan, Tab.1). Der Lernerfolg wurde unmittelbar nach dem Training über einen Leistungstest (Nachtest 1) gemessen, der Items vom Typ A0 bis A4 enthielt.[8] Nach ca. 12 Wochen wurde der Leistungstest in derselben Weise wiederholt (Nachtest 2), um die Nachhaltigkeit des Lernerfolges zu prüfen.

Tab. 1: Versuchsplan, Laborstudie

Gruppe (Stichprobe)	*Vortest*	*Training*	*Nachtest 1*	*Nachtest 2*
Prob II (SP II)	unmittelbar vor Training	Wahrscheinlichkeits-version	unmittelbar nach Training	ca. 12 Wochen nach Training
Freq II (SP II)		Häufigkeitsversion		
Prob I (SP I)		Wahrscheinlichkeits-version		
Freq I (SP I)		Häufigkeitsversion		
Combi I (SP I)		„Kombi“version		

3.3 Ergebnisse

Analyse des Problemlöseprozesses bei Bayes-Problemen
Ein Erkennen von individuellen Verarbeitungsprozessen aus den rein numerischen Wahrscheinlichkeitslösungen der Probanden ist problematisch. Verschiedene Prozesse können zur selben Lösung führen (z.B. Scholz, 1988). Eine Analyse durch Sequenzierung der individuellen Problemlösungsprozesse war deshalb zunächst nötig. Diese Sequenzierung war Grundlage für die Operationalisierung der Testleistung über einen erreichten Punktsummenwert. Die Punktwerte pro Aufgabenitem ergaben sich also nicht nur anhand des normativen „richtigen“ oder „falschen“ Ergebnisses, sondern die einzelnen zur Lösung führenden Schritte wurden entsprechend klassifiziert, vorliegendes Verständnis individuell überprüft und nach festgelegten Krite-

8 Der Nachtest enthielt jeweils 3 Items vom Typ A2 (mit A0, A0+, A1, A1+), 1 Item vom Typ A3 und 1 Item vom Typ A4 (nur in Stichprobe II).

rien bewertet. Die erreichten Testleistungen spiegeln also prozessuales Verständnis der Schüler wider.

Bei den Problemstellungen im Bereich der Bayesschen Inferenz (Itemtypen A2 bis A4) wurden nach eingehender Analyse vier Hauptschritte des Problemlösungsprozesses klassifiziert:

a. *Identifizierung der Informationen aus dem Aufgabentext und Übersetzung in numerische Werte mit entsprechenden Bezeichnungen („labeln").*

Beispiel: Folgende Informationen sind verfügbar:

A. Wenn der Motor aussetzt, deutet in 90% der Flüge aufsteigender Rauch vorher darauf hin.

B. Während 10% der Flüge, bei denen der Motor nicht aussetzt, kommt Rauch aus dem Motor.

C. Der Motor seines Flugzeugtyps setzt in 1% aller Flüge aus.

Wie wahrscheinlich ist es, dass der Motor aussetzt, wenn Rauch von ihm aufsteigt?

Erwartetes Verständnis nach PROB-Training sinngemäß:

- C) P(Motor aus) = 0,01, A) P(Rauch | Motor aus) = 0,99, B) P(Rauch | Motor nicht aus) = 0,1

Erwartetes Verständnis nach FREQ-Training sinngemäß:

- C) gibt an, bei wie vielen Flügen von einer Gesamtzahl (z.B. 10 von 1000) im Durchschnitt der Motor aussetzt. A) gibt an, bei wievielen Flügen mit Motoraussetzer, Rauch aus dem Motor kommt (z.B. 90 von 100) ... usw.

b. *Verknüpfung der verfügbaren Informationen und Ableitung von Lösungselementen („strukturieren").*

Beispiel (wie bei a.): Folgende Informationen sind verfügbar:

A. Wenn der Motor aussetzt, deutet in 90% der Flüge aufsteigender Rauch vorher darauf hin.

B. Während 10% der Flüge, bei denen der Motor nicht aussetzt, kommt Rauch aus dem Motor.

C. Der Motor seines Flugzeugtyps setzt in 1% aller Flüge aus.

Wie wahrscheinlich ist es, dass der Motor aussetzt, wenn Rauch von ihm aufsteigt?

Erwartetes Verständnis nach PROB-Training sinngemäß:

- Nötig für die Lösung ist eine Verküpfung zwischen Information A und C bzw. B und (1-C). Das Produkt ist jeweils die Wahrscheinlichkeit für die Konjunktion, d.h. „Motor aus" *und* „Rauch" bzw. „Motor nicht aus" *und* „Rauch".

Erwartetes Verständnis nach FREQ-Training sinngemäß:

- Ausgehend von 1000 Flügen, geht der Motor im Durchschnitt bei 10 aus (C). Bei 9 dieser 10 Flüge kommt Rauch aus dem Motor (A) und bei 1 Flug nicht. Bei 990 Flügen (= 1000 - 10) geht der Motor nicht aus. Bei 99 dieser 990 Flüge kommt Rauch aus dem Motor (B) und

c. Aufstellen eines kompletten Lösungsalgorithmus („verarbeiten")

Im obigen Beispiel:

Erwartetes Verständnis nach PROB-Training:

- Aufstellen einer kompletten „Bayesformel", also ein Bruch, der im Zähler die Wahrscheinlichkeit für die Konjunktion von „Motor aus" und „Rauch" und im Nenner eine Summe aus der Wahrscheinlichkeit für die Konjunktion von „Motor aus" und „Rauch" und der Wahrscheinlichkeit für die Konjunktion „Motor nicht aus" und „Rauch" enthält. Im Allgemeinen sollte auch verstanden sein, dass es sich bei der Summe im Nenner um die Wahrscheinlichkeit für „Rauch kommt aus dem Motor" handelt. Die richtigen Werte sollen eingesetzt werden, evtl. auch ohne nochmals einen allgemeinen Zusammenhang zu formulieren.

Erwartetes Verständnis nach FREQ-Training:

- Aufstellen eines Bruches, der im Zähler die Anzahl der Flüge mit Motoraussetzer und Rauch enthält und im Nenner eine Summe aus der Anzahl der Flüge mit Motoraussetzer und Rauch und der Flüge ohne Motoraussetzer und Rauch. Im Allgemeinen sollte auch verstanden sein, dass es sich im Nenner um die Anzahl aller Flüge, bei denen Rauch aus dem Motor kommt, handelt. Die richtigen Werte sollen eingesetzt werden, evtl. auch ohne nochmals einen allgemeinen Zusammenhang zu formulieren.

d. Berechnung der Lösung („berechnen")

Erwartetes Verständnis nach PROB und FREQ -Training:

Berechnung des aufgestellten Lösungsbruches und Angabe eines numerischen Wertes für die gesuchte Wahrscheinlichkeit, evtl. Konvertierung des numerischen Formates.

Testleistungen

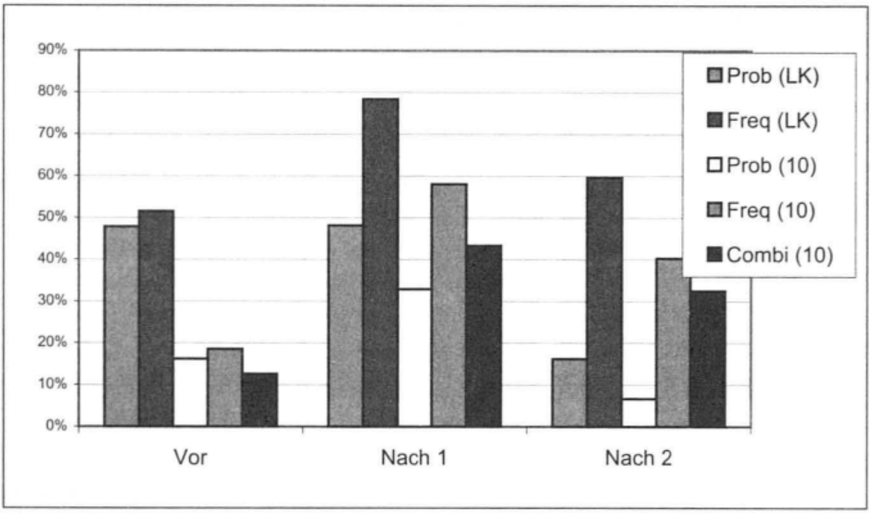
Zur quantitativen Analyse von Lernerfolgen verglichen wir die Unterschiede zwischen den Trainingsgruppen in den Testleistungen (erreichte Punktsummenwerte im Vortest, Nachtest 1, Nachtest 2). Abbildung 8a zeigt die durchschnittlich erreichten Anteile an den erreichbaren Punktsummen („Lösungsrate").

Im Vortest erreichten die *Zehntklässler* in den drei Versuchsgruppen Prob, Freq bzw. Combi 16,2%, 18,5% bzw 12,6% der erreichbaren Punktsumme.[9]

9 Eine Varianzanalyse ergibt keinen signifikanten Unterschiedseffekt ($F = 1{,}864$; $df = 2$; $p = 0{,}162$).

Im unmittelbaren Nachtest 1 ergaben sich Leistungen von 32,9%, 58,1% bzw. 43,3%.[10] Im späteren Nachtest 2 zeigten sich deutlich zunehmende Unterschiede zwischen Prob (6,6%) und Freq (40,3%) bzw. Combi (32,5%).[11] Nach Wahrscheinlichkeitstraining ist eine deutliche „Verfallskurve“ im Lernerfolg zu beobachten, während dieser Verfall in den Häufigkeits- und Combigruppen weniger stark auftritt.

Abb. 8a: Mittlere Testleistungen
(in % der jeweils erreichbaren Punktsummen)



Die *LK-Schüler* erreichten im Vortest in den beiden Versuchsgruppen Prob und Freq 47,8% bzw. 51,5% der erreichbaren Punktsumme.[12] Im unmittelbaren Nachtest 1 ergaben sich Leistungen von 48,1% bzw. 78,3%.[13] Im spä-

10 Eine Varianzanalyse mit der Berücksichtigung der unterschiedlichen „Vortestleistungen“ als Kovariate ergibt einen hochsignifikanten Unterschiedseffekt ($F = 5{,}318$; $df = 2$; $p = 0{,}007$; Effektgrößenschätzer partielles $\eta^2 = 0{,}127$). Bei paarweisen Vergleichen erweist sich der Unterschied zwischen Freq und Prob als hochsignifikant (Mittl. Diff. = 4,596; $p = 0{,}006$), die Unterschiede zwischen Freq und Combi bzw. Prob und Combi erweisen sich als nicht signifikant.

11 Eine Varianzanalyse mit der Berücksichtigung der „Vortestleistungen“ als Kovariate ergibt einen hochsignifikanten Unterschiedseffekt ($F = 17{,}062$; $df = 2$; $p < 0{,}001$; $\eta^2 = 0{,}319$). Bei paarweisen Vergleichen erweist sich der Unterschied zwischen Freq und Prob als hochsignifikant (Mittl.Diff. = 6,264; $p < 0{,}001$) sowie der Unterschied zwischen Combi und Prob (Mittl.Diff. = 4,506; $p < 0{,}001$). Der Unterschied zwischen Combi und Freq ist nicht signifikant (Mittl. Diff = 1,757; $p = 0{,}342$).

12 Eine Varianzanalyse ergibt keinen signifikanten Unterschiedseffekt ($F = 0{,}548$; $df = 1$; $p = 0{,}463$).

13 Eine Varianzanalyse mit der Berücksichtigung der unterschiedlichen „Vortestleistungen“ als Kovariate ergibt einen hochsignifikanten Unterschiedseffekt (Mittl.Diff = 4,112; $F = 22{,}608$; $df = 1$; $p < 0{,}001$; Effektgrößenschätzer partielles $\eta^2 = 0{,}339$).

teren Nachtest 2 verringert sich der Lernerfolg bei Prob (16,2%) deutlich stärker als bei Freq (59,7%)[14].

Abbildung 8b zeigt die Testleistungen bei den trainierten, dichotomen Bayes-Problemen (Itemtyp A2). Die Zehntklässlern zeigten hier kaum bewertbares Vorwissen, während die LK-Schüler bei etwa 40% liegen. Unterschiede in den Leistungen in Nachtest 1 und der Stabilität der Lernerfolge sind etwa mit dem obigen Gesamttestergebnis vergleichbar.

Abb. 8b: Mittlere Testleistungen (in % der jeweils erreichbaren Punktsummen), nur Itemtyp A2

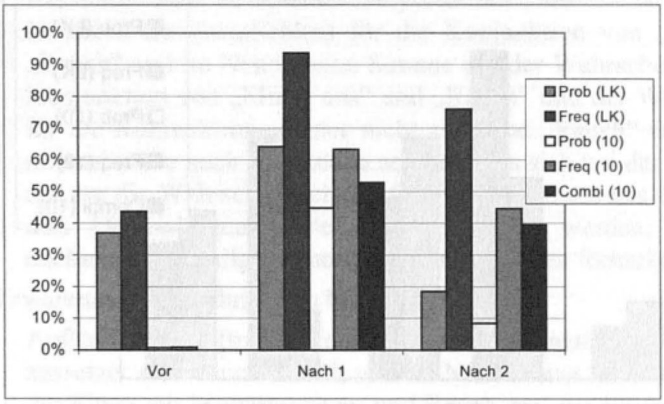
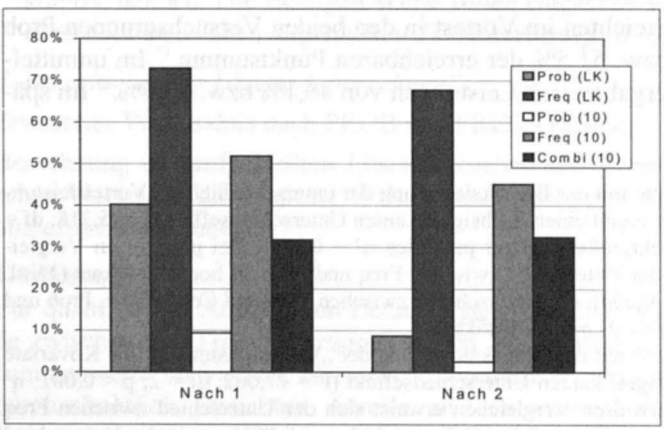


Abb. 8c: Mittlere Testleistungen (in % der jeweils erreichbaren Punktsummen), nur Itemtyp A3



Transferprobleme des Itemtyps A3, die nicht im Vortest enthalten waren, wurden in den Nachtests unterschiedlich gut bearbeitet. Es ergab sich ein

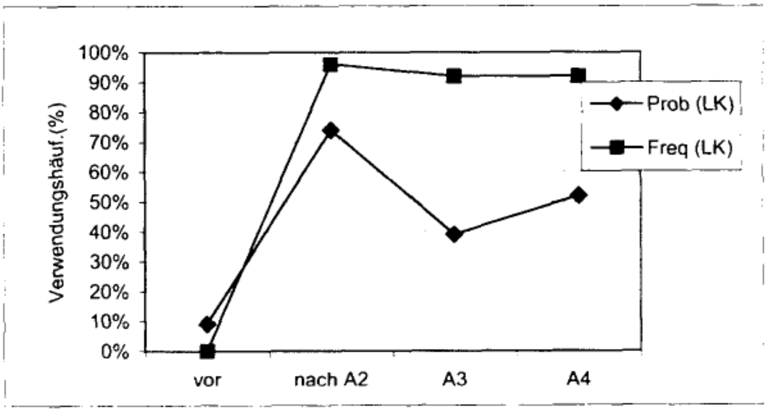
14 Eine Varianzanalyse mit der Berücksichtigung der „Vortestleistungen" als Kovariate ergibt einen hochsignifikanten Unterschiedseffekt (Mittl. Diff = 6,294; F = 121,917; df = 1; p < 0,001; $\eta^2 = 0,735$).

ähnliches Bild in den Unterschieden der Testleistungen wie im Gesamttest (Abbildung 8c). Die Wahrscheinlichkeitsgruppen schnitten im Vergleich sogar noch schlechter als bei „A2-Items" ab.

Kognitive Zugänglichkeit der Darstellungsformen
Die „paper & pencil"-Tests mit offenen Antwortformaten erlaubten die Analyse der tatsächlichen Verwendung von Darstellungs- bzw. Lösungsmethoden. Die Verwendungsraten können als Maß für die „kognitive Zugänglichkeit" einer Repräsentation interpretiert werden, d.h. die Bereitschaft ein Repräsentationsmodell bzw. -methode als adäquat für den Problemlösungsprozess zu akzeptieren und sie deshalb tatsächlich zu benutzen (vgl. Bea, 1995, S. 161). Inwieweit die Akzeptanz tatsächlich als genereller Indikator für die kognitive Eignung einer Darstellungsmethode gelten kann, ist aber nicht eindeutig geklärt. Diese Frage muss hier ausgeklammert werden.

Der Beschreibung von „kognitiver Zugänglichkeit" über Akzeptanz folgend, überprüften wir, *wie oft* die Schüler im Nachtest 1 exakt die trainierte Methode verwendeten. Wir fanden (vgl. Abb. 9a), dass nach dem Wahrscheinlichkeitstraining 74% der LK-Schüler bei den dichotomen Bayes-Problemen (A2) die gelernte Methode benutzten und sogar nur 39% bei A3-Problemen bzw. 52% bei A4-Problemen. Demgegenüber benutzten nach Häufigkeitstraining bei A2-Problemen ca. 96% und sonst 92% die gelernte Methode.

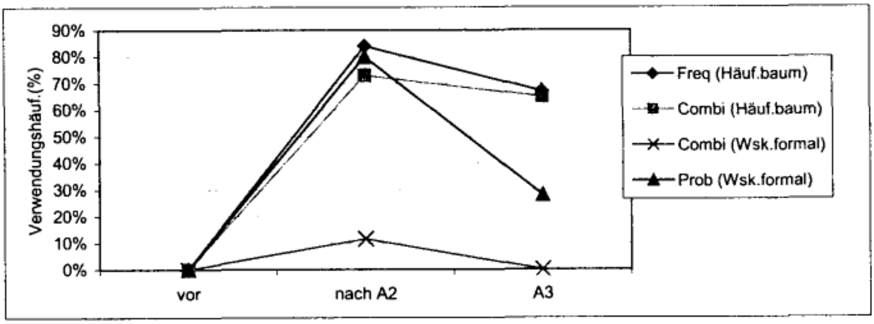
Abb. 9a: Verwendungshäufigkeiten der jeweiligen trainierten Darstellungs- bzw. Lösungsmethode (Stichprobe II; vor Training bzw. nach Training in Anforderungsstufen A2, A3, A4)



Bei den Schülern der 10. Klasse ohne Vorwissen aus dem Unterricht (Stichprobe I) waren im Vortest keine tragfähigen Lösungs- oder Darbietungsmethoden[15] zu erkennen (vgl. Abb. 9b).

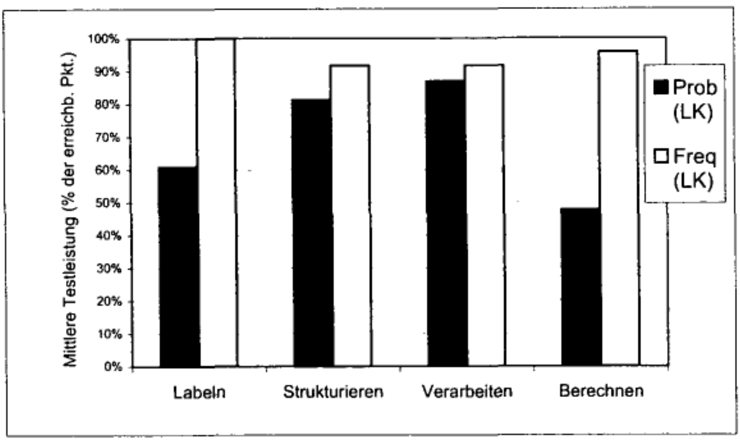
15 Ein von Null verschiedenes Grundniveau von Vortestleistungen (siehe oben) ergibt sich aus basalen, intuitiven Vorstellungen. Tatsächliche methodische Lösungen oder Darstellungen von Bayes-Problemen waren nicht erkennbar.

Abb. 9b: Verwendungshäufigkeiten verschiedener Darstellungs- bzw. Lösungsmethoden (Stichprobe I, 10. Klasse; vor bzw. nach Training in Anforderungsstufen A2, A3)



Nach FREQ-Training verwendeten die Schüler im unmittelbaren Nachtest zum größten Teil den Häufigkeitsbaum zur Lösung (für A2-Aufgaben 84%, für A3-Aufgaben 67%)[16]. Nach COMBI-Training zeigte sich eine ähnlich hohe Verwendung von Lösungsmethoden mit Häufigkeitsformaten und Baumdarstellungen (A2: 73%, A3: 65%) und eine nur sehr geringe Verwendung von Lösungsmethoden mit Wahrscheinlichkeitsformaten mit formaler Darstellung (A2: 12%, A3: 0%). Nach PROB-Training verwendeten die Schüler bei A2-Aufgaben zu 80% die gelernte formale Wahrscheinlichkeitsmethode, jedoch bei A3-Aufgaben nur noch zu 28%.

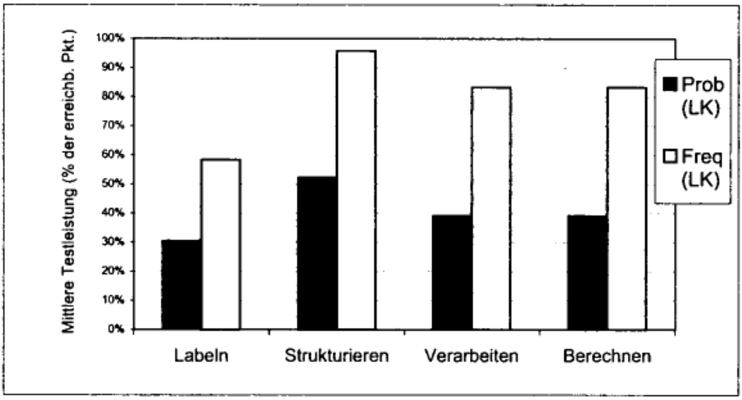
Abb. 10a: Mittlere Testleistungen (in % der jeweils erreichbaren Punktsumme), differenziert nach Lösungsprozessschritten, bei dichotomen Bayes-Problemen (A2)



16 Der Rest verwendete jeweils eine andere oder keine tragfähige Methode bzw. Darstellung.

Differenzierung nach Lösungsprozessschritten
Eine Analyse der schriftlichen Überlegungen der *LK-Schüler* (Stichprobe II) im Nachtest 1 ergab folgende Unterschiede zwischen den Trainingsgruppen in den erbrachten (mittleren) Leistungen in den vier Lösungsprozessschritten (siehe oben). Bei A2-Problemen zeigen sich die Unterschiede am deutlichsten in den Schritten „Labeln" und „Berechnen". Bei A3-Problemen ergeben sich deutliche Unterschiede in allen Prozessschritten (vgl. Abb. 10a, b).

Abb. 10b: Mittlere Testleistungen (in % der jeweils erreichbaren Punktsumme), differenziert nach Lösungsprozessschritten, bei trichotomen Bayes-Problemen (A3)



3.4 Diskussion und Implikationen
Die Ergebnisse dieser Studie und früherer Studien von Sedlmeier und Gigerenzer sind Belege dafür, dass die Darbietungsform entscheidenden Einfluss auf individuelle Verarbeitungsprozesse und damit auf den Lernerfolg hat. Die Hypothese, dass auf Häufigkeitsdarstellung basierendes Training gegenüber üblichem, formalen Wahrscheinlichkeitstraining zu höherem Lernerfolg führt, wurde für verschiedene Schülergruppen der Sek. I (10. Klasse, ohne unterrichtlichem Vorwissen) und der Sek. II (LK Mathematik, mit unterrichtlichem Vorwissen) bestätigt. Die Schüler erreichten bereits nach einem nur etwa 1½ - stündigen Training mit spezieller Häufigkeitsdarstellung nicht nur bei trainierten dichotomen Bayes-Problemen, sondern auch bei Transferproblemen deutlich höhere Leistungen als nach vergleichbarem Wahrscheinlichkeitstraining. Dieser Lernerfolg blieb im Vergleich zu den Wahrscheinlichkeitsgruppen auch längerfristig deutlich stabiler. Die Befunde stützen die kognitionspsychologische Theorie, dass „natürliche Häufigkeiten" als evolutorisch ursprüngliche Form für die Mitteilung von Wahrscheinlichkeitsinformationen besonders adaptiv sind („adaptive algorithms") und somit eine Häufigkeitsdarstellung gewisser stochastischer Probleme das Verständnis erleichtern kann.

Die qualitative Analyse der Akzeptanz der Darstellungsformen ergab weitere Evidenz für die kognitiven Vorteile der Häufigkeitsmethode. Die Schüler der Häufigkeitsgruppe akzeptierten die gelernte Darstellungsform in hohem Maße als angemessen für die Problemlösung, während die Schüler der Wahrscheinlichkeitsgruppe die trainierte Methode deutlich weniger oft zur Lösung verwendeten. Die Schüler der Combi-Gruppe, die sich hinsichtlich der Methode entscheiden mussten, präferierten eindeutig die Verwendung der Häufigkeitsmethode. Diese Befunde interpretieren wir als weitere Indikatoren für ihre höhere „kognitive Zugänglichkeit".

Bei der Analyse des individuellen Lösungsprozesses von Bayes-Problemen wurden vier sequentielle Lösungsschritte klassifiziert. Ein Vergleich der differentiellen Leistungen bei den einzelnen Lösungsprozessschritten zeigte Defizite der Wahrscheinlichkeitsgruppe in allen Lösungsschritten, vor allem aber in den Prozessschritten „Labeln" und „Berechnen". Der Schritt „Labeln" bezieht sich auf die Informationsklassifizierung, d.h. die Bedeutung der Informationen im Aufgabentext und ihre Notwendigkeit für die mathematische Modellierung. Der Schritt „Berechnen" beinhaltet die technischen Fertigkeiten, die zur Umformung in eine Wahrscheinlichkeitsaussage nötig sind (z.B. Bruch-, Prozentrechnung). Diese Befunde deuten darauf hin, dass - neben technischen Grundfertigkeiten - die jeweilige kontextbezogene Modellierung mit der Wahrscheinlichkeitsmethode die größten Schwierigkeiten verursacht. Für eine tiefgreifendere Analyse dieser Lösungsprozesse und Deskription und Klassifizierung der Lösungsstrategien verweisen wir auf gerade laufende Arbeiten (Wassner, Dissertationsarbeit).

Aus den deutlichen Befunden der Studie kann für den gewinnbringenden Stochastikunterricht gefolgert werden, dass für die Schüler Wahrscheinlichkeitsprobleme durch eine Übersetzung in Häufigkeiten leichter verständlich werden und dadurch nachhaltigere Lernerfolge möglich sind (vgl. Sedlmeier, 1999; Wassner et al., 2002). Ein veränderter Zugang zu stochastischen Problembereichen, wie etwa Problemen mit bedingten Wahrscheinlichkeiten und Bayesscher Inferenz, mit den gezeigten Häufigkeitsdarstellungen verspricht eine Verbesserung der Effektivität des Unterrichts hinsichtlich *stochastischen Denkens* der Schüler.

4. Der Schritt in den Unterricht - eine explorative Unterrichtsstudie im Themengebiet „Anwenden der Bayesschen Regel"

4.1 Rahmenbedingungen und Untersuchungsgegenstand

Die Vortestleistungen der untersuchten LK-Schüler sowie Aussagen von befragten Lehrenden bestätigten unsere Überzeugung, dass bisher üblicher Stochastikunterricht z.B. im Themenbereich bedingte Wahrscheinlichkeit nicht zu optimalen Lernerfolgen führt, nicht einmal in einem Leistungskurs Mathematik. Shaughnessy stellte bereits 1992 fest (S.475): „To my knowledge, there has not been any research reported that has attempted to change students´ misconceptions or beliefs specifically on conditional tasks. The literature is long on excellent didactical suggestions, but short on hard research in the area". An diesem Mangel an systematischen Unterrichtsuntersuchungen hat sich in Deutschland bis heute nichts geändert, und das trifft durchaus auf alle Themengebiete der Stochastik zu. Es ist dringend Handlungsbedarf geboten.

Wir entwickelten im Projekt ab August 2002 zum Zweck einer Evaluation in der Sekundarstufe I eine Unterrichtsreihe[17] in Kooperation mit dem Mathematiklehrer Stefan Schweynoch, der auch die Realisierung im Unterricht koordinierte. Die Reihe basiert auf den bisher dargestellten Grundideen und enthält Veränderungen zum „üblichen" Unterricht im Theorieaufbau, bei den Darstellungsformen, der Realitätsnähe der Themen und in Unterrichtsmethoden. Die Frage nach geeigneten Problemstellungen und ihrer Bearbeitung und Erschließung im Unterricht war eine Kernfrage bei der Gestaltung.

Der Untersuchungsgegenstand dieser Unterrichtsstudie war die Realisierung, Exploration und Optimierung eines auf der entworfenen Unterrichtsreihe basierenden Lehr-Lernszenarios. Die Studie kann als Basis für weitere quasi-experimentelle Vergleichsuntersuchungen dienen.

Wir geben im Folgenden einen Überblick über die Unterrichtsreihe, beschreiben das methodische Vorgehen der Explorationsstudie und berichten einige Ergebnisse.

4.2 Beschreibung der Unterrichtsreihe

Einstieg in das Themengebiet

Unsere grundsätzliche Überzeugung ist, dass ein sinnvoller und gewinnbringender Unterricht in Stochastik über den aufwendigeren Weg möglichst authentischer und konkreter Anwendungen im täglichen Leben gehen sollte. So steht am Beginn der Unterrichtseinheit keine mathematische Frage, sondern ein reales Problem, das sich auch Schülern durchaus stellen könnte: Wie gut ist der AIDS-Test? Dass sich diese Problematik besonders gut als Beispiel für eine reale Anwendung des Satzes von Bayes eignet, ist nicht neu. Z.B. im Unterrichtsvorschlag von Böer (1997) oder Israel (2001) findet es Verwendung. Bei uns dient der folgende relativ offene Medientext als Einstieg:

17 Es existiert ein bisher unveröff. Manuskript: Wassner, C., Biehler, R., Schweynoch, S. & Martignon, L. Authentisches Anwenden der Bayesschen Regel - Arbeitsmaterialien und didaktische Kommentare für den Stochastikunterricht der Sekundarstufe I.

Wie gut ist der AIDS-Test?

Der sogenannte AIDS-Test ist einer der zuverlässigsten Tests, die jemals entwickelt wurden. Er wird eingesetzt, um eine Infektion mit HIV festzustellen*. Wegen der hohen Gefahr der Verbreitung der tödlichen HIV-Infektion war sogar lange Zeit in der Diskussion, ob nicht die gesamte Bevölkerung zum AIDS-Test gezwungen werden soll.

Der AIDS-Test ist aber nicht perfekt. Wenn jemand HIV-infiziert ist, soll der Test positiv sein. Zu 99,9% fällt er dann auch positiv aus. Andererseits wenn jemand nicht HIV-infiziert ist, soll der Test natürlich negativ sein. Zu 99,7% fällt er dann tatsächlich negativ aus.

Nehmen wir mal an, dass für alle Menschen in NRW ein AIDS-Test durchgeführt werden soll. Laut Schätzung des Robert-Koch-Instituts sind bundesweit 0,05% der Bevölkerung HIV-infiziert, die Quote kann auch für NRW angenommen werden. Die Bevölkerungsstatistik gibt an, dass in NRW 18.000.000 Menschen leben.

* Im Sprachgebrauch hat sich AIDS-Test eingebürgert. AIDS bezeichnet eigentlich die Krankheit, die man bekommen kann, wenn man mit HIV infiziert ist. HIV kommt vom engl. „human immunodeficiency virus" = „Immunschwäche-Virus beim Menschen".

Nach kurzer Diskussion über die Textinhalte soll versucht werden, eine Fragestellung zu konkretisieren: „Stell dir vor, eine beliebige Person aus NRW bekommt mitgeteilt, dass ihr Test positiv ist. Wie sicher kann sie sein, dass sie tatsächlich HIV-infiziert ist?" Nächster Schritt ist die Formulierung des zugehörigen wahrscheinlichkeitstheoretischen Problems, das gelöst werden soll: „Wie hoch ist die Wahrscheinlichkeit für eine HIV-Infektion, wenn der Test positiv ist: _______%?"

Vor dem Beginn des Lösungsprozesses sollen *alle* Schüler das Ergebnis schätzen. Die Schätzungen werden notiert. Die meisten Menschen schätzen in dieser Situation ein viel zu hohes Ergebnis (siehe oben). Wir glauben, dass über die Bewusstmachung der kognitiven Fehlleistung Motivation zur Problemlösung geweckt werden kann.

Die Lösungsfindung erfolgt über den im Labor erfolgreich untersuchten Weg der Übersetzung der Informationen in „natürliche Häufigkeiten" und Aufstellung von Häufigkeitsbäumen. Um die Schüler zum aktiven und sorgfältigen „Nachdenken über die Situation" zu bewegen, werden in der Anfangsstunde zur Häufigkeitslösung hinführende Fragen gestellt, die in eigenaktiver Arbeit (z.B. als Gruppenaufträge) mit Unterstützung des Lehrers bearbeitet werden sollen, z.B.:

- Was kann passieren, wenn ein HIV-Infizierter getestet wird? Was, wenn ein nicht HIV-Infizierter getestet wird? Schreibe alle Möglichkeiten auf!
- Welche Möglichkeiten würdest du als „Fehler des Tests" bezeichnen und wo lag der Test richtig?

- Verteile die Bevölkerung von NRW auf die verschiedenen Möglichkeiten. Wie viele Personen sind es jeweils?

Die eigentliche Schwierigkeit ist die Berücksichtigung der „Fehler des Tests". Die Schüler sollen in der ersten Auseinandersetzung mit dem realen Problem erkennen, dass vier mögliche „Testsituationen" auftreten können, von denen zwei „Fehler des Tests" darstellen. Ziel ist die Bestimmung einer Wahrscheinlichkeit für die Richtigkeit des Tests. Dass alle Menschen aus NRW am Test teilnehmen, ist zugegebenermaßen eine unrealistische Annahme. Die Einschränkung der Realität umgeht hier zunächst die schwierige Frage der Schätzung der Basisrate (Vertiefung später).

Im Anschluss erfolgt fragend-entwickelnd die Modellierung und Visualisierung der Problemstellung mit einem Häufigkeitsbaum (vgl. Abb. 11a).

Abb.11a: Modellierung des AIDS-Test Problems im Häufigkeitsbaum (ohne Lösungsstufe)

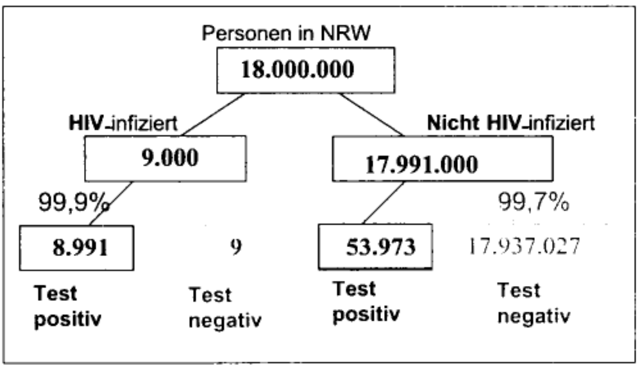
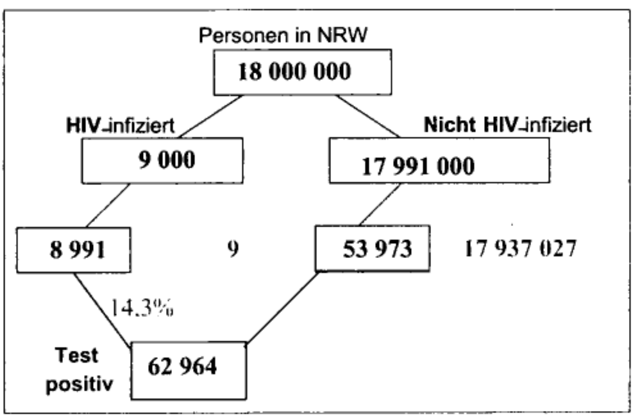


Abb.11b: Modellierung der Lösung des AIDS-Test Problems im Häufigkeitsbaum



Lösung: „Die Wahrscheinlichkeit für eine HIV-Infektion bei positivem Test" = „Der Anteil der HIV-Fälle unter allen positiven Tests" = 8 991 / 62 964 = 14,3%

Das Ziel der 1. Unterrichtsstunde sollte die Lösung des konkret gestellten Problems sein, die nun sehr viel leichter und intuitiver zu finden ist als mit der Bayesformel (vgl. Abb. 11b), und der Vergleich mit den Schülerschätzungen.

Weiterer Aufbau der Unterrichtsreihe [18]

Das weitere Vorgehen soll im Folgenden überblicksartig wiedergegeben werden:

Die 2. Unterrichtsstunde ist der Ergebnisvertiefung gewidmet. Das Ergebnis wirft eine Reihe von weiteren Fragen auf, die mit Hilfe relativ offener Materialien („Das AIDS-Testverfahren im Detail") geklärt werden sollen, z.B. „Warum wird der Test überhaupt durchgeführt, wenn er so „schlecht" ist?" oder „Was ist zu tun, um die Sicherheit für Getestete zu erhöhen?".

In Stunde 3 soll die Kontextvertiefung im Vordergrund stehen. Anhand von Medientexten oder Schilderung von Fällen wird den Schülern die praktische, individuelle und gesellschaftliche Bedeutung der Problematik des Themas AIDS und der Diagnose nahegebracht.

In Stunde 4 wird das Thema des Einflusses der Basisrate auf das Ergebnis vertieft. Die Schüler sollen ausgehend von verschiedenen Risikogruppen selbstständig entdecken, wie sich die Aussagekraft des Testes ändert. Das Vergleichen im Plenum verschiedener Ergebnisse, die in Kleingruppen erarbeitet wurden, soll diesen Einfluss für die Schüler plausibel machen. Nebenziel ist das Üben des Umgangs mit der Darstellungsform „Häufigkeitsbaum". In dieser Phase wird eine weitere Darstellungsform eingeführt, die Häufigkeitstabelle (siehe Abb. 12a), die ebenfalls sehr intuitiv erarbeitet werden kann. Die Gründe liegen einerseits in der Förderung einer höheren Flexibilität der Modellierung, andererseits in der tatsächlichen Benutzungsrelevanz im Alltag bzw. in der gymnasialen Oberstufe. Im Vergleich mit der Tabelle wird nun auch erstmals ein „vollständiger Häufigkeitsbaum" entwickelt („Doppelbaum", siehe Abb. 12b; vgl. Wassner et al., 2002). Im Folgenden können die Schüler beide Repräsentationen gleichberechtigt verwenden.

Abb. 12a: Modellierung des AIDS-Test Problems mit der Häufigkeitstabelle.

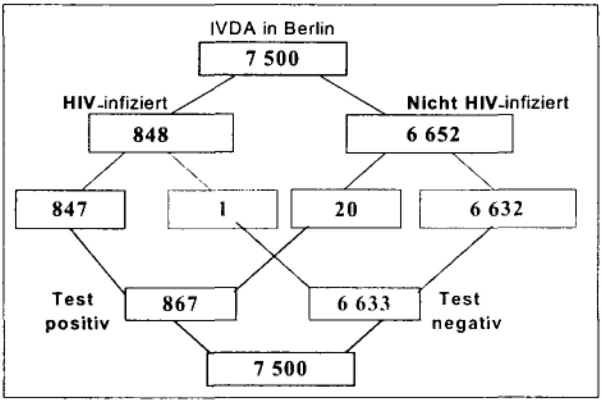
	Gesamt	HIV-infiziert	Nicht HIV-infiziert
Gesamt	7 500	848	6 652
TEST positiv	867	847	20
TEST negativ	6633	1	6632

Gesamtheit bezieht sich auf „intravenös Drogenabhängige (IVDA) in Berlin"

18 Die folgende Beschreibung gibt einen idealisierten zeitlichen Ablauf des Unterrichts wieder. Die Verteilung des Stoffes auf Schulstunden muss in der Realisierung flexibel gehandhabt werden.

In Stunde 5 wird das AIDS-Test Problem aus der „umgekehrten" Sichtweise eines Testentwicklers behandelt. Die gegebenen Basisrateninformationen sind nun auf die Testergebnisse bezogen, gesucht ist die Wahrscheinlichkeit für ein positives Testergebnis bei tatsächlich vorliegender HIV-Infektion. Es ist festzustellen, dass im Doppelbaum und auch in der Tabelle diese Sichtweise bereits integriert ist; man muss den Baum nur von unten „lesen", die Tabelle in gewisser Weise „in Zeilen".[19] Den Schülern kann so vor Augen geführt werden, dass stochastische Schlussfolgerungen nicht an kausale Richtungen (Krankheit - Testergebnis) gebunden sind. Diese besondere Eigenschaft führt erfahrungsgemäß sehr leicht zu Täuschungen (z.B. Falk-Paradoxon, Falk, 1979).

Abb. 12b: Modellierung des AIDS-Test Problems im vollständigen Häufigkeitsbaum



Gesamtheit bezieht sich auf „intravenös Drogenabhängige (IVDA) in Berlin"

Stunde 6 enthält einen „Überblick" mit Übungsaufgaben, in dem bewusst das bisher am konkreten Beispiel entwickelte nur „so weit wie nötig" verallgemeinert und einfache Begriffe eingeführt werden, die die Kommunikation erleichtern.

Die folgenden Stunden dienten dazu die verwendeten Repräsentationen und Regeln auf andere Sachkontexte zu verallgemeinern. Eine allgemeine Bayes-Regel als solche wurde bisher nicht entwickelt. Anhand eines Kriminalitätsfalles (Indizien dafür, ob jemand Mörder ist), wird einerseits der Häufigkeitsbaum zur Lösung verwendet, andererseits ist hier ein natürlicher Kontext gegeben, um Begriffe wie Hypothesen, Indizien sowie Apriori- und Aposteriori-Wahrscheinlichkeit einzuführen. Ziele sind das Begreifen der Problemsituation als eine Neubewertung einer Wahrscheinlichkeit unter neuen Daten (Revision der Wahrscheinlichkeit), das Einführen wichtiger, neuer Begriffe,

---

19 Trotz der hohen Analogie der Darstellungen Baum und Tabelle können Vorteile der Baumstruktur in der „sequenziellen" Art der Darstellung gesehen werden (vgl. Wassner et al., 2002, S.15).

unter denen die Diagnoseaufgaben aus der Medizin reformuliert und neu gedeutet werden sollen. Hiermit gewinnt man auch Anschluss an das Hypothesentesten. Dies war auch die ursprüngliche Intention der Arbeiten von Riemer (1985, 1991). Bei der späteren Behandlung von Hypothesentests in der Sekundarstufe II werden bedingte Wahrscheinlichkeiten berechnet (Signifikanz: P(Daten | Nullhypothese)). Dem Bayesianischen Zugang zur Statistik liegt dabei die Neubewertungsdenkweise zugrunde (Berechnung von P(Hypothese | Daten)). Diese curriculare Verbindung von bedingter Wahrscheinlichkeit und Hypothesentesten wird auch z.B. bei Krauss & Wassner (2001) als didaktische Veränderung vorgeschlagen und wäre eine Fortsetzung der Reihe in der Sekundarstufe II.

Wir haben bewusst darauf verzichtet, eine weitere Formalisierung der Bayes-Regel in der Sek. I, und sei es auch nur in Form des in Kapitel 2 erwähnten verkürzten umgekehrten Wahrscheinlichkeitsbaums, im Unterricht anzustreben. Die weiteren Unterrichtsstunden werden für die Erarbeitung von realen Anwendungsbeispielen genutzt. (z.B. zum Thema Schwangerschaftstest, Vaterschaftstest, Krebsfrüherkennung, BSE-Krise, Drogenkontrollen im Straßenverkehr, Urteile und DNA-Evidenz vor Gericht). Anhand dieser Beispiele sollen die Schüler erlernen, sowohl in Situationen zu denken, Material zu elaborieren und Fragestellungen vernünftig zu modellieren, als auch Darstellungsmittel zur Lösungsfindung zu benutzen und numerische Ergebnisse zu interpretieren.

4.3 Methodisches Vorgehen

Stichprobe und Unterrichtsplanung

Auf Grundlage der entwickelten Unterrichtsreihe wurden 144 Gymnasialschüler (fünf 9. Klassen aus NRW; 71 weiblich, 73 männlich[20]) von jeweils verschiedenen Lehrpersonen im Themengebiet „Anwenden der Bayesschen Regel" unterrichtet. Für die Realisierung wurden 15 Unterrichtsstunden bereitgestellt (5 Wochen à 3 x 45 Min.). Die Vorstellung der Reihe erfolgte in einer Fortbildungsveranstaltung und die konkrete Unterrichtsvorbereitung erfolgte von den Lehrenden auf Grundlage der vorgestellten Entwürfe. Die Ziele und Prozessplanung, spezielle Arbeitsmaterialien (Arbeitsblätter, Medientexte, Computerprogramme, Übungsaufgaben) und didaktische Hinweise wurden von uns vorgegeben. Didaktische Entscheidungen im Realisationsprozess waren jeweils in der Verantwortung der Lehrenden als Unterrichtsexperten.

20 Die Stichprobe bestand aus der gesamten 9. Jahrgangsstufe des Freiherr-vom-Stein Gymnasiums in Bünde (Westfalen), an dem der Mitautor der Reihe Stefan Schweynoch als Mathematiklehrer unterrichtet.

Erhebungen

Tab. 2: Ablaufplan, Unterrichtsstudie

Stichprobe	*Unterricht*	*Klassenarbeit*	*Befragungen*	*BAYES-Test*
Fünf 9. Klassen, Gymnasium NRW (N=144)	Beobachtung (Protokolle/ Video)	unmittelbar nach Training	Schülerfragebogen Lehrerfragebogen off. Gruppeninterview	ca. 4 Monate nach Unterrichtsende

Analyse des Unterrichts

Der Unterricht wurde teilweise (in einer der Klassen vollständig) per Videoaufzeichnung protokolliert und teilweise wurden schriftliche Verlaufs- und Inhaltsprotokolle vom Unterricht erstellt.

Tests zur Erfassung von Lernerfolgen

Am Ende der Unterrichtsreihe sollte bei allen Schülern eine Leistungsfeststellung mittels derselben Klassenarbeit durchgeführt werden, die in Absprache der Autoren der Reihe mit den teilnehmenden Lehrern entwickelt wurde. Die Klassenarbeit bestand aus zwei kontextuell verschiedenen Aufgaben, die jeweils in 5 Teilitems (a-e) untergliedert waren. Die Auswertungskriterien für die Bepunktung wurden exakt festgelegt. Es konnte ein Punktsummenwert von 48 erreicht werden (einzelne erreichbare Punktewerte hinter den Items).

Klassenarbeit

Aufgabe 1 (*27 Pkt.*):

Wenn jemand eine Katzenhaarallergie hat, so fällt der entsprechende Test auch zu 99% positiv aus. Umgekehrt erhalten 98% der Personen, die keine Katzenhaarallergie haben, ein negatives Testergebnis. Man vermutet, dass ca. 1% der Bevölkerung allergisch gegen Katzenhaare ist. 1.000 000 Personen nehmen an dem Allergietest teil.

a Vervollständige mit Hilfe der gegebenen Daten das abgebildete Doppelbaumdiagramm. Beschrifte alle Felder! (Struktur des Baumes ohne Beschriftung vorgegeben) *2 Pkt. (Beschriftung) + 5 Pkt. (Werte einsetzen)*

b Mit welcher Wahrscheinlichkeit hat eine Person eine Katzenhaarallergie, wenn der Test positiv ausfällt? *3 Pkt.*

c Berechne P(keine Allergie | Test negativ), P(Test negativ | keine Allergie) und P(keine Allergie und Test negativ). *6 Pkt. (jeweils 2 Pkt.)*

d Berechne eine mögliche Apriori-Wahrscheinlichkeit zu dem Allergietest und eine dazugehörige Aposteriori-Wahrscheinlichkeit. *6 Pkt. (jeweils 3 Pkt.)*

e Auf Blatt 2 siehst du den Baum eines Testdurchgangs, bei dem 50% der positiv Getesteten auch wirklich Allergiker waren. Wieviel Prozent der Nicht-Allergiker haben ein positives Ergebnis erhalten? *5 Pkt.*

Aufgabe 2 (*21 Pkt.*):
In München liegt die Wahrscheinlichkeit, dass ein abgestellter PKW aufgebrochen wird, bei 0,03%. Wir gehen davon aus, dass alle PKWs mit einer Alarmanlage ausgestattet sind, die laut Hersteller mit 99%iger Wahrscheinlichkeit anschlägt, wenn der PKW aufgebrochen wird. Leider geht die Alarmanlage auch manchmal los, wenn sich niemand am Auto zu schaffen macht. Die Wahrscheinlichkeit für einen Fehlalarm beträgt laut Hersteller 0,1%. (Hinweis: Gehe von 1 000 000 abgestellter PKWs aus)

a Ein PKW ist in München geparkt und die Alarmanlage aktiviert. Mit welcher Wahrscheinlichkeit wurde der PKW aufgebrochen, wenn der Alarm ertönt? *6 Pkt. (Repräsentation) + 3 Pkt. (Lösung)*

b Wie müssen die Wahrscheinlichkeitsangaben des Herstellers jeweils geändert werden, damit sich der Wert aus a) erhöht? (Begründung, keine Rechnung!) *2 + 2Pkt. (jeweils pro Änderung mit Begr.)*

c Vervollständige die in Berlin erhobene Statistik, für die sehr viele Autos über einen langen Zeitraum beobachtet wurden. (Tabellenstruktur mit einigen Häufigkeitswerten vorgegeben) *3 Pkt.*

d Mit welcher Wahrscheinlichkeit wurde in Berlin ein PKW aufgebrochen, wenn der Alarm ertönt? *3 Pkt.*

e Wie groß ist die Aufbruchswahrscheinlichkeit in Berlin? *2 Pkt.*

Es sollte ebenfalls in der Studie überprüft werden, ob Lernerfolge längerfristig stabil bleiben. Zur Untersuchung der Stabilität des Lernerfolges und Möglichkeit einer Verankerung mit den Laborstudien erfolgte nach etwa 4 Monaten mit allen Schülern ein *BAYES*-Leistungstest (identisch mit Labornachtest, vgl. 3.2)[21].

Befragungen (retrospektiv) zum erfolgten Unterricht
Nach Durchführung der Unterrichtsreihe erfolgten retrospektive Schüler- und Lehrerbefragungen. Der Schülerfragebogen enthielt eine Reihe von nichtleistungsbezogenen Items (z.B. zum eigenen Sachinteresse, zu selbst- und fremdbestimmter Motivation, Selbststeuerung, Tätigkeitsanreizen) und einige spezielle Einschätzungen zu verwendeten Darstellungsformen und Problemkontexten. Der Lehrerfragebogen hatte v.a. die Erfahrungen aus der durchgeführten Unterrichtsreihe sowie allgemeine Erfahrungen und Vorstellungen zum Lehren und Lernen von Stochastik zum Inhalt. Nachdem der Fragebogen durch die Lehrenden ausgefüllt worden war, wurde außerdem ein etwa 1,5-stündiges offenes Gruppeninterview mit allen Beteiligten

21 Aufgrund unterrichtstechnischer Gegebenheiten war nur *eine* Leistungsfeststellung zusätzlich zur üblichen Klassenarbeit möglich. Wir entschieden uns für den späteren „Stabilitätstest", da bereits die Klassenarbeit Befunde zum unmittelbaren Erfolg zeigte und wir die längerfristige Betrachtung für interessanter hielten.

geführt, um durch gemeinsame Reflektion noch eine Verfeinerung bzw. Spezifizierung des Feedbacks zu erreichen.

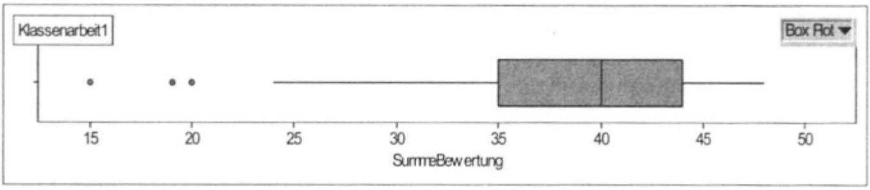
4.4 Ergebnisse

Wir stellen im Folgenden einige ausgewählte Ergebnisse dar, die Ergebnisse der Tests und Befunde aus dem Schülerfragebogen betreffen. Eine vollständige Auswertung aller erhobenen Daten erfolgt im Dissertationsprojekt von Wassner.

Testleistungen (Klassenarbeit)

In der Klassenarbeit konnte eine maximale Summe von 48 Testpunkten erreicht werden. Die Auswertung erfolgte nach der oben dargestellten Punkteverteilung pro Item.[22] Der folgende Boxplot zeigt die Verteilung über die Punkteskala und die Tabelle Kennwerte:

Abb. 13: Verteilung der Summe der Testpunkte in der Klassenarbeit, Kennwerte



Tab. 3: Verteilung der Summe der Testpunkte in der Klassenarbeit, Kennwerte

Variable	Summe Bewertung	Summe Aufg1	Summe Aufg2
Mittelwert	38,8	22,8	16,1
Median	40,0	23,0	18
Modus	42	27	19
Minimum	15	6	0
Maximum	48	27	21
Standardabweichung	6,54	3,89	4,10
Erstes Quartil	35,0	21,0	14,0
drittes Quartil	44,0	26,0	19,0

Mit dem Bepunktungsschema der Arbeit war auch die Analyse nach Schritten des Lösungsprozesses (vgl. 3.3) möglich.[23] Es ergab sich, dass die Schüler sehr kompetent das Aufstellen der Baumdiagramme (bzw. Tabelle) beherrschten.[24] Bei der Bildung des „Lösungsbruches“[25] aus dem aufgestellten Baumdiagramm traten häufiger Schwierigkeiten auf.[26] Dieser Befund ist auf verschiedene Fehlkonzepte zurückzuführen, die in diesem Zusammenhang

22 Die Auswertung wurde von uns durchgeführt und unabhängig davon auch jeweils für die Klasse von den Lehrenden. Wir berichten hier die eigene Auswertung.

23 Z.B. bei Aufgabe 2a: 6 Pkt. auf „labeln und strukturieren“, 3 Pkt. für „verarbeiten“

24 93,8% = (mittl.) Anteil der erreichbaren Punktsumme (max.13) für Schritte „labeln“ und „strukturieren“.

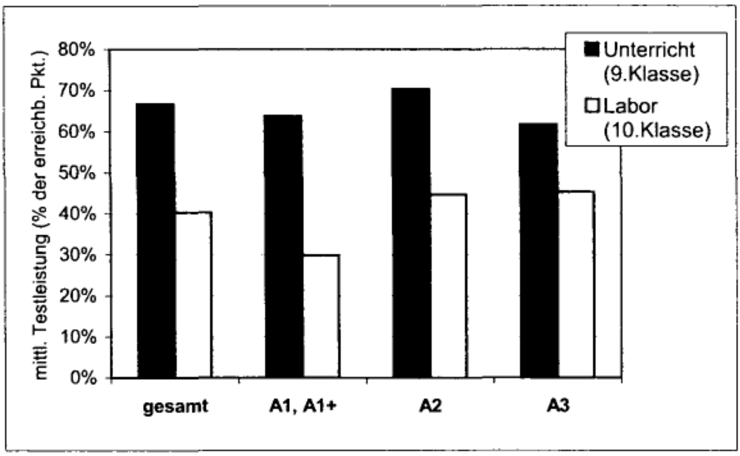
25 formal: $P(A|B) = \#(A \text{ und } B) / \#B$

26 78,6% = (mittl.) Anteil an der erreichbaren Punktsumme (max.14) für den Schritt „verarbeiten“.

auch in vorherigen Studien beobachtet wurden (vgl. z.B. Gigerenzer & Hoffrage, 1995, „non-Bayesian algorithms").[27]

Lernerfolgsstabilität (Bayes-Test)
Die Stabilität der Lernerfolge nach etwa 4 Monaten überprüften wir mit demselben Leistungstest, den bereits die Laborschüler absolviert hatten. Abbildung 14 zeigt den Vergleich der mittleren Testleistungen (als Anteile an den erreichbaren Punktsummen) zwischen den unterrichteten Gymnasialschülern der 9. Klasse und den Gymnasialschülern der 10. Klasse nach Häufigkeitstraining im Labor (Gruppe Freq (10), Nachtest 2 siehe Abb. 8a). Die Schüler erreichten nach 12-stündigem Unterricht noch nach 4 Monaten durchschnittlich etwa 2/3 der maximalen Testleistung. Im Vergleich dazu, hatten die Schüler im Labor ca. 12 Wochen nach Häufigkeitstraining etwa 40% der maximalen Testleistung gezeigt (nach Wahrscheinlichkeitstraining nur 7%). Nach verschiedenen Itemtypen differenziert, ergeben sich bei A1(+)-Items die größten Abstände, bei A3-Items die geringsten.

Abb. 14: Mittlere Testleistungen (in % der jeweils erreichbaren Punktsummen), gesamt und differenziert nach Itemtypen.



Schülerbefragung
Aus der Schülerbefragung gingen u.a. Befunde zur Bewertung der Darstellung Häufigkeitsbaum hervor.[28] In der Tabelle 4 sind entsprechende Items aus dem Schülerfragebogen (4-stufige Skala, 1: volle Zustimmung .... 4: volle Ablehnung) angegeben und jeweilige Kennwerte. Grundlegende Verstehensprobleme bei Bildung und Verwendung der Darstellungsform „Häufigkeitsbaum" werden in der Selbstbewertung eher nicht gesehen (D1, D2). Auch Wirksamkeits- und generelle Anreize der Häufigkeitsbaumverwendung als Lösungswerkzeug sind deutlich erkennbar (D3, D4, D5). Bei der

27 Die verschiedenen Lösungsstrategien werden in der Dissertationsarbeit von Wassner diskutiert.
28 Wir gehen hier aus Platzgründen nur auf die Items zu Darstellungsformen ein.

drei-alternativigen Frage einer Präferenz (PRI) für eine der beiden im Unterricht verwendeten Darstellungsformen (Häufigkeitsbaum bzw. -tabelle) sprechen sich 69% für den Baum, 5,8% für die Tabelle aus, und 25,2% wollen sich nicht festlegen.

Tab. 4: Items und Kennwerte des Schülerfragebogens zu Darstellungsformen

Häufigkeiten					
1	*2*	*3*	*4*	*Mittelwert*	*Item (Rating: 1:volle Zustimmung...4:volle Ablehnung)*
5	1	16	118	3.8	D1. Die Verwendung des Baumes habe ich bis jetzt nicht richtig verstanden
66	59	11	4	1.7	D2. Das Übersetzen der Textangaben in Werte im Baum habe ich leicht verstanden
67	57	9	6	1.7	D3. Ich konnte gut sehen, wie der Baum mir half, die Probleme zu lösen
73	39	18	10	1.8	D4. Es hat mehr Spaß gemacht, mit einem Baum zu arbeiten als mit irgendwelchen Formeln wie sonst in Mathe
57	66	11	6	1.8	D5. Im Baum wurde alles klarer, weil man da jeweils die Anzahlen von Personen erkennen konnte

1	*2*	*3*	*Item (mult.choice, 3-stufig)*
96	*8*	*35*	*PRI. Hast Du lieber den Baum oder die Tabelle verwendet? 1 Baum / 2 Tabelle / 3 beide etwa gleich*

4.5 Diskussion

Obwohl Aussagekraft und Testgüte der Klassenarbeit sicher diskussionswürdig sind, weist das hervorragende Gesamtergebnis doch auf einen deutlichen Lernerfolg durch den Unterricht hin. Die Stabilität des Lernerfolges ist sehr deutlich gegeben und im Vergleich mit im Labor erhaltenen Ergebnissen deutlich höher. In der Klassenarbeit bereitete das Finden der richtigen Lösung mit der erstellten Modellierung größere Schwierigkeiten als die Modellierung selbst. Es muss deutlich darauf geachtet werden, dass alle Schritte des Modellierungs- und Lösungsprozesses gleichermaßen verstanden und verinnerlicht werden. Das erkennbare Auftreten von systematischen Fehlkonzepten konnte nicht völlig eliminiert werden. Einige Fehlvorstellungen gründeten jedoch eindeutig auf fehlendem Grundkompetenzen, die in Klasse 9 bereits erworben sein müssten, wie etwa Grundeigenschaften der Wahrscheinlichkeit oder Prozentrechnung.

Aus Verwendungspräferenzen in der Klassenarbeit und Antworten der Schülerbefragung kann gefolgert werden, dass für die Schüler die Häufigkeitsbaumdarstellung kognitiv gut zugänglich war und als sehr hilfreich für die Lösung angesehen wurde.

Die Analyse und Bewertung der Unterrichtsreihe aus der retrospektiven Befragung der Lehrenden und aus Beobachtung des Unterrichts ist nach ersten

Analysen als sehr positiv einzustufen. Eine vollständige Darstellung dieser Ergebnisse ist Gegenstand weiterer Arbeiten (vgl. Wassner, Dissertation).

5. Ausblicke

Im Vergleich zu den (wenigen) anderen didaktischen Vorschlägen für die Behandlung der Bayes-Probleme in der Sekundarstufe I sehen wir in unsem Zugang wesentliche Vereinfachungen. Deutlich sichtbar wurde, dass anstatt der Einübung von Lösungsroutinen vor allem das Verständnis mit dem Häufigkeitszugang gefördert wurde. Allerdings erfolgte in unserer Untersuchungen kein systematischer Vergleich verschiedener Zugänge unter Unterrichtsbedingungen. Das muss gegebenenfalls einer zukünftigen Untersuchung vorbehalten bleiben.

In unseren Studien haben die Schülerinnen und Schüler der Sekundarstufe I erfolgreich gelernt, mit den von uns vorgeschlagenen Repräsentationsformen Bayes-Probleme zu lösen, ohne allerdings die Regel als solche in abstrakt-formaler Fassung kennenzulernen. Ein weiterer Bereich für zukünftige didaktische Untersuchungen wäre die Sekundarstufe II, in der die formalisierte Darstellung der Bayes-Regel in den meisten Lehrplänen Unterrichtsgegenstand ist. Die Hypothese, dass der Einstieg über natürliche Häufigkeiten auch zu einem besseren Verständnis der abstrakten Regel führt, ist unterrichtlich noch nicht untersucht worden[29]. Wir schlagen in jedem Falle vor, Häufigkeitsrepräsentationen als ersten Zugang zu Wahrscheinlichkeitsproblemen zu verwenden. Mit dieser Leitidee kann ein entsprechender Unterricht auch mit viel jüngeren Schülern erfolgreich sein.

Literatur

AK Stochastik in der Schule (2002). Empfehlungen zu Zielen und zur Gestaltung des Stochastikunterrichts. GDM Mitteilungen, Dezember 2002, 75-83.

Bea, W. (1995). Stochastisches Denken - Analysen aus kognitionspsychologischer und didaktischer Perspektive. In: Crott, H.W., Scholz, R.W. (Hrsg.). Psychologie des Entscheidungsverhaltens und des Konflikts. Frankfurt a.M.: Europäischer Verlag der Wissenschaften.

Böer, H. (1997). AIDS - Was ist von einem positiven Test-Ergebnis zu halten? In: Blum, W., König, G. & Schwehr, S. (Hrsg.). Materialien für einen realitätsbezogenen Mathematikunterricht, ISTRON-Reihe Bd.4, Hildesheim: Franzbecker. 38-57

Borovcnik, M. (1984). Der Problemkreis BAYESsche Formel. Mathematica didactica, 7, 207-224.

Borovcnik, M. (1992). Stochastik im Wechselspiel von Intuitionen und Mathematik. Mannheim: Bibliographisches Institut.

Daston, L. (1988). Classical probability in the enlightenment. Princeton University Press.

29 Laboruntersuchungen sind erfolgt (vgl. 3.1 u. Fußnote 4)

Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In: D. Kahneman, P. Slovic & A. Tversky (Eds.), Judgment under uncertainty: Heuristics and biases. Cambridge: Cambridge University Press, 249-267.

Falk, R. (1979). Revision of probability and the time axis. Tall., D. (ed.). Proceedings of 3rd International Conf. for the Psychology of Mathematics Education. Coventry: Warwick Univ., 64-66.

Fischbein, E. (1975). The intuitive sources of probabilistic thinking in children. Dordrecht: D. Reidel.

Gigerenzer, G. (1993). Die Repräsentation von Information und ihre Auswirkung auf statistisches Denken. In: Hell, W./Fiedler, K./Gigerenzer, G. (Hrsg.): Kognitive Täuschungen - Fehl-Leistungen und Mechanismen des Urteilens, Denkens und Erinnerns. Heidelberg: Spektrum, S. 99-127.

Gigerenzer, G. & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. Psychological Review, 102, 684-704.

Hacking, I. (1984). The emergence of probability. Cambridge University Press.

Harten, G. von, & Steinbring, H. (1984). Stochastik in der Sekundarstufe I. Köln: Aulis.

Hasher, L. & Zacks, R.T. (1979): Automatic and effortful processes in memory. Journal of Experimental Psychologie: General, 108, 356-388.

Israel, S. (2001). Was hat Aids mit Mathe zu tun? - Hinführung zum Satz von Bayes mit offenen Materialien. Mathematik lehren (104). 62-66.

König, G. (1991). Aids und Mathematikunterricht. Zentralblatt für Didaktik der Mathematik, 23(6), 207-220

Kolmogorov, A. (1933). Grundbegriffe der Wahrscheinlichkeitstheorie. Berlin: Springer.

Krauss, S. & Wassner, C. (2001). Wie man das Testen von Hypothesen einführen sollte. Stochastik in der Schule, Band 21, Heft 1. S.29-34.

Kütting, H. (1981). Didaktik der Wahrscheinlichkeitsrechnung. Freiburg: Herder.

Kütting, H. (1994). Didaktik der Stochastik. Mannheim: BI Wissenschaftsverlag.

Kuypers, W., Lauter, J. & Wuttke, H. (1999, Hrsg.). Mathematik 10. Schuljahr. Berlin: Cornelsen

Laplace, P.S. de (1812). Théorie analytique des probabilités. Paris: Veuve Courcier.

NRW (1993). Richtlinien und Lehrpläne für das Gymnasium, Sekundarstufe I, in Nordrhein-Westfalen, Mathematik. Frechen: Ritterbach-Verlag

Riemer, W. (1985). Neue Ideen zur Stochastik. Mannheim: B.I. Wissenschaftsverlag.

Riemer, W. (1991). Stochastische Probleme aus elementarer Sicht. Mannheim: BI-Wissenschaftsverlag.

Schmid, A & Weidig, I. (1996, Hrsg.). Lambacher Schweizer 9 - Mathematisches Unterrichtswerk für das Gymnasium, Ausgabe NRW. Stuttgart: Klett

Sedlmeier, P. (1997). BasicBayes: A tutor system for simple Bayesian inference. Behavior Research Methods, Instruments, & Computers. 29, 328-336.

Sedlmeier, P. (1999). Improving statistical reasoning: Theoretical models and practical implications. Mahwah: Lawrence Erlbaum Associates.

Sedlmeier, P. & Gigerenzer, G. (2001). Teaching Bayesian reasoning in less than two hours. Journal of Experimental Psychology: General 130(3). 380-400.

Shaughnessy, J.M. (1992). Research in Probability and Statistics: Reflections and Directions. D. Grouws (Ed.). Handbook on Research in Mathematics Teaching and Learning. New York: Macmillan Publ. Company, 465-494.

Tversky, A. & Kahneman, D. (1972). Subjective probability: A judgment of representativeness. Cognitive Psychology 3, 430-454.

Wassner, C., Krauss, S. & Martignon, L. (2002). Muss der Satz von Bayes schwer verständlich sein? Praxis der Mathematik 1/44. 12-16

Wassner C., Martignon L. & Sedlmeier, P. (2003). Nutzen alltagsorientierter Darbietungsformen für den Stochastikunterricht. In: Prenzel, M. & Doll, J. (Hrsg.). Bildungsqualität von Schule: Bedingungen mathematischer, naturwissenschaftlicher und überfachlicher Kompetenzen. 43. Beiheft der Zeitschrift für Pädagogik. 35-50.

Wickmann, D. (1990). Bayes-Statistik: Einsicht gewinnen und entscheiden bei Unsicherheit. Mannheim: BI-Wissenschaftsverlag.

Danksagung

Wir danken Stefan Schweynoch für die tatkräftige Mitarbeit bei der Erstellung der Unterrichtsreihe und der Durchführung der Unterrichtsstudie, insbesondere der Koordination der Unterrichtsstudie an der Schule. Außerdem den Lehrenden Frau Niederkleine, Herrn Gößling, Herrn Schweynoch, Herrn Weller und Herrn Hötger und allen Schülern des Freiherr-vom-Stein Gymnasiums Bünde (Westfalen), die bei der Unterrichtstudie mitgewirkt haben.

Anschrift der Autoren

Christoph Wassner
Prof. Dr. Rolf Biehler, Universität Kassel, Fachbereich 17,
Postfach 10 13 80, 34109 Kassel,
E-Mail: wassner@mathematik.uni-kassel.de
biehler@mathematik.uni-kassel.de

Prof. Dr. Laura Martignon, Pädagogische Hochschule Ludwigsburg,
Postfach 220, 71602 Ludwigsburg
E-Mail: martignon_laura@ph-Ludwigsburg.de