



Von Baumdiagrammen über Doppelbäume zu Häufigkeitsnetzen – kognitive Überlastung oder didaktische Unterstützung?

Karin Binder · Nicole Steib · Stefan Krauss

Eingegangen: 30. März 2020 / Angenommen: 24. Oktober 2022 / Online publiziert: 1. Dezember 2022
© Der/die Autor(en) 2022

Zusammenfassung In stochastischen Situationen mit zwei dichotomen Merkmalen erlauben weder die schulüblichen Baumdiagramme noch Vierfeldertafeln die simultane Darstellung sämtlicher in der Situation möglicher Wahrscheinlichkeiten. Das im vorliegenden Beitrag vorgestellte Netz hat die Kapazität, alle vier möglichen Randwahrscheinlichkeiten, alle vier Schnittwahrscheinlichkeiten sowie alle acht bedingten Wahrscheinlichkeiten *gleichzeitig* darzustellen. Darüber hinaus ist – aufgrund der Knoten-Ast-Struktur des Netzes – die simultane Darstellung von Wahrscheinlichkeiten *und* absoluten Häufigkeiten mit dieser Visualisierung ebenfalls möglich. Bei der sukzessiven Erweiterung des typischen Baumdiagramms zunächst zum Doppelbaum und schließlich zum Netz sinkt der Inferenzgrad (d.h. weniger kognitive Schritte sind erforderlich) z.B. für Fragen nach bedingten Wahrscheinlichkeiten, aber gleichzeitig steigt die Komplexität der Darstellung und somit die extrinsische kognitive Belastung. Im vorliegenden Artikel erfolgt zunächst ein theoretischer Vergleich dieser Knoten-Ast-Strukturen. Eine anschließende Studie illustriert, dass sich die sukzessive Erweiterung bereits vollständig ausgefüllter Diagramme positiv auf die Performanz von $N=269$ Schülerinnen und Schülern auswirkt. Obwohl *Häufigkeitsdoppelbäume* und *Häufigkeitsnetze* den Schülerinnen und Schülern gänzlich unbekannt waren, unterstützten diese Visualisierungen die Schülerinnen und Schüler bei der Bearbeitung der Aufgaben am meisten.

Schlüsselwörter Häufigkeitsnetz · Bedingte Wahrscheinlichkeiten · Extrinsische kognitive Belastung · Baumdiagramm · Doppelbaum

✉ Karin Binder

Ludwig-Maximilians-Universität München, Theresienstraße 39, 80333 München, Deutschland
E-Mail: Karin.Binder@lmu.de

Nicole Steib · Stefan Krauss

Universität Regensburg, Universitätsstraße 31, 93053 Regensburg, Deutschland

Moving from tree diagrams to double trees to net diagrams—cognitively overwhelming or educationally supportive?

Abstract In stochastic situations with two dichotomous events, neither typical tree diagrams nor 2×2 tables allow the simultaneous representation of all possible probabilities in the situation. The net diagram presented in this paper has the capacity to represent all four possible marginal probabilities, all four joint probabilities, and all eight conditional probabilities simultaneously. Furthermore, due to the node-branch structure of the frequency net, the simultaneous representation of probabilities and absolute frequencies is also possible with this visualization. With the successive extension of the typical tree diagram to the double tree and finally to the net diagram, the inference degree, e.g., for questions about conditional probabilities, decreases (i.e., less mental steps are required), however, at the same time the complexity of the representation increases and thus the extrinsic cognitive load. In the present article, a theoretical comparison of these node-branch-structures is made. Furthermore, we demonstrate with an empirical study that the successive extension of these node-branch structures, which were already completely worked out, positively affects the performance of $N = 269$ students. Although *frequency double trees* and *frequency nets* were entirely unfamiliar to the students, these visualizations provided the best support to the students in completing the tasks.

Keywords Natural frequency net · Conditional probabilities · Extraneous cognitive load · Tree diagram · Double tree

1 Einleitung

Der kompetente Umgang mit Visualisierungen mathematischer Situationen (mathematisch darstellen) ist in den Bildungsstandards als prozessbezogene Kompetenz für Lernende verankert (Kultusministerkonferenz 2022). Lehrkräfte wiederum sollen in Bezug auf mathematische Sachverhalte Visualisierungen begründet auswählen und einsetzen können, die für Schülerinnen und Schüler besonders verständnisfördernd wirken. Im Stochastikunterricht werden an verschiedenen Stellen der Lehrpläne beispielsweise Situationen mit zwei dichotomen Merkmalen visualisiert (z. B. Merkmal 1: Mädchen vs. Jungen; Merkmal 2: im Sportverein vs. nicht im Sportverein). Im Sinne des Spiralcurriculums werden solche Situationen in der Unterstufe zunächst auf der Basis absoluter und dann auch relativer Häufigkeiten behandelt, bis in der Regel in der gymnasialen Oberstufe eine Thematisierung mit dem Kalkül bedingter Wahrscheinlichkeiten erfolgt (z. B. Krauss et al. 2020).

Typische simultane Visualisierungen zweier dichotomer Merkmale sind zum Beispiel Vierfeldertafeln oder Baumdiagramme, auf deren Basis sich eine große Klasse verschiedener Aufgabentypen inklusive deren Lösungen illustrieren lassen. So können mit beiden Diagrammtypen in der Oberstufe beispielsweise Aufgaben zu Schnittwahrscheinlichkeiten oder bedingten Wahrscheinlichkeiten veranschaulicht werden. In jüngster Zeit werden in Forschungsarbeiten zur Visualisierung von Situationen mit zwei dichotomen Merkmalen vermehrt auch graphische Erweiterungen

von Baumdiagrammen wie der Doppelbaum oder das Netz vorgeschlagen (siehe Abschn. 2.5, Abb. 3), mit denen sich alle acht auftretenden bedingten Wahrscheinlichkeiten (statt nur vier wie beim Einfachbaum) oder im Netz sogar zusätzlich auch noch Schnittwahrscheinlichkeiten an den dafür vorgesehenen Ästen darstellen lassen (Baumdiagramme, Doppelbäume und Netze werden im Folgenden auch als „Knoten-Ast-Strukturen“ bezeichnet).

In allen bislang genannten Visualisierungen können absolute Häufigkeiten oder Wahrscheinlichkeiten (bzw. relative Häufigkeiten) eingetragen werden, wobei sich in empirischen Studien wiederholt vor allem solche Knoten-Ast-Strukturen für das Verständnis als hilfreich erweisen, die mit absoluten Häufigkeiten gefüllt sind (Binder et al. 2015, 2018a). Dessen ungeachtet finden sich jedoch auch Stimmen, die eine zu hohe Komplexität und Informationsdichte des Doppelbaumes oder des Netzes vermuten (z. B. Henze und Vehling 2021). In der Tat enthält ein im Häufigkeitsformat vollständig (d. h. mit allen stochastischen Informationen) ausgefüllter Doppelbaum und ein Netz jeweils neun verschiedene absolute Häufigkeiten, während im Wahrscheinlichkeitsformat ein entsprechend komplett ausgefüllter Doppelbaum 12 und ein Netz sogar 16 Wahrscheinlichkeiten enthält (siehe Abschn. 2.5, Abb. 3).

Im vorliegenden Beitrag wird die *sukzessive Erweiterung* des Baumdiagramms hin zum Doppelbaum und weiter zum Netz aus mathematischer und aus psychologischer Perspektive zunächst theoretisch beschrieben – getrennt für beide Informationsformate (Häufigkeiten vs. Wahrscheinlichkeiten). In einer anschließenden empirischen Studie gehen wir der Frage nach, ob die beiden Visualisierungen Doppelbaum und Netz – für den Fall, dass sie bereits vollständig ausgefüllt sind – die Beantwortung einer Frage nach einer bedingten Wahrscheinlichkeit erleichtern oder (im Vergleich zum schulüblichen „Einfachbaum“ bzw. einer reinen textuellen Präsentation der Informationen) eventuell aufgrund ihrer Komplexität sogar erschweren. Um dabei anschlussfähig zum bisherigen Forschungsstand zu sein, wählen wir für die empirische Studie exemplarisch sogenannte „Bayesianische Aufgaben“ (siehe Abschn. 2.2), die sich in vorangegangenen Studien als besonders anspruchsvoll erwiesen haben. Da frühere Studien nachweisen konnten, dass vor allem Häufigkeitsvisualisierungen verständnisfördernd sind, liegt der Fokus des vorliegenden Beitrags in erster Linie auf den Häufigkeitsvisualisierungen.

Die Formulierung einer Hypothese (bzgl. der Frage „Mit welcher bereits vollständig ausgefüllten Visualisierung lassen sich solche Aufgaben am leichtesten beantworten?“) erfolgt dann vor dem Hintergrund eines Widerstreits zweier Theorien: Aus psychologischer Sicht erläutern wir zunächst die zunehmende extrinsische kognitive Belastung bei solchen Erweiterungen von Knoten-Ast-Strukturen im Sinne der Cognitive-Load-Theorie (z. B. Chandler und Sweller 1991). Bei der sukzessiven Erweiterung ausgehend vom Baum über den Doppelbaum bis hin zum Netz werden auf der einen Seite zunehmend auch für die Lösung einer bestimmten Aufgabe unnötige Informationen dargestellt und die Visualisierungen werden somit komplexer. Diese dadurch höhere extrinsische Belastung dürfte auch möglichen Befürchtungen im Hinblick auf eine zu hohe Komplexität des Netzes zugrunde liegen (Henze und Vehling 2021). Aus mathematischer Sicht wird anschließend in Anlehnung an Curcios (1987) Schwierigkeitstaxonomie (*Level 1*: read the data; *Level 2*: read between the data; *Level 3*: read beyond the data) auf der Grundlage einer vollständig aus-

gefüllten Visualisierung das entsprechende Level (je niedriger, desto geringer der „Inferenzgrad“) für eine Frage nach einer bedingten Wahrscheinlichkeit bestimmt. Hierbei stellt sich auf der anderen Seite heraus, dass bei der sukzessiven Erweiterung natürlich auch immer mehr Lösungen verschiedenster Aufgabenstellungen enthalten sind, wodurch der Inferenzgrad abnimmt. Bei der gleichzeitigen Betrachtung beider Einflussfaktoren (zunehmende extrinsische kognitive Belastung vs. sinkender Inferenzgrad) haben Knoten-Ast-Strukturen die Eigenschaft, dass der Inferenzgrad umso niedriger ist, je mehr statistische Informationen in der bereits vollständig ausgefüllten Visualisierung enthalten sind.

In der empirischen Studie wird untersucht, welcher der beiden Effekte (erhöhte extrinsische kognitive Belastung vs. Reduktion des Inferenzgrads) bei Bayesianischen Aufgabenstellungen überwiegt. Die Ergebnisse können zu einer Klärung der Frage beitragen, ob erweiterte Knoten-Ast-Strukturen bereits „intuitiv erfassbar sind“ (d.h. ohne vorherigen Unterricht). Nur wenn der Effekt der extrinsischen kognitiven Belastung bei bereits ausgefüllten Häufigkeitsnetzen nicht überwiegt, sind Anschlussstudien sinnvoll, in denen der konkrete unterrichtliche Einsatz von Häufigkeitsnetzen untersucht wird und Schülerinnen und Schüler zum Beispiel eigenständig entsprechende Visualisierungen anfertigen müssen.

2 Theoretischer Hintergrund

2.1 Stochastische Situationen mit zwei dichotomen Merkmalen

In den schulüblichen Situationen mit zwei dichotomen Merkmalen (in denen zwei Ereignisse A und B und deren Gegenereignisse betrachtet werden) gibt es im Wesentlichen 16 verschiedene relevante Wahrscheinlichkeiten¹ (Böcherer-Linder und Eichler 2017; Bruckmaier et al. 2019; Krauss et al. 2020):

- Vier *Rand*wahrscheinlichkeiten:
 $P(A)$, $P(\bar{A})$, $P(B)$, $P(\bar{B})$
- Vier *Schnitt*wahrscheinlichkeiten:
 $P(A \cap B)$, $P(A \cap \bar{B})$, $P(\bar{A} \cap B)$, $P(\bar{A} \cap \bar{B})$
- Acht *bedingte* Wahrscheinlichkeiten:
 $P(A|B)$, $P(A|\bar{B})$, $P(\bar{A}|B)$, $P(\bar{A}|\bar{B})$, $P(B|A)$, $P(B|\bar{A})$, $P(\bar{B}|A)$, $P(\bar{B}|\bar{A})$

Als besonders schwierig haben sich in diesem Zusammenhang in zahlreichen Studien sogenannte Bayesianische Aufgaben erwiesen (Hoffrage et al. 2000; Kahneman et al. 1982; Operskalski und Barbey 2016; Spiegelhalter et al. 2011).

¹ Neben den 16 genannten Wahrscheinlichkeiten gibt es noch die Wahrscheinlichkeiten der trivialen Ereignisse $P(\emptyset)$, $P(\Omega)$ oder die Wahrscheinlichkeiten von Vereinigungen von Ereignissen, die im vorliegenden Beitrag nicht berücksichtigt werden.

2.2 Bayesianische Aufgaben und natürliche Häufigkeiten als alternatives Darstellungsformat

Unter einer Bayesianischen Aufgabe verstehen wir im Folgenden Aufgabentypen, in denen verschiedene statistische Informationen gegeben sind, die beispielsweise in der Medizin auch als Basisrate, Sensitivität und Falsch-Positiv-Rate (bzw. die Wahrscheinlichkeiten entsprechender Gegenereignisse) bezeichnet werden und ein sogenannter „positiver Vorhersagewert“ berechnet werden soll (vgl. Tab. 1). Eine berühmte Bayesianische Aufgabe ist die sogenannte Mammographie-Aufgabe (Eddy 1982). Diese kann ausschließlich Wahrscheinlichkeiten beinhalten (Tab. 1, links; z. B. „80 %“) oder mit sogenannten „natürlichen Häufigkeiten“ formuliert werden (Tab. 1, rechts; z. B. „80 von 100“), die erstmals von Gigerenzer und Hoffrage (1995) eingeführt wurden. Unter natürlichen Häufigkeiten versteht man ein geordnetes Paar natürlicher Zahlen in der Schreibweise „a von b“ (eine präzise Definition sowie eine ausführliche Einordnung dieser Schreibweise in mathematische und psychologische Konzepte findet sich in Krauss et al. 2020).

Die richtige Lösung (etwa 7,8 %) kann beispielsweise im Wahrscheinlichkeitsformat (links) mithilfe der *Formel von Bayes* oder mit den *Pfadregeln* berechnet werden und erscheint auf den ersten Blick überraschend niedrig. Im Häufigkeitsformat (rechts) ist die Lösung deutlich einfacher zu berechnen: 1030 Frauen (nämlich 80 + 950) erhalten ein positives Testergebnis, aber nur 80 dieser 1030 Frauen sind tatsächlich auch an Brustkrebs erkrankt. Eine Meta-Analyse von McDowell und Jacobs (2017) zeigt, dass etwa 24 % der Versuchspersonen derartige Textaufgaben (ohne zusätzliche Visualisierung) im Häufigkeitsformat richtig lösen können (im Wahrscheinlichkeitsformat sind es dagegen nur etwa 4 %). Diese Lösungsrate von 24 % ist allerdings immer noch bemerkenswert niedrig (Weber et al. 2018). Obwohl durch zahlreiche Studien seit über 25 Jahren bekannt ist, dass natürliche Häufig-

Tab. 1 Bayesianische Aufgabe, links: Wahrscheinlichkeitsversion der Mammographie-Aufgabe; rechts: Häufigkeitsversion der Mammographie-Aufgabe (nach Eddy 1982 und Gigerenzer & Hoffrage 1995)

	Wahrscheinlichkeitsversion der Mammographie-Aufgabe	Häufigkeitsversion der Mammographie-Aufgabe
Basisrate	Die Wahrscheinlichkeit, dass eine Frau, die zu einer Routineuntersuchung geht, Brustkrebs hat, beträgt 1 %.	100 von 10.000 Frauen, die zu einer Routineuntersuchung gehen, haben Brustkrebs.
Sensitivität	Wenn eine Frau, die zu einer Routineuntersuchung geht, Brustkrebs hat, dann beträgt die Wahrscheinlichkeit, dass sie ein positives Testergebnis erhält, 80 %.	80 von 100 Frauen, die zu einer Routineuntersuchung gehen und die Brustkrebs haben, erhalten ein positives Testergebnis.
Falsch-Positiv-Rate	Wenn eine Frau, die zu einer Routineuntersuchung geht, keinen Brustkrebs hat, dann beträgt die Wahrscheinlichkeit, dass sie ein positives Testergebnis erhält, 9,6 %.	950 von 9900 Frauen, die zu einer Routineuntersuchung gehen und die keinen Brustkrebs haben, erhalten ein positives Testergebnis.
Positiver Vorhersagewert	<i>Frage:</i> Wie groß ist die Wahrscheinlichkeit, dass eine Frau, die zu einer Routineuntersuchung geht und ein positives Testergebnis erhält, tatsächlich Brustkrebs hat?	<i>Frage:</i> Wie viele der Frauen, die zu einer Routineuntersuchung gehen und ein positives Testergebnis erhalten, haben Brustkrebs?

keiten gerade in Bayesianischen Situationen verständnisfördernd wirken (Binder et al. 2015; Hoffrage et al. 2000; Hoffrage und Gigerenzer 1998; McDowell et al. 2018; McDowell und Jacobs 2017), werden sie bislang nur im saarländischen (Saarländisches Ministerium für Bildung und Kultur 2016) und bayerischen Lehrplan thematisiert (Staatsinstitut für Schulqualität und Bildungsforschung ISB 2021).

2.3 Visualisierung der statistischen Informationen

Das Lösen solcher Bayesianischen Aufgaben kann – neben natürlichen Häufigkeiten – auch durch die zusätzliche Präsentation einer bereits vollständig ausgefüllten Visualisierung der statistischen Informationen unterstützt werden (McDowell und Jacobs 2017), jedoch wirken nicht alle Visualisierungen gleichermaßen verständnisfördernd: Bezüglich der in der Schule eher unüblichen Visualisierungen *Euler-Diagramm* (Micallef et al. 2012; Sloman et al. 2003) und *Rouletterad-Diagramm* (Brase 2008; Yamagishi 2003) gibt es beispielsweise gemischte Ergebnisse in Bezug auf die Wirksamkeit im Vergleich zu reinen Textvarianten (wie in Tab. 1). Robuste und stabile verständnisfördernde Effekte in Bayesianischen Situationen können vor allem bei den folgenden Visualisierungen nachgewiesen werden: *Bildgitter* (auch *icon arrays* genannt; siehe z. B. Sirota et al. 2014; Zikmund-Fisher et al. 2014), *Häufigkeitsgitter* (Cosmides und Tooby 1996; Garcia-Retamero und Hoffrage 2013), *Einheitsquadrate* (Bea 1995; Böcherer-Linder und Eichler 2017), *Baumdiagramme* (Binder et al. 2015, 2018a; Yamagishi 2003), *Doppelbäume* (Khan et al. 2015; Wassner 2004) und *Vierfeldertafeln* (Binder et al. 2015; Steckelberg et al. 2004). Ein Überblick über verschiedene Visualisierungen findet sich in Spiegelhalter et al. (2011) oder Binder et al. (2015).

Grundsätzlich können Visualisierungen unterschieden werden, die numerische Informationen enthalten (Wahrscheinlichkeiten, absolute Häufigkeiten oder beides) und Visualisierungen, die die Anteilswerte geometrisch oder über die Anzahl dargestellter Figuren darstellen. Im schulischen Stochastikunterricht kommen im deutschen Sprachraum bislang vor allem Baumdiagramme und Vierfeldertafeln zum Einsatz (siehe Abb. 1 und 3 oben), die jeweils numerische Informationen enthalten.

		Wahrscheinlichkeiten		Absolute Häufigkeiten			
Vierfeldertafel		T+	T-				
	B	0,8%	0,2%	1%	80	20	100
	\bar{B}	9,5%	89,5%	99%	950	8.950	9.900
		10,3%	89,7%	100%	1.030	8.970	10.000

Abb. 1 Schulübliche Vierfeldertafeln; *links*: mit Wahrscheinlichkeiten (bzw. relativen Häufigkeiten); *rechts*: mit absoluten Häufigkeiten

Baumdiagramme und Vierfeldertafeln können – im Gegensatz zu manchen der anderen Visualisierungen – schnell und mit geringem Aufwand auf dem Papier oder an der Tafel angefertigt werden (für ein Bildgitter müssten z. B. bei obigem Mammographie-Problem eine große Anzahl an Symbolen gezeichnet werden). Diese „Anwenderfreundlichkeit“ sowie die Möglichkeit, bequem numerische Informationen eintragen zu können, dürften Gründe für die Verbreitung dieser beiden Visualisierungen im Stochastikunterricht sein.

Während in den in der Schule üblicherweise verwendeten Vierfeldertafeln *Schnittwahrscheinlichkeiten* (aber keine bedingten Wahrscheinlichkeiten) fokussiert werden, stehen bei den schulisch typischerweise eingesetzten Baumdiagrammen umgekehrt *bedingte Wahrscheinlichkeiten* (aber keine Schnittwahrscheinlichkeiten) im Fokus. Es gibt allerdings verschiedene Varianten von Vierfeldertafeln, die unterrichtlich ebenfalls fruchtbar sein könnten (z. B. Vierfeldertafeln, die bedingte Wahrscheinlichkeiten enthalten und sich deshalb zeilen- bzw. spaltenweise zu 100% addieren, siehe Onlinematerial 1). Im Weiteren fokussieren wir ausschließlich die schulüblichen Vierfeldertafeln.

Im Folgenden sprechen wir von *Häufigkeitsvisualisierungen*, wenn eine Visualisierung lediglich absolute Häufigkeiten beinhaltet und von *Wahrscheinlichkeitsvisualisierungen*, wenn lediglich Wahrscheinlichkeiten abgebildet werden. In Vorbereitung auf das Wahrscheinlichkeitskonzept werden auch schon in früheren Jahrgangsstufen entsprechende Visualisierungen mit relativen Häufigkeiten betrachtet (für eine Unterscheidung siehe z. B. Krauss et al. 2020). In der Sekundarstufe II werden die abgebildeten relativen Häufigkeiten (bei großem N) schließlich auch als Schätzer für Wahrscheinlichkeiten interpretiert.

Vierfeldertafeln werden im Stochastikunterricht routinemäßig entweder mit Wahrscheinlichkeiten (bzw. relativen Häufigkeiten) oder mit absoluten Häufigkeiten (Abb. 1) eingesetzt. Baumdiagramme mit absoluten Häufigkeiten in den Knoten (siehe Abschn. 2.4, Abb. 2) werden dagegen bislang kaum unterrichtlich verwendet, sondern finden sich in Schulbüchern und Lehrplänen fast ausschließlich mit Wahrscheinlichkeiten (bzw. relativen Häufigkeiten) an den Ästen. Aus didaktischer Sicht ist es bedauerlich, dass Häufigkeitsbäume schulisch nur eine untergeordnete Rolle spielen, da sich in empirischen Studien gerade Visualisierungen mit absoluten Häufigkeiten (Abb. 1, rechts) als hilfreich erwiesen haben. Während in Binder et al. (2015) nur etwa 6% der untersuchten Schülerinnen und Schüler eine Bayesianische Aufgabe korrekt lösen konnten, die mit einer Wahrscheinlichkeitsvisualisierung (also einem Baum oder einer Vierfeldertafel mit Wahrscheinlichkeiten) dargestellt wurde, lag die Lösungsrate mit den entsprechenden Häufigkeitsvisualisierungen (d. h. Baum oder Vierfeldertafel mit absoluten Häufigkeiten) bei etwa 51%.

Obschon Erkenntnisse aus der kognitionspsychologischen Forschung zeigen, dass natürliche Häufigkeiten und Häufigkeitsvisualisierungen das Verständnis besser unterstützen als die entsprechenden Wahrscheinlichkeitsvisualisierungen, stehen Mathematiklehrkräfte natürlich vor der Aufgabe, im Unterricht das Wahrscheinlichkeitskonzept in seiner Breite zu entfalten und das entsprechende Kalkül zu vermitteln. Aus dieser Sicht ist die Frage für Lehrkräfte also nicht nur, welches Format leichter verständlich ist, sondern auch, wie das Häufigkeitskonzept und Häufigkeitsvisualisierungen genutzt werden können, um die Bedeutung von und das Rechnen

mit *Wahrscheinlichkeiten* zu verstehen (und letztendlich so auch die Konzepte der Schnitt- und der bedingten Wahrscheinlichkeit).

Eine Möglichkeit hierfür wäre beispielsweise die beiden in Abb. 1 gezeigten Vierfeldertafeln nebeneinander zu präsentieren, um die Zusammenhänge zwischen absoluten Häufigkeiten und Wahrscheinlichkeiten herauszuarbeiten. Der Effekt der geteilten Aufmerksamkeit („split attention effect“) aus der Theorie des multimedialen Lernens (Chandler und Sweller 1992) gibt jedoch Hinweise darauf, dass – aber das ist eine offene empirische Frage – eine solche *getrennte* Präsentation zu einer erhöhten kognitiven Belastung führen könnte.

Ein didaktischer Schlüssel für ein besseres Verständnis des Wahrscheinlichkeitskonzepts unter dezidiertem Nutzung des Häufigkeitskonzepts könnte demnach darin liegen, auf solche Visualisierungen zurückzugreifen, die die *gleichzeitige* Darstellung von Häufigkeiten und Wahrscheinlichkeiten und somit einen Repräsentationswechsel innerhalb *einer* Visualisierung erlauben. Dies ist bei *Knoten-Ast-Strukturen* wie Baumdiagrammen, Doppelbäumen oder den hier im Artikel vorgestellten Netzdiagrammen der Fall (siehe Abschn. 2.4, Abb. 2 und Abschn. 2.5, Abb. 3, bzw. Binder et al. 2018b, 2019). Wahrscheinlichkeiten können in Knoten-Ast-Strukturen immer an den Ästen und absolute Häufigkeiten in den Knoten eingetragen werden. Unter einem „Ast“ soll dabei im Folgenden immer die Verbindungslinie zwischen zwei Knoten verstanden werden. Durch eine Überlagerung eines Wahrscheinlichkeitsbaums und eines Häufigkeitsbaums (jeweils die obere Zeile in Abschn. 2.4, Abb. 2 und Abschn. 2.5, Abb. 3) entsteht so eine Visualisierung, die beide Informationsformate simultan repräsentiert (siehe Abschn. 2.6, Abb. 5).

Knoten-Ast-Strukturen eignen sich daher in besonderem Maße, um eine Brücke zwischen der Welt der Wahrscheinlichkeiten und der Welt der absoluten Häufigkeiten zu schaffen. Die absoluten Häufigkeiten in den Knoten erlauben bereits eine Behandlung von Situationen mit zwei dichotomen Merkmalen in der Sprache der natürlichen Häufigkeiten und die relativen Häufigkeiten an den Ästen können später problemlos ergänzt werden. Im weiteren Verlauf des Stochastikunterrichts kann schließlich mithilfe von Knoten-Ast-Strukturen die Beziehung zwischen relativen Häufigkeiten und Wahrscheinlichkeiten im Sinne einer frequentistischen Interpretation von Wahrscheinlichkeiten (Kolmogoroff 1933) thematisiert werden. Aus diesen Gründen wurden als experimentelle Stimuli für die vorliegende Studie Knoten-Ast-Strukturen und nicht etwa Vierfeldertafeln gewählt.

2.4 Sukzessive Erweiterung von Knoten-Ast-Strukturen im Häufigkeitsformat: Vom Häufigkeitsbaum über den Häufigkeitsdoppelbaum zum Häufigkeitsnetz

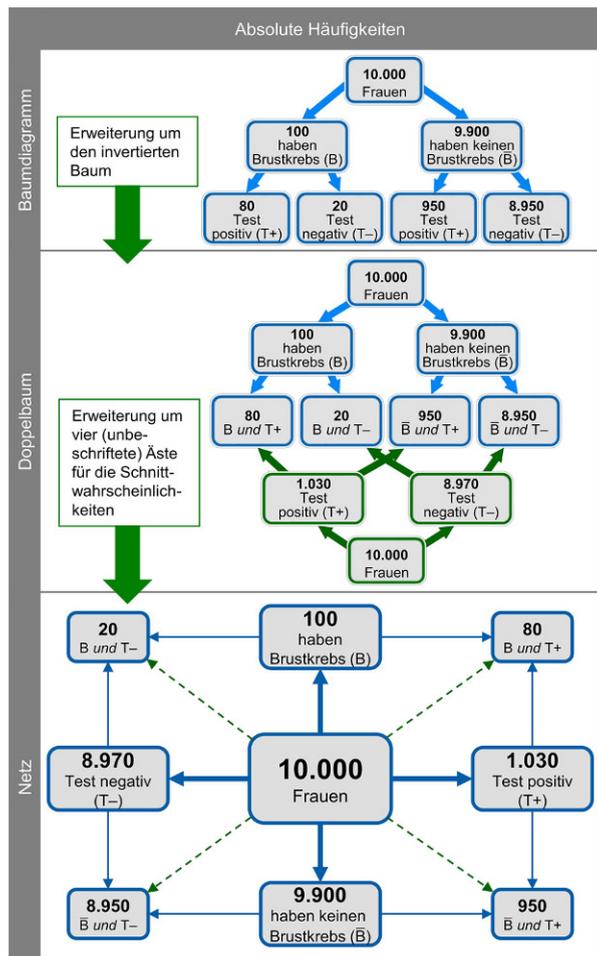
Im Folgenden sollen die aufeinander aufbauenden Knoten-Ast-Strukturen, die sich durch sukzessive Erweiterungen ergeben, näher erläutert werden. Dabei wird zunächst auf Häufigkeitsvisualisierungen eingegangen und im Anschluss werden die entsprechenden Wahrscheinlichkeitsvisualisierungen thematisiert. Durch die getrennte Betrachtung der Knoten-Ast-Struktur-Erweiterungen in beiden Informationsformaten soll verdeutlicht werden, was im jeweiligen Format konkret geschieht und was die Erweiterungen jeweils sowohl für den Inferenzgrad – in Bezug auf

oben beschriebene Bayesianische Aufgaben – als auch für die extrinsische kognitive Belastung bedeuten. Für beide Erweiterungen wählen wir exemplarisch den Kontext und die Zahlen aus der Mammographie-Aufgabe (Tab. 1).

2.4.1 Vom Häufigkeitsbaum zum Häufigkeitsdoppelbaum

Häufigkeitsbäume bestehen aus sieben Knoten und sechs Ästen (Abb. 2, oben), wobei sich absolute Häufigkeiten in den Knoten befinden, aber keine Informationen an den Ästen. Das in Abb. 2 (oben) dargestellte (einfache) Baumdiagramm enthält in der unteren Ebene – wie in der Schule üblich – nur die Bezeichnung des zweiten Merkmals (z. B. Test positiv), wobei auch eine ausführlichere Beschriftung denkbar wäre, die beide Merkmale berücksichtigt (z. B. Brustkrebs und Test positiv; vgl. hierzu auch die Beschriftung in der mittleren Ebene des Doppelbaums, Abb. 2, Mitte).

Abb. 2 Häufigkeitsbaum (oben), Häufigkeitsdoppelbaum (Mitte) und Häufigkeitsnetz (unten); die abgebildeten Visualisierungen wurden in leicht modifizierter Form (ohne die hier gezeigte Farbgebung) in der später beschriebenen empirischen Studie eingesetzt



Erweitert man den Häufigkeitsbaum unten um den invertierten Häufigkeitsbaum und somit um zwei neue absolute Häufigkeiten (genaugenommen werden drei Knoten und sechs Äste ergänzt), erhält man den Häufigkeitsdoppelbaum (Abb. 2, Mitte, vgl. auch Onlinematerial 1). Dieser enthält bereits alle neun absoluten Häufigkeiten, die in einer Situation mit zwei dichotomen Merkmalen von Relevanz sind.

Für Lernende ist es oft ein Problem zu entscheiden, welcher der beiden theoretisch möglichen Einfachbäume für eine bestimmte Aufgabenstellung zu wählen ist. Der Doppelbaum löst dieses Problem. Der Preis dafür ist, dass a) sich im unteren Teil zwei Äste kreuzen, und b) der Knoten mit der Gesamtstichprobe nun doppelt vorkommt (Binder et al. 2018b).

2.4.2 Vom Häufigkeitsdoppelbaum zum Häufigkeitsnetz

Wird der Häufigkeitsdoppelbaum auf bestimmte Weise umstrukturiert, so dass vier neue Äste, jedoch keine neuen absoluten Häufigkeiten ergänzt werden (der „doppelte“ Knoten mit der Gesamtstichprobe im Doppelbaum entfällt sogar), erhält man das sogenannte „Häufigkeitsnetz“ (Abb. 2, unten). Bei dieser Visualisierung gibt es im Vergleich zum Doppelbaum keine sich überschneidenden Äste, die falsche Antworten provozieren können (Binder et al. 2020b). Diese Häufigkeitsvisualisierung enthält somit (ebenfalls) alle relevanten absoluten Häufigkeiten einer Situation mit zwei dichotomen Merkmalen und bietet prinzipiell an den 16 Ästen Platz für alle Randwahrscheinlichkeiten, bedingten Wahrscheinlichkeiten und Schnittwahrscheinlichkeiten.

2.5 Sukzessive Erweiterung von Knoten-Ast-Strukturen im Wahrscheinlichkeitsformat: Vom Wahrscheinlichkeitsbaum über den Wahrscheinlichkeitsdoppelbaum zum Wahrscheinlichkeitsnetz

2.5.1 Vom Wahrscheinlichkeitsbaum zum Wahrscheinlichkeitsdoppelbaum

Der in Abb. 3 dargestellte Wahrscheinlichkeitsbaum (oben) enthält neben zwei von vier Randwahrscheinlichkeiten vier von acht möglichen bedingten Wahrscheinlichkeiten. Das einfache Baumdiagramm hat wieder den Nachteil, dass Schülerinnen und Schülern bei der Erstellung der Visualisierung nicht immer klar ist, ob in Bezug auf die obige Aufgabenstellung in der ersten Stufe des Baumes das Ereignis Erkrankungszustand oder das Ereignis Testergebnis stehen sollte. Wie auch im Häufigkeitsformat kann der Doppelbaum dieses Problem lösen, indem die Strukturen beider Bäume integriert werden (Binder et al. 2018b).

Bei Wahrscheinlichkeitsdoppelbäumen (Abb. 3, Mitte) findet man neben allen vier Randwahrscheinlichkeiten alle acht bedingten Wahrscheinlichkeiten, allerdings gibt es weiterhin keine entsprechenden Äste für Schnittwahrscheinlichkeiten. Die Komplexität der Darstellung steigt auch in diesem Format beim Übergang vom Baum zum Doppelbaum, was wieder durch die zwei sich überkreuzenden Äste im Doppelbaum deutlich wird, die nun darüber hinaus sogar mit Zahlen belegt sind.

Man könnte nun die vier fehlenden Schnittwahrscheinlichkeiten noch zusätzlich in den Doppelbaum einpflegen, darunter leidet aber die Übersichtlichkeit (vgl. z. B.

Abb. 3 Wahrscheinlichkeitsbaum (*oben*), Wahrscheinlichkeitsdoppelbaum (*Mitte*) und Wahrscheinlichkeitsnetz (*unten*); die abgebildeten Visualisierungen wurden in leicht modifizierter Form (ohne die hier gezeigte Farbgebung) in der später beschriebenen empirischen Studie eingesetzt

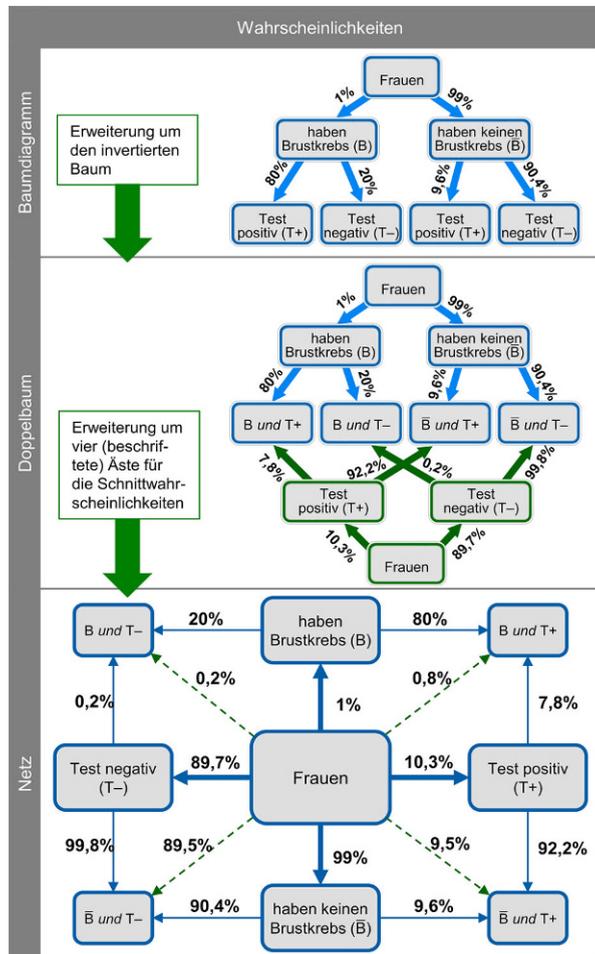
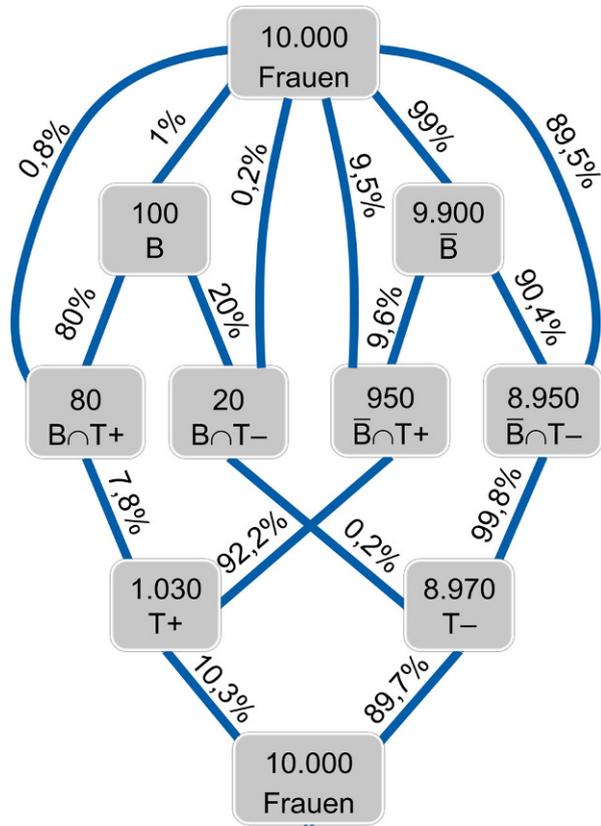


Abb. 4). Weitere denkbare Varianten von (Doppel-)Baumdiagrammen, die zusätzlich Schnittwahrscheinlichkeiten repräsentieren und die im Folgenden nicht thematisiert werden, finden sich in Onlinematerial 1.

2.5.2 Vom Wahrscheinlichkeitsdoppelbaum zum Wahrscheinlichkeitsnetz

Im Wahrscheinlichkeitsnetz (vgl. Abb. 3, unten) werden alle 16 Wahrscheinlichkeiten (bzw. relativen Häufigkeiten) dargestellt und dabei wieder alle Probleme des Doppelbaums bereinigt (wie die Überschneidung von Ästen und der doppelte Knoten für die Stichprobe). Im Gegensatz zu den in Onlinematerial 1 dargestellten Visualisierungen ist im Wahrscheinlichkeitsnetz für jede Wahrscheinlichkeit ein eigener Ast vorgesehen, so dass eine direkte Verbindung zwischen der Welt der absoluten Häufigkeiten, der relativen Häufigkeiten und der Wahrscheinlichkeiten hergestellt werden kann. Durch die erneute Strukturierung vom Doppelbaum zum Netz

Abb. 4 Doppelbaum mit zusätzlichen Schnittwahrscheinlichkeiten



im Wahrscheinlichkeitsformat ergibt sich eine Visualisierung, die einen vollständigen Informationssatz bereithält, was positive aber auch negative Aspekte mit sich bringen kann.

2.6 Die Struktur von Netzen mit beiden Informationsformaten

Abb. 5 zeigt die Struktur des Netzes, das aus neun Knoten und 16 Ästen besteht. Neben reinen Häufigkeitsnetzen (vgl. Abb. 2, unten) und reinen Wahrscheinlichkeitsnetzen (siehe Abb. 3, unten) ergibt sich durch eine „Überlagerung“ eine Struktur, in der beide Informationsformate gleichzeitig dargestellt werden können. Anhand von Abb. 5 sollen nun noch einmal kurz wesentliche Eigenschaften (inklusive empfohlener Beschriftung) rekapituliert werden.

Während die den Ereignissen entsprechenden Bezeichnungen in der unteren Hälfte der Knoten dargestellt werden, können die zugehörigen absoluten Häufigkeiten jeweils in die obere Hälfte der Knoten eingetragen werden. Die Beschriftung der Struktur kann im Unterricht beispielsweise im mittleren Knoten beginnen, in dem die Merkmalsträger (d.h. die Bezeichnung der Gesamtstichprobe, z.B. „Frauen“) abgebildet werden. In den weiteren Knoten werden (jeweils unten) analog die Be-

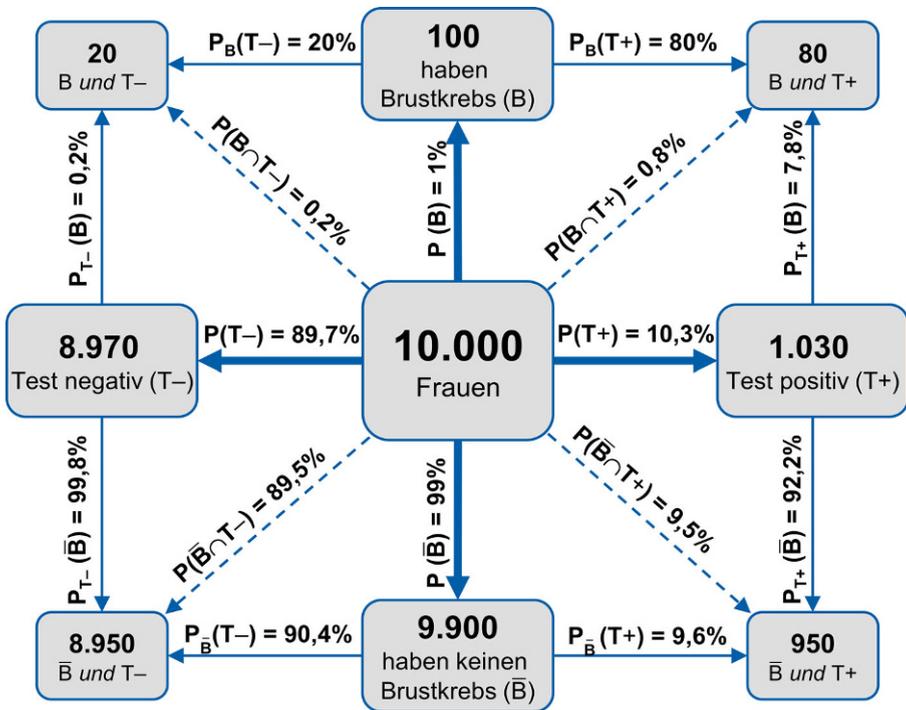


Abb. 5 Netz zum Mammographie-Problem, das neben allen absoluten Häufigkeiten und Wahrscheinlichkeiten auch die formalen Bezeichnungen der 16 auftretenden Wahrscheinlichkeiten beinhaltet

schreibungen der Ereignisse eingetragen. In den vier Eckknoten stehen die formalen Bezeichnungen der vier entstehenden Schnittereignisse: $B \cap T+$, $B \cap T-$, $\bar{B} \cap T+$ und $\bar{B} \cap T-$. Vertikal eingetragen ist in Abb. 5 der Gesundheitszustand (Brustkrebs vs. kein Brustkrebs), während das Testergebnis horizontal eingetragen ist (natürlich wären auch entsprechend gedrehte oder gespiegelte Darstellungen von Abb. 5 möglich; dass dies keine Rolle mehr spielt, ist gerade ein Vorteil des Netzes im Vergleich zum Einfachbaum).

An den Ästen können alle 16 Wahrscheinlichkeiten notiert werden, die in Abschn. 2.1 beschrieben wurden. Die vier *Randwahrscheinlichkeiten* befinden sich auf den vom mittleren Knoten jeweils vertikal und horizontal ausgehenden Ästen (siehe Abb. 5). Die *Schnittwahrscheinlichkeiten*, die üblicherweise nicht im Baumdiagramm enthalten sind, finden sich entsprechend an den vier diagonalen Ästen. An den acht Ästen, die den Rand des Netzes bilden, werden schließlich die acht *bedingten Wahrscheinlichkeiten* eingetragen. Die Wahrscheinlichkeit, dass der Test positiv wird, wenn die Frau Brustkrebs hat, steht demnach oben rechts auf dem Pfeil der vom Knoten des Ereignisses B ausgeht und im Knoten des Ereignisses $B \cap T+$ mündet. Die jeweils invertierten bedingten Wahrscheinlichkeiten (z. B. $P(B|T+)$ und $P(T+|B)$) befinden sich dabei immer „über Eck“. Didaktische Überlegungen zur konkreten Gestaltung des Netzes finden sich in Onlinematerial 2.

Zusammengefasst erfüllt das Netz (Abb. 5) die folgenden Eigenschaften:

- Beide Informationsformate (Wahrscheinlichkeiten und absolute Häufigkeiten) können aufgrund der Knoten-Ast-Struktur gleichermaßen dargestellt werden.
- Alle 16 Wahrscheinlichkeiten haben einen festen Platz im Diagramm.
- Es gibt keine Überschneidungen von Ästen wie im Doppelbaum (in graphentheoretischer Ausdrucksweise bedeutet das: Der Graph – bestehend aus Knoten und Kanten – ist planar bzw. plättbar).
- Keiner der Knoten wird wie im Doppelbaum doppelt dargestellt.

Rudimentäre Vorformen von Netzdiagrammen finden sich in Soto-Andrade (2019) und auf Wikipedia (ohne Datum): https://de.wikipedia.org/wiki/Satz_von_Bayes#/media/Datei:Bayes%27_Theorem_2D.svg.

2.7 Entlastung oder Überforderung durch Erweiterungen der Knoten-Ast-Strukturen?

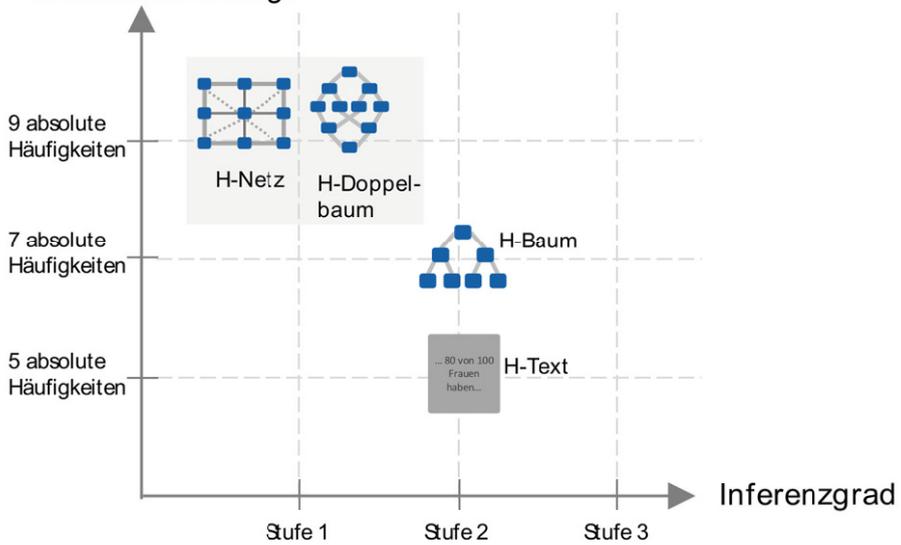
2.7.1 Reduzierter Inferenzgrad vs. extrinsische kognitive Belastung in Knoten-Ast-Strukturen

Zur theoretischen Analyse der im Titel des Beitrags aufgeworfenen Frage betrachten wir wiederum exemplarisch die in Tab. 1 dargestellte Bayesianische Aufgabe (in Wahrscheinlichkeiten bzw. Häufigkeiten), die entweder in der reinen Textversion vorliegt (Tab. 1 links bzw. rechts) oder zusätzlich mit einem bereits ausgefüllten Baumdiagramm, Doppelbaum oder Netz visualisiert ist (siehe Abb. 2 und 3). Prinzipiell lassen sich die folgenden Überlegungen analog für Aufgabenstellungen bezüglich jeder der 16 oben skizzierten Wahrscheinlichkeiten anstellen, wofür die Analysen bezogen auf eine Bayesianische Aufgabenstellung eine Vorlage bieten.

Die sukzessive Erweiterung der Knoten-Ast-Strukturen vom Baumdiagramm zum Doppelbaum und weiter zum Netz hat bei bereits vollständig ausgefüllten Diagrammen prinzipiell zwei Konsequenzen, die sich auf die Leistung von Schülerinnen und Schülern bei der Nutzung solcher vorausgefüllten Visualisierungen für entsprechende Aufgaben gegenläufig auswirken könnten: 1. Einerseits geht die Strukturverweiterung in vielen Fällen mit einer Reduktion des Inferenzgrads einher, was sich positiv auf die Performanz auswirken sollte. 2. Andererseits könnte die Strukturverweiterung aufgrund der zunehmenden Anzahl präsentierter Informationen und der zunehmenden Komplexität der Struktur zu einer erhöhten extrinsischen kognitiven Belastung führen, die sich negativ auf die Performanz von Versuchspersonen auswirken dürfte. Abb. 6 illustriert diesen Sachverhalt getrennt nach den beiden Informationsformaten Häufigkeiten (A) und Wahrscheinlichkeiten (B).

Inferenzgrad (kognitive Schritte) bei Knoten-Ast-Strukturen In seiner Schwierigkeitstaxonomie für statistische Diagramme unterscheidet Curcio (1987) die Level 1 „read the data“, Level 2 „read between the data“ und Level 3 „read beyond the data“. Daran angelehnt schlagen wir einen „Inferenzgrad“ für bereits ausgefüllte Knoten-Ast-Strukturen im Hinblick auf Bayesianische Aufgaben mit drei Stufen vor. In allen drei Stufen, die sogleich eingeführt werden, ist zunächst ein Zurechtfinden im dargestellten Diagramm entscheidend. Es müssen überdies die gegebenen statis-

a Informationsmenge



b Informationsmenge

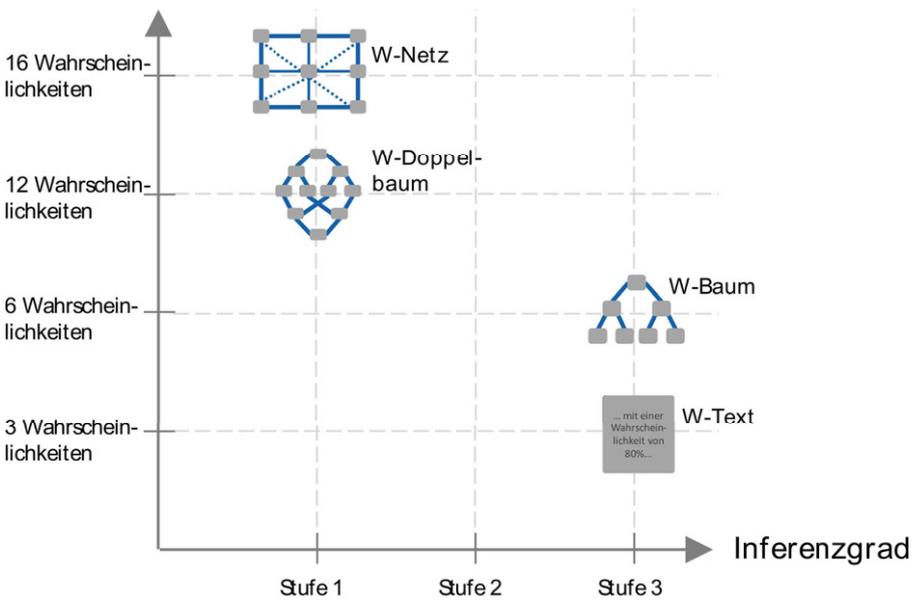


Abb. 6 Der zugehörige Inferenzgrad und die extrinsische Belastung durch die Informationsmenge bei den ausgefüllten Knoten-Ast-Strukturen Baum, Doppelbaum und Netz, wenn eine Bayesianische Aufgabe zugrunde liegt; **a** Häufigkeitsformat; **b** Wahrscheinlichkeitsformat

tischen Informationen erfasst werden, ggfs. unnötige Informationen „ausgeblendet“ werden und es muss erkannt werden, welche Information überhaupt gesucht wird. Die leichteste *Stufe 1* beschränkt sich dann auf die Identifikation einer im Diagramm abgebildeten Information. *Stufe 2* besteht (neben dem Erfassen gegebener Informationen und der Identifikation, welche Information gesucht ist) zusätzlich in einer einfachen additiven Verknüpfung. *Stufe 3* erfordert hingegen (neben dem Erfassen gegebener Informationen und der Identifikation, welche Information gesucht ist) eine genuine Inferenzbildung, in der mehrere Informationen (auch multiplikativ) miteinander in Beziehung gesetzt werden müssen (für Details siehe Abschn. 2.7.2 und 2.7.3). Je nach gewählter Stufe können Lernende unterschiedliche Kompetenzen entwickeln.

Extrinsische kognitive Belastung bei Knoten-Ast-Strukturen Auch wenn eine Knoten-Ast-Struktur wie das Netz allein aus fachmathematischer Perspektive durch ihre Vollständigkeit einen gewissen Charme aufweist und den Grad der erforderlichen Inferenz senkt (siehe Abschn. 2.7.2), legt die Cognitive-Load-Theorie nahe (Chandler und Sweller 1991; Sweller und Chandler 1991), dass die gezeigte Informationsmenge und im Besonderen die zusätzliche Präsentation überflüssiger Informationen auch von Nachteil sein kann. In der Cognitive-Load-Theorie wird unterschieden zwischen der intrinsischen Belastung (*intrinsic cognitive load*), die sich aus der Aufgabenschwierigkeit ergibt, der extrinsischen Belastung (*extraneous cognitive load*), die sich in der Präsentation der Lerninhalte niederschlägt und der lernbezogenen kognitiven Belastung (*Germane cognitive load*), der die freien kognitiven Ressourcen beschreibt (Niegemann 2008). Ziel jeglicher Instruktion sollte die Minimierung der sogenannten extrinsischen Belastung sein (*extraneous cognitive load*). Übertragen auf die verschiedenen Knoten-Ast-Strukturen (Baum, Doppelbaum und Netz) bedeutet das, dass die extrinsische Belastung durch die zusätzliche Präsentation irrelevanter Informationen steigt und die Aufgabenbearbeitung somit erschwert werden könnte. Das Bestimmen der gesuchten Wahrscheinlichkeit könnte also vor allem im Netzdiagramm scheitern, da hier ein Maximum an Informationen präsentiert wird und darüber hinaus sogar typische Fehler provozieren könnte, da übliche Falschantworten nun explizit abgebildet sind.

Im Folgenden wenden wir beide Konzepte zunächst auf Häufigkeitsvisualisierungen und dann auf Wahrscheinlichkeitsvisualisierungen an, wobei wir jeweils mit dem höchsten Inferenzgrad beginnen.

2.7.2 Reduzierter Inferenzgrad vs. extrinsische kognitive Belastung – Anwendung auf Häufigkeitsvisualisierungen

Inferenzgrad (kognitive Schritte) in den Häufigkeitsvisualisierungen Der höchste Inferenzgrad (nämlich Stufe 2) ergibt sich in den Häufigkeitsversionen in der reinen Textversion und bei einem ausgefüllten Häufigkeitsbaum (ein Inferenzgrad auf Stufe 3 ergibt sich erst in den Wahrscheinlichkeitsversionen; vgl. Abb. 6). Möchte man in diesen beiden Versionen die Frage beantworten, wie viele der positiv getesteten Frauen erkrankt sind, erfordert dies die folgenden kognitiven Schritte: Zunächst müssen die Informationen aus dem Text bzw. dem Häufigkeits-

baum richtig erfasst werden und auch verstanden werden, welche Information gesucht ist. Die erkrankten Frauen mit positivem Mammogramm (80) und gesunden Frauen mit positivem Mammogramm (950) müssen identifiziert und addiert werden, um die Gesamtzahl der Frauen mit positivem Mammogramm (1030) zu erhalten. Anschließend kann bereits die Lösung (z. B. im Häufigkeitsformat) „80 von 1030“ notiert werden. Der Inferenzgrad liegt also sowohl bei der reinen Textvariante als auch mit einem zusätzlich ausgefüllten Häufigkeitsbaum auf Stufe 2 da ein einfaches Identifizieren der gesuchten Information nicht möglich ist (das wäre Stufe 1), sondern zunächst eine additive Verknüpfung erfolgen muss.

Durch die Strukturweiterung auf einen Häufigkeitsdoppelbaum ist keine Addition mehr notwendig, da die Lösung aus der Visualisierung abgelesen werden kann (vgl. Abb. 2). Dies ist möglich, da der Häufigkeitsdoppelbaum (im Mammographie-Problem) die benötigte Zahl aller Frauen mit positivem Testergebnis bereits enthält. Somit müssen zur Formulierung der Antwort „80 von 1030“ lediglich die Zahlen „80“ und „1030“ aus der Häufigkeitsvisualisierung entnommen werden. Dies entspricht dem Inferenzgrad von Stufe 1.

Da auch das Häufigkeitsnetz alle neun absoluten Häufigkeiten der Bayesianischen Aufgabe beinhaltet, gelten hier alle Überlegungen analog zum Häufigkeitsdoppelbaum (Stufe 1).

Extrinsische kognitive Belastung bei Häufigkeitsvisualisierungen Während in der reinen Textvariante zum Mammographie-Problem (im Häufigkeitsformat) lediglich fünf absolute Häufigkeiten gegeben sind (von denen zur Darstellung der Lösung nur zwei benötigt werden), beinhaltet der einfache Häufigkeitsbaum noch zwei zusätzliche absolute Häufigkeiten, die zur Bestimmung der Wahrscheinlichkeit $P(\text{Brustkrebs}|\text{positives Mammogramm})$ ebenfalls nicht benötigt werden (nämlich 20 und 8950; vgl. Abb. 2, oben). Im Häufigkeitsdoppelbaum kommen noch einmal drei weitere absolute Häufigkeiten hinzu, wobei eine dieser drei absoluten Häufigkeiten die Gesamtanzahl der Personen darstellt und somit doppelt in der Visualisierung auftritt.

Die Anzahl gezeigter absoluter Häufigkeiten fällt bei der Erweiterung vom Häufigkeitsdoppelbaum (10 absolute Häufigkeiten) zum Häufigkeitsnetz (9 absolute Häufigkeiten) ab, da der im Häufigkeitsdoppelbaum gezeigte „doppelte“ Knoten im Häufigkeitsnetz entfällt. Allerdings könnten auch die vier zusätzlichen Äste für die Schnittwahrscheinlichkeiten (vgl. Abb. 2) die extrinsische kognitive Belastung erhöhen. Diese enthalten zwar keine Wahrscheinlichkeiten, laden aber zur mentalen Verknüpfung der entsprechenden absoluten Häufigkeiten ein und provozieren so möglicherweise ebenfalls Falschantworten.

2.7.3 Reduzierter Inferenzgrad vs. extrinsische kognitive Belastung – Anwendung auf Wahrscheinlichkeitsvisualisierungen

Inferenzgrad (kognitive Schritte) in den Wahrscheinlichkeitsvisualisierungen

In der Wahrscheinlichkeitsversion der Visualisierungen ergibt sich durch die Strukturweiterungen ebenfalls eine Reduktion des Inferenzgrads (Abb. 6): Sowohl in der reinen Textversion als auch bei der zusätzlichen Präsentation eines ausgefüllten

Wahrscheinlichkeitsbaums ist zur Bestimmung von $P(\text{Brustkrebs}|\text{positives Mammogramm})$ eine genuine Inferenz durch Multiplikation, Addition und Quotientenbildung erforderlich (Stufe 3).

Wenn die statistischen Informationen anhand des Wahrscheinlichkeitsdoppelbaums und des Wahrscheinlichkeitsnetzes und die dadurch illustrierte Gesamtsituation richtig interpretiert werden und im Besonderen verstanden wird, welche Information in der Aufgabenstellung gesucht ist, kann diese Wahrscheinlichkeit in beiden Visualisierungen direkt abgelesen werden. Durch die Präsentation eines Wahrscheinlichkeitsdoppelbaumes oder Wahrscheinlichkeitsnetzes reduziert sich der Inferenzgrad auf Stufe 1 (der Inferenzgrad Stufe 2 wird somit in den Wahrscheinlichkeitsversionen nicht angenommen).

Extrinsische kognitive Belastung bei Wahrscheinlichkeitsvisualisierungen

Während in der reinen Textversion drei Wahrscheinlichkeiten gegeben sind, werden im Baumdiagramm bereits sechs, im typischen Doppelbaum (vgl. Abb. 3) zwölf und im Netz 16 Wahrscheinlichkeiten abgebildet. Die Strukturweiterungen im Wahrscheinlichkeitsformat haben durch die zusätzlich präsentierten statistischen Informationen also ebenfalls ein zunehmendes Potenzial für eine erhöhte extrinsische kognitive Belastung.

Vor allem die vier zusätzlich eingezeichneten Äste für die Schnittwahrscheinlichkeiten mit den zugehörigen Zahlenwerten im Wahrscheinlichkeitsnetz könnten im Vergleich zum Wahrscheinlichkeitsdoppelbaum nicht nur zu einer höheren extrinsischen Belastung führen, sondern die Präsentation der Schnittwahrscheinlichkeit $P(\text{Brustkrebs} \cap \text{positives Mammogramm})$ könnte wieder eine typische Falschantwort provozieren (Binder et al. 2020b; Eichler et al. 2020; Gigerenzer und Hoffrage 1995).

2.8 Bisherige empirische Befunde zu Häufigkeits- und Wahrscheinlichkeitsnetzen

Erste empirische Befunde zum Häufigkeitsnetz – allerdings jeweils mit Studierendenstichproben – finden sich in Binder et al. (2020b) sowie in Kunzelmann et al. (2022). In beiden Studien hatten Studierende die Aufgabe, bereits vollständig ausgefüllte Diagramme zu Bayesianischen Situationen zu interpretieren und die Lösung auf die Frage nach einer (invertierten) bedingten Wahrscheinlichkeit anzugeben.

Die Studierenden verschiedener Fachrichtungen erhielten in Binder et al. (2020b) vollständig ausgefüllte Netze, Vierfeldertafeln oder Doppelbäume, die entweder nur mit absoluten Häufigkeiten gefüllt waren oder nur mit Wahrscheinlichkeiten. Die richtige Interpretation des ausgefüllten Diagramms gelang den Studierenden bei Häufigkeitsnetzen ähnlich gut wie bei Häufigkeitsdoppelbäumen (jeweils ca. 60% Lösungsrate). Ein ähnliches Bild zeigte sich auch für Wahrscheinlichkeitsnetze und Wahrscheinlichkeitsdoppelbäume, wobei die Lösungsrate hier insgesamt deutlich niedriger war (jeweils 34% Lösungsrate) als in den entsprechenden Häufigkeitsvarianten. Binder et al. (2020b) präsentieren in diesem Zusammenhang erste Erkenntnisse bezüglich eines systematischen Vergleichs verschiedener Wahrscheinlichkeits- und Häufigkeitsvisualisierungen sowie einen Vergleich mit Vierfelderta-

feln, die schulisch häufig im Zusammenhang mit Bayesianischen Aufgaben herangezogen werden. Während 78% der Studierenden beispielsweise die gesuchten Zahlenangaben in einer vollständig ausgearbeiteten Vierfeldertafel mit Häufigkeiten identifizieren konnten, gelang es nur 19% der Studierenden in einer Vierfeldertafel mit Wahrscheinlichkeiten, die entsprechenden Zahlenangaben zu identifizieren und miteinander in Beziehung zu setzen. In Binder et al. (2020b) wird weiterhin die Abhängigkeit der Häufigkeit typischer Fehler vom Informationsformat (Wahrscheinlichkeiten vs. Häufigkeiten) und der jeweils eingesetzten Visualisierungen diskutiert.

Im Fokus der Studie von Kunzelmann et al. (2022) standen vollständig ausgefüllte Doppelbäume und Netze, die von Medizinstudierenden interpretiert wurden. Der Effekt von Häufigkeitsvisualisierungen im Vergleich zu Wahrscheinlichkeitsvisualisierungen war hier deutlich geringer, was dafür spricht, dass Medizinstudierende möglicherweise bereits ein ausreichendes konzeptuelles Wissen für den Unterschied von $P(A|B)$ und $P(B|A)$ aufgebaut haben, jedoch das prozedurale Wissen für die Berechnung entsprechender Wahrscheinlichkeiten fehlt.

Inwieweit bereits Schülerinnen und Schüler ohne vorherige Instruktion von (ausgefüllten) Doppelbäumen und Netzen profitieren können, ist noch eine offene Frage, ebenso steht der Vergleich mit den schulüblichen Einfachbäumen noch aus. Diese Fragen werden in der vorliegenden Studie untersucht und die Ergebnisse anschließend im Lichte der beiden widerstreitenden Theorien interpretiert. Bislang wurde noch nicht untersucht, welcher dieser beiden Effekte bei der sukzessiven Erweiterung von ausgefüllten Knoten-Ast-Strukturen von Baumdiagramm über Doppelbaum hin zum Netzdiagramm überwiegt.

3 Forschungsfragen und Hypothesen

Betrachtet man nun Doppelbäume und Netze als sukzessive Erweiterungen von (einfachen) Baumdiagrammen im Vergleich, stellt sich die Frage, welcher der beiden eben beschriebenen Effekte überwiegt, wenn Schülerinnen und Schüler für eine Bayesianische Aufgabe (im Sinne von Tab. 1) neben den üblichen statistischen Informationen (Basisrate, Sensitivität, Falsch-Positiv-Rate) zusätzlich eine bereits vollständig ausgefüllte Knoten-Ast-Struktur erhalten: Der reduzierte Inferenzgrad durch die Präsentation der Lösung in einer Visualisierung oder die erhöhte extrinsische kognitive Belastung durch teilweise unnötige zusätzliche Informationen.

In der vorliegenden Studie wird untersucht, welchen Einfluss bereits vollständige ausgefüllte Häufigkeitsvisualisierungen – nämlich *Häufigkeitsbaum*, *Häufigkeitsdoppelbaum* und *Häufigkeitsnetz* – auf die Performanz von Schülerinnen und Schülern bei Bayesianischen Aufgaben (im obigen Sinne) haben. Häufigkeitsvisualisierungen stehen im Fokus (sie werden in der Studie durch zwei Kontexte operationalisiert), da sie sich in früheren Studien als besonders fruchtbar erwiesen haben. Analog sollen (als untergeordnete Fragestellung und lediglich als Additum) auch entsprechende Wahrscheinlichkeitsvisualisierungen untersucht werden (mit einem Stimulus). Der sonst bei solchen Studien fokussierte Performanzvergleich zwischen den beiden Informationsformaten ist nicht intendiert.

3.1 Hypothese zu Häufigkeitsvisualisierungen

Es wird erwartet, dass sich die zusätzliche Präsentation eines ausgefüllten Häufigkeitsbaumes im Vergleich zur reinen Textvariante (in natürlichen Häufigkeiten) vorteilhaft auf die Performanz der Versuchspersonen auswirkt, weil dieser Effekt bereits in zahlreichen früheren Untersuchungen nachgewiesen werden konnte (Binder et al. 2015; McDowell und Jacobs 2017). Die Erweiterung zum Häufigkeitsdoppelbaum sollte sich darüber hinaus positiv auf die Performanz auswirken, da in dieser Variante der Inferenzgrad reduziert wird und ein einfaches Identifizieren der beiden absoluten Häufigkeiten ausreicht. Eine zusätzliche Erweiterung zum Häufigkeitsnetz sollte dann keinen zusätzlichen positiven Effekt mehr bewirken, weil der Inferenzgrad identisch bleibt. Da die Anzahl der verschiedenen absoluten Häufigkeiten zwischen Doppelbaum und Netz gleich bleibt, sollten die vier zusätzlich eingezeichneten Äste zu keiner erhöhten extrinsischen kognitiven Belastung führen. Diese zusätzlichen Äste (ohne Wahrscheinlichkeiten) könnten allerdings dazu verleiten, die beiden Knoten auch gedanklich zu verbinden und so versehentlich eine Schnittwahrscheinlichkeit statt der gesuchten bedingten Wahrscheinlichkeit als Antwort anzugeben.

3.2 Hypothese zu Wahrscheinlichkeitsvisualisierungen

Hier wird erwartet, dass die Erweiterung von einem ausgefüllten Wahrscheinlichkeitsbaum auf einen ausgefüllten Wahrscheinlichkeitsdoppelbaum einen positiven Einfluss darauf haben dürfte, dass Schülerinnen und Schüler die Aufgabe korrekt lösen, weil diese Erweiterung mit einer Reduktion des Inferenzgrads einhergeht. Die Erweiterung von einem ausgefüllten Wahrscheinlichkeitsdoppelbaum auf ein ausgefülltes Wahrscheinlichkeitsnetz sollte hingegen eher einen negativen Einfluss haben, da zwar der Inferenzgrad identisch bleibt (es muss nach wie vor lediglich die korrekte Zahl identifiziert werden), sich aber die Anzahl zusätzlich abgebildeter (und für die Frage irrelevanter) Wahrscheinlichkeiten erhöht. Hierdurch wird nicht nur die kognitive Belastung erhöht, sondern durch die zusätzliche Abbildung der vier Schnittwahrscheinlichkeiten werden darüber hinaus typische Falschantworten zur Wahl angeboten.

4 Methode

4.1 Stichprobe

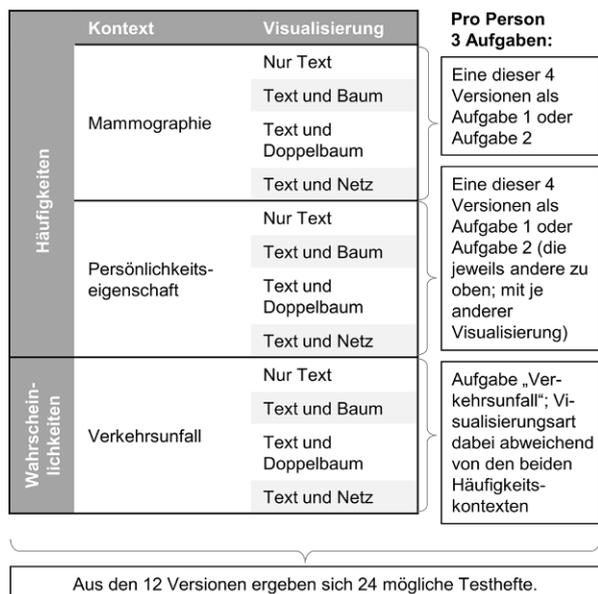
Eine Gelegenheitsstichprobe von $N=269$ Schülerinnen und Schülern aus 13 verschiedenen 10. Klassen von vier bayerischen Gymnasien wurde im Sommer 2018 mit dem Instrument (siehe Abschn. 4.3) untersucht. Die Erhebungen fanden jeweils im Klassenverband statt. Das Alter der Schülerinnen und Schüler lag zwischen 15 und 17 Jahren. Die Versuchspersonen wurden zunächst darüber informiert, dass die Teilnahme an der Studie freiwillig ist und Anonymität zugesichert wird. Alle Schülerinnen und Schüler gaben ihre schriftliche Einwilligung zur Teilnahme an der Studie.

4.2 Design der Studie

Das Design der Studie ist in Abb. 7 zusammengefasst. Alle Versuchspersonen bearbeiteten in einem Papier-und-Bleistift-Test insgesamt drei Bayesianische Aufgaben (als reine Textaufgabe oder zusätzlich mit vollständig ausgefüllten Knoten-Ast-Strukturen). Zunächst hatten die Schülerinnen und Schüler zwei Fragen im Häufigkeitsformat zu den beiden Kontexten „Mammographie“ und „Persönlichkeitseigenschaft“ und anschließend eine Frage im Wahrscheinlichkeitsformat zum Kontext „Verkehrsunfall“ zu beantworten. Die drei Kontexte sind etabliert und wurden bereits in früheren Studien eingesetzt. Die beiden Kontexte „Mammographie“ und „Persönlichkeitseigenschaft“ dienen der wechselseitigen Validierung etwaiger Befunde im Häufigkeitsformat. Die konkreten Aufgabenformulierungen können Tab. 2 und 3 entnommen werden. Alle Aufgaben wurden dabei entweder rein textuell präsentiert (in Aufgabe 1 und 2 mit natürlichen Häufigkeiten und in Aufgabe 3 mit Wahrscheinlichkeiten) oder zusätzlich zum Aufgabentext mithilfe einer vollständig ausgefüllten Knoten-Ast-Struktur (im selben Informationsformat wie der Text). Den Schülerinnen und Schülern bislang unbekannt Visualisierungen wie der Doppelbaum oder das Netz wurden dabei vorher nicht erläutert.

Aus den vier Visualisierungsarten (inklusive „nur Text“) und drei verschiedenen Kontexten ergeben sich somit 12 verschiedene Aufgabenversionen, die nach folgendem Schema zu 24 verschiedenen Testheften mit systematisch variiertes Reihenfolge der Visualisierung kombiniert wurden (siehe Abb. 7): Die Hälfte der Versuchspersonen erhielt als erste Aufgabe den Kontext „Mammographie“, die andere Hälfte den Kontext „Persönlichkeitseigenschaft“. Als zweite Aufgabe wurde dann der jeweils andere Kontext mit abweichender Visualisierungsart präsentiert. Die dritte Aufgabe, die im Wahrscheinlichkeitsformat zum Kontext „Verkehrsunfall“ gestellt wurde,

Abb. 7 Design der Studie



enthielt schließlich eine der beiden verbleibenden Visualisierungen, die die jeweilige Versuchsperson noch nicht in Aufgabe 1 und 2 gesehen hatte. Die 24 verschiedenen Testhefte wurden zufällig auf die gesamte Stichprobe verteilt. Dabei wurde insbesondere darauf geachtet, dass jeweils zwei Personen, die nebeneinander saßen (und durch Sichtschutz getrennt wurden) die Kontexte von Aufgabe 1 und 2 permutiert erhielten. Der Aufwand der Erstellung von 24 verschiedenen Testheften dient einer maximal randomisierten Zuordnung von Aufgabenversionen (d. h. verschiedene Kontexte, verschiedene Visualisierung) auf Versuchspersonen, um potenzielle Einflussgrößen wie Zugehörigkeit zu einer bestimmten Klasse, Zugehörigkeit zu einer bestimmten Schule oder der Note in Mathematik im Zwischenzeugnis möglichst gut kontrollieren zu können.

Insgesamt wurde jede der 12 Aufgabenversionen von etwa 67 Versuchspersonen bearbeitet. Die Durchführung der Studie erfolgte durch einen geschulten Testleiter. Für die Bearbeitung der Aufgaben gab es keinerlei zeitliche Beschränkung. Die Schülerinnen und Schüler benötigten im Schnitt 30 min zur Komplettierung aller Aufgaben. Um bei der Auswertung etwaige Reihenfolgeeffekte kontrollieren zu können, durfte jede Versuchsperson die Bearbeitung von Aufgabe 2 erst beginnen, wenn Aufgabe 1 eingesammelt wurde (gleiches gilt für Aufgabe 3).

Die Schülerinnen und Schüler hatten zur Bearbeitung der Aufgaben Taschenrechner zur Verfügung, die sie zu jedem Zeitpunkt der Erhebung nutzen durften.

4.3 Instrument

Tab. 2 und 3 geben einen Überblick über die drei verwendeten Textversionen (die unabhängig von der Visualisierung in jedem Fall präsentiert wurden). Die Versuchspersonen bearbeiteten Aufgaben entweder als reine Textversion oder als Textversion, die zusätzlich eine bereits ausgefüllte Visualisierung beinhaltet. Die in der Studie eingesetzten Visualisierungen entsprechen (außer der Farbgebung) den Häufigkeitsvisualisierungen in Abb. 2 und den Wahrscheinlichkeitsvisualisierungen in Abb. 3. Die Originalversionen der Testbögen inklusive der implementierten Visualisierungen können im Onlinematerial 3 eingesehen werden.

Aufgrund der unterschiedlichen Kontexte für Häufigkeits- und Wahrscheinlichkeitsformat (*Mammographie* und *Persönlichkeitseigenschaft* bei Häufigkeiten und *Verkehrsunfall* bei Wahrscheinlichkeiten) ist ein direkter Vergleich zwischen den beiden Informationsformaten weder intendiert noch valide möglich. Ziel der Wahrscheinlichkeitsaufgabe war es lediglich herauszufinden, in welchem Maße Schülerinnen und Schülern die Beantwortung einer Bayesianischen Frage aus Wahrscheinlichkeitsvisualisierungen gelingt.

4.4 Kovariaten

Von den Schülerinnen und Schülern wurde als Kontrollvariable zusätzlich die Mathematiknote des letzten Zwischenzeugnisses erhoben. Darüber hinaus liegen Informationen über die Zugehörigkeit der einzelnen Schülerinnen und Schüler zu bestimmten Klassen sowie der jeweiligen Schule vor. Außerdem wird in den späteren statistischen Analysen die Reihenfolge der beiden Häufigkeitsaufgaben modelliert –

Tab. 2 Häufigkeitsversionen zur Untersuchung der Performanz zu den beiden Kontexten Mammographie und Persönlichkeitseigenschaft

	Kontext: Mammographie	Kontext: Persönlichkeitseigenschaft
Einleitung	Stellen Sie sich bitte vor, Sie sind Reporter/Reporterin einer Frauenzeitschrift und möchten einen Artikel über Brustkrebs schreiben. Sie recherchieren auch darüber, was von den Tests zu halten ist, die im Rahmen von Routine-untersuchungen eingesetzt werden, um Brustkrebs zu entdecken. Ihr besonderes Interesse gilt der Frage, was es bedeutet, wenn eine Frau bei einem solchen Test ein positives Ergebnis (welches Brustkrebs anzeigt) erhält	Stellen Sie sich bitte vor, Sie interessieren sich für die Frage, ob karriere-orientierte Schüler eher den Wirtschaftskurs wählen. Der schulpsychologische Dienst untersucht daher für Sie den Zusammenhang von Persönlichkeitseigenschaften bei Schülern und der Wahl von Unterrichtsfächern
Statistische Informationen (auch in den Versionen mit Visualisierung gezeigt)	<ul style="list-style-type: none"> – 100 von 10.000 Frauen, die zu einer Routineuntersuchung gehen, haben Brustkrebs – Von 100 Frauen, die zu einer Routineuntersuchung gehen und die Brustkrebs haben, erhalten 80 ein positives Testergebnis – Von 9900 Frauen, die zu einer Routineuntersuchung gehen und die keinen Brustkrebs haben, erhalten 950 dennoch ein positives Testergebnis 	<ul style="list-style-type: none"> – 325 von 1000 Schülern wählen den Wirtschaftskurs – Von 325 Schülern, die den Wirtschaftskurs wählen, sind 208 karriere-orientiert – Von 675 Schülern, die nicht den Wirtschaftskurs wählen, sind 405 dennoch karriere-orientiert
Visualisierungen	<ul style="list-style-type: none"> – Keine zusätzliche Visualisierung, oder – Häufigkeitsbaum (Abb. 2 oben), oder – Häufigkeitsdoppelbaum (Abb. 2 Mitte), oder – Häufigkeitsnetz (Abb. 2 unten) 	<ul style="list-style-type: none"> – Keine zusätzliche Visualisierung, oder – Häufigkeitsbaum, oder – Häufigkeitsdoppelbaum, oder – Häufigkeitsnetz
Frage	Wie viele der Frauen, die zu einer Routineuntersuchung gehen und ein positives Testergebnis erhalten, haben Brustkrebs? Antwort: _____ von _____	Wie viele der Schüler, die karriereorientiert sind, wählen den Wirtschaftskurs? Antwort: _____ von _____

also ob eine Häufigkeitsaufgabe als erste Aufgabe bearbeitet wurde oder als zweite Aufgabe. Dies ermöglicht eine Kontrolle etwaiger Lern- oder Ermüdungseffekte.

4.5 Vorwissen der Schülerinnen und Schüler

Die Lehrkräfte aller Schülerinnen und Schüler gaben an, dass – wie vermutet – den Schülerinnen und Schülern Vierfeldertafeln mit Häufigkeiten, Vierfeldertafeln mit Wahrscheinlichkeiten und Baumdiagramme mit Wahrscheinlichkeiten bereits bekannt waren. Diese wurden in vorherigen Schuljahren thematisiert und auch von den Versuchspersonen bereits eigenständig erstellt. Doppelbäume und Netze waren den Schülerinnen und Schülern dagegen unbekannt. Der Begriff der *bedingten Wahrscheinlichkeit* sowie die 1. und 2. Pfadregel war den Schülerinnen und Schülern zum Zeitpunkt der Erhebung aus dem vorherigen Schuljahr bekannt. Bayesianische Aufgaben wurden aber noch nicht gelöst.

Keine der Visualisierungen wurde während der Durchführung der Studie erläutert. Im Sinne der Forschungsfrage soll überprüft werden, ob der ausgefüllte Doppelbaum

Tab. 3 Wahrscheinlichkeitsversionen zur Untersuchung der Performanz zum Kontext Verkehrsunfall

Kontext: Verkehrsunfall	
Einleitung	Stellen Sie sich bitte vor, Sie sehen, wie sich eine Person nach einer Feier betrunken ans Steuer ihres Autos setzt. Sie interessiert nun besonders das Risiko eines von dieser Person verursachten Verkehrsunfalls. Im Internet finden Sie diese Informationen:
Statistische Informationen (auch in den Versionen mit Visualisierung gezeigt)	<ul style="list-style-type: none"> – Die Wahrscheinlichkeit, dass ein Autofahrer einen Unfall verursacht, beträgt 1 % – Wenn ein Autofahrer einen Unfall verursacht, so beträgt die Wahrscheinlichkeit, dass dieser dabei betrunken ist, 55 % – Wenn ein Autofahrer keinen Unfall verursacht, beträgt die Wahrscheinlichkeit, dass er dennoch betrunken ist, 5 %
Visualisierung	<ul style="list-style-type: none"> – Keine zusätzliche Visualisierung, oder – Wahrscheinlichkeitsbaum, oder – Wahrscheinlichkeitsdoppelbaum, oder – Wahrscheinlichkeitsnetz
Frage	Wie groß ist die Wahrscheinlichkeit, dass ein Autofahrer einen Unfall verursacht, wenn er betrunken ist? Antwort: _____

oder das ausgefüllte Netz bereits ohne vorangegangene Erklärung richtig interpretiert werden.

4.6 Korrekte Lösungen und Kodierung der Antworten

Die korrekte Lösung lautet für die Aufgabe im *Häufigkeitsformat* beim Kontext Mammographie „80 von 1030“ und beim Kontext Persönlichkeitseigenschaft „208 von 613“. Eine Antwort wurde jeweils nur dann als korrekt gewertet, wenn tatsächlich beide Zahlen (z. B. 80 und 1030) richtig angegeben waren.

Die korrekte Lösung für die Aufgabe im *Wahrscheinlichkeitsformat* zum Kontext Verkehrsunfall lautet 10 %. Aufgrund möglicher Rundungsfehler wurde allerdings jede Antwort als korrekt gelöst kategorisiert, die zwischen 9 % und 11 % lag. Dabei wurde auch jede dazu äquivalente Angabe als gewöhnlicher Bruch oder Dezimalbruch als richtig gewertet – andere numerische Darstellungsarten wurden von den Versuchspersonen nicht als Antwort gegeben.

4.7 Datensatz

Der Datensatz kann in Onlinematerial 4 eingesehen werden.

5 Ergebnisse

5.1 Performanz bei sukzessiver Erweiterung der Häufigkeitsvisualisierungen

Abb. 8 zeigt den Anteil korrekter Antworten der Schülerinnen und Schüler für die verschiedenen Häufigkeitsversionen der Bayesianischen Aufgabe. Die sukzessive Erweiterung der Knoten-Ast-Struktur führt deskriptiv zu zunehmend höheren Lö-

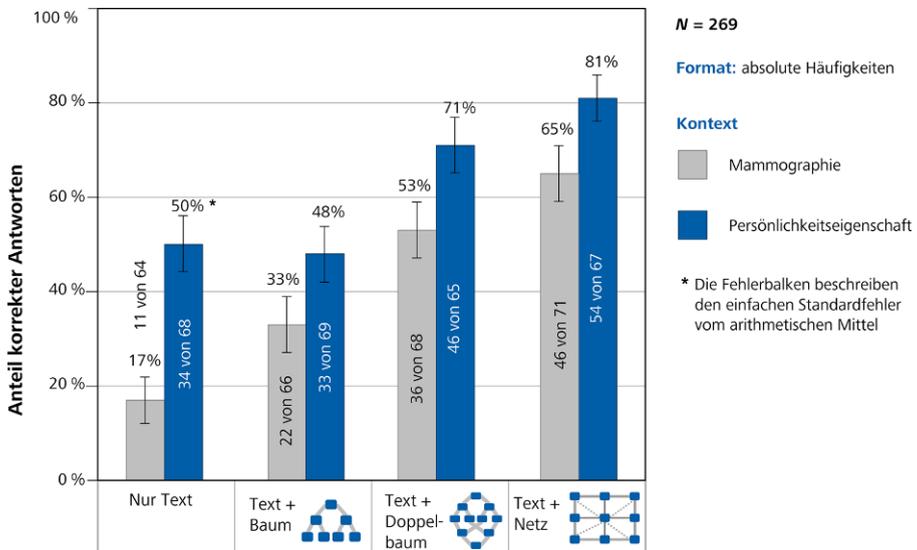


Abb. 8 Anteil korrekter Antworten bei Häufigkeitsaufgaben (getrennt nach Visualisierung und Kontext)

sungsraten. Während über beide Kontexte gemittelt die reine Textvariante von 34 % der Schülerinnen und Schüler richtig gelöst wird, liegt die Lösungsrate beim Häufigkeitsbaum bei 41 % (ebenfalls über beide Kontexte gemittelt). Außerdem führt die Erweiterung von Häufigkeitsbaum zu Häufigkeitsdoppelbaum auf deskriptiver Ebene zu einer erhöhten Performanz (62 % über beide Kontexte gemittelt). Weiterhin wird die Aufgabe für beide Kontexte dann am besten bearbeitet, wenn zusätzlich zum Text mit den statistischen Informationen das Häufigkeitsnetz präsentiert wurde (72 % über beide Kontexte gemittelt), obschon dieses den Lernenden unbekannt war und ihnen auch nicht erläutert wurde.

Um den Effekt der Knoten-Ast-Struktur-Erweiterung auf die Performanz inferenzstatistisch zu überprüfen, wurde ein Verallgemeinertes Lineares Gemischtes Modell (GLMM, General Linear Mixed Model) mit einer logistischen Link-Funktion gerechnet. Wie auch binäre logistische Regressionsanalysen eignen sich GLMM zur Modellierung der Verteilung abhängiger diskreter Variablen. Die abhängige Variable im vorliegenden Modell ist die Korrektheit der Lösung (0: falsch, 1: richtig). Im Gegensatz zu binären logistischen Regressionsanalysen können in GLMM auch Zufallsfaktoren mitmodelliert werden – im vorliegenden Fall die Zugehörigkeit zu einer bestimmten Schule sowie Klasse als auch der Fehler, der typisch für die Versuchsperson ist (diese Zufallsfaktoren können mitmodelliert werden, da jede Schülerin und jeder Schüler zwei Häufigkeitsaufgaben bearbeitet hat). Im vorliegenden GLMM wurde jeweils die vorherige Knoten-Ast-Struktur als Referenz verwendet, so dass additive Veränderungen von einer zur jeweils nächsten Struktur abgebildet werden können. Als Prädiktoren wurden demzufolge die Erweiterung auf den Häufigkeitsbaum (d.h. zusätzlich zum Text), die Erweiterung auf den Häufigkeitsdoppelbaum und die Erweiterung auf das Häufigkeitsnetz modelliert. Als Kontrollvariablen wur-

den der Kontext der Aufgabe, die Mathematiknote und die Reihenfolge der Aufgaben implementiert.

Während die Erweiterung auf das Baumdiagramm kein signifikanter Prädiktor zur Vorhersage der Lösewahrscheinlichkeit ist (unstandardisierter Regressionskoeffizient $\beta_1 = 0,40$; $SE = 0,38$; $z = 1,06$; $p = 0,29$), erhöht die Erweiterung von Häufigkeitsbaum auf Häufigkeitsdoppelbaum die Performanz signifikant (unstandardisierter Regressionskoeffizient $\beta_2 = 1,64$; $SE = 0,42$; $z = 3,93$; $p < 0,001$). Eine Erweiterung der Knoten-Ast-Struktur von Häufigkeitsdoppelbaum zum Häufigkeitsnetz führt zu keiner signifikanten Erhöhung der Performanz der Versuchspersonen (unstandardisierter Regressionskoeffizient $\beta_3 = 0,55$; $SE = 0,38$; $z = 1,46$; $p = 0,14$).

Signifikante Prädiktoren für die Lösewahrscheinlichkeit sind weiterhin die Mathematiknote (bessere Schülerinnen und Schüler lösten die Aufgabe häufiger richtig), der Kontext der Aufgabe (die Aufgabe zum Kontext Persönlichkeitseigenschaft wurde häufiger richtig gelöst als zum Kontext Mammographie) als auch die Reihenfolge der Aufgabe (wie bereits in früheren Studien wurde Aufgabe 1 – unabhängig von der Visualisierung und dem Kontext – häufiger korrekt gelöst als wenn die identische Aufgabe erst als Aufgabe 2 gestellt wurde, siehe z. B. Binder et al. 2018a).

Die Verbesserung der Performanz von Häufigkeitsbaum auf Häufigkeitsdoppelbaum ist erwartungskonform, da nur in diesem Vergleich eine Erweiterung der Knoten-Ast-Struktur auch mit einer Reduktion des Inferenzgrads einhergeht (die direkte Identifikation der gesuchten Informationen im Häufigkeitsdoppelbaum ist möglich). Die deskriptive (aber nicht signifikante) Verbesserung der Performanz von Häufigkeitsdoppelbaum zu Häufigkeitsnetz widerspricht hingegen den Erwartungen, da hier die Knoten-Ast-Struktur durch das Einfügen vier zusätzlicher Äste keine weiteren für die Aufgabe relevanten Informationen präsentiert werden, sondern sich lediglich die Komplexität der Visualisierung erhöht und mögliche Falschantworten angeboten werden (vgl. Binder et al. 2020b).

Der systematische Unterschied in der Performanz hinsichtlich der beiden Kontexte wurde bereits in früheren Studien beobachtet und ist daher erwartungskonform (siehe z. B. Siegrist und Keller 2011). Mögliche Gründe für diesen Unterschied könnten sein (vgl. Binder et al. 2015): 1. Die erhöhte sprachliche Komplexität des Kontexts Mammographie; 2. Die Kontraintuitivität des Ausdrucks „positives Testergebnis“, das für ein negatives Ereignis im Sinne einer ernstzunehmenden Diagnose steht; 3. Der nähere Lebensweltbezug beim Kontext zu Persönlichkeitseigenschaften, oder 4. Die Tatsache, dass das Ergebnis beim Mammographie-Problem überrascht und so einen kognitiven Konflikt erzeugt. Allerdings zeigt sich für beide Kontexte prinzipiell das gleiche Bild: Eine sukzessive Erweiterung der bereits ausgefüllten Knoten-Ast-Struktur bewirkt eine höhere Lösungsrate.

5.2 Performanz der Schülerinnen und Schüler bei Wahrscheinlichkeitsvisualisierungen

Bezüglich der Bayesianischen Aufgabe im Wahrscheinlichkeitsformat gibt es drei interessante Ergebnisse, die Abb. 9 entnommen werden können: Zunächst wurde das Ergebnis repliziert, dass Schülerinnen und Schüler nur sehr selten Bayesianische Aufgaben korrekt lösen, wenn die statistischen Informationen in Wahrscheinlich-

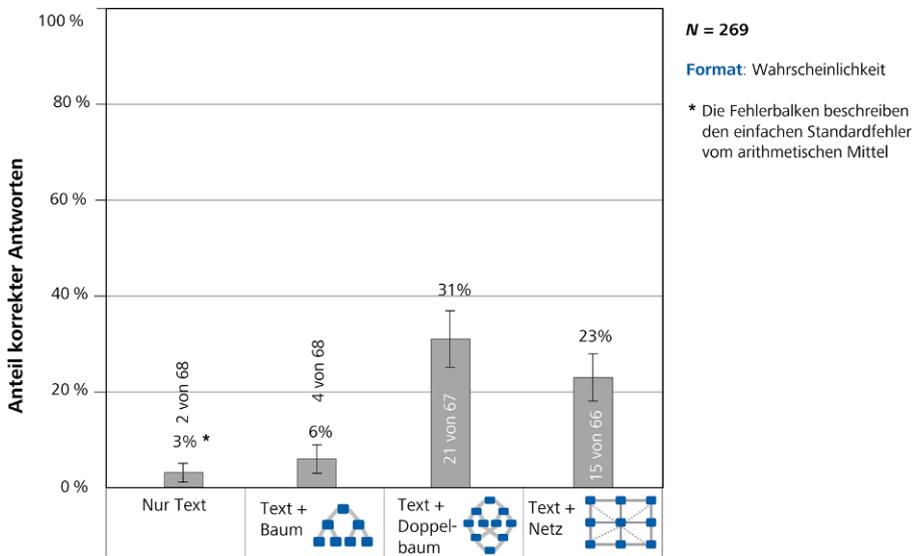


Abb. 9 Anteil korrekter Antworten bei der Aufgabe im Wahrscheinlichkeitsformat (getrennt nach Visualisierung)

keiten präsentiert werden (Anteil korrekter Lösungen ohne Visualisierung 3 % bzw. 6 % mit Wahrscheinlichkeitsbaum). Die Erweiterung der Knoten-Ast-Struktur von einem Wahrscheinlichkeitsbaum zu einem Wahrscheinlichkeitsdoppelbaum führt jedoch deskriptiv zu einer Erhöhung der Performanz (Anteil korrekter Lösungen mit dem Wahrscheinlichkeitsdoppelbaum 31 %). Darüber hinaus fällt auf, dass eine zusätzliche Erweiterung der Knoten-Ast-Struktur von Wahrscheinlichkeitsdoppelbaum zum Wahrscheinlichkeitsnetz wieder zu einer deskriptiv *niedrigeren* Lösungsrate führt (Anteil korrekter Lösungen 23 %). Man beachte, dass die Lösungsrate selbst in den beiden Varianten mit Wahrscheinlichkeitsdoppelbaum und Wahrscheinlichkeitsnetz mit maximal 31 % noch erstaunlich niedrig ist, obwohl in beiden Fällen gar keine echte Inferenz erforderlich ist, sondern die korrekte Lösung lediglich identifiziert werden muss.

Um den Effekt der sukzessiven Erweiterung der Knoten-Ast-Strukturen in der Wahrscheinlichkeitsversion auf die Lösewahrscheinlichkeit inferenzstatistisch zu überprüfen, wurde erneut ein Verallgemeinertes Lineares Gemischtes Modell (GLMM) mit einer logistischen Link-Funktion gerechnet. Die vorhergehende Knoten-Ast-Struktur wurde dabei jeweils wieder als Referenz verwendet und als Prädiktoren die Erweiterung auf das zusätzliche Baumdiagramm, die Erweiterung vom Baum auf den Doppelbaum und die Erweiterung von Doppelbaum hin zum Netz zusammen mit der Kontrollvariable Mathematiknote und den Zufallsfaktoren Zugehörigkeit zu einer bestimmten Schule sowie Klasse implementiert. Während die Erweiterung zum Baumdiagramm kein signifikanter Prädiktor ist (unstandardisierter Regressionskoeffizient $\beta_1 = 0,80$; $SE = 0,89$; $z = 0,89$; $p = 0,37$), führt die Erweiterung von Baum zu Doppelbaum zu einer signifikanten Verbesserung der Lösewahrscheinlichkeit (unstandardisierter Regressionskoeffizient $\beta_2 = 2,39$; $SE = 0,62$;

$z = 3,83$; $p < 0,001$). Die Erweiterung vom Doppelbaum zum Netz hingegen führt wiederum zu einer nicht signifikanten Reduktion der Lösungsrate (unstandardisierter Regressionskoeffizient $\beta_3 = -0,45$; $SE = 0,43$; $z = -1,04$; $p = 0,30$). Auch bei der Wahrscheinlichkeitsversion hat eine bessere Note in Mathematik einen positiven Einfluss auf die Lösefähigkeit.

In der Wahrscheinlichkeitsversion entspricht die Verbesserung der Performanz von Baumdiagramm auf Doppelbaum ebenfalls den Erwartungen, da eine Erweiterung der Knoten-Ast-Struktur eine vereinfachte Lösung bedeutet (die direkte Identifikation ist ab dem Doppelbaum möglich). Der (nicht-signifikante) Unterschied in der Performanz von Doppelbaum zu Netzdiagramm ist ebenfalls erwartungskonform: Es werden vier zusätzliche Äste mit für die Lösung nicht relevanten Schnittwahrscheinlichkeiten präsentiert. Die extrinsische Belastung steigt bei unverändertem Inferenzgrad, was zumindest deskriptiv in den Wahrscheinlichkeitsversionen zu einer Reduktion der Performanz der Versuchspersonen führt.

6 Diskussion

Der vorliegende Beitrag stellt dar, wie die schulüblichen (Einfach-)Baumdiagramme – und zwar auf der Basis beider Informationsformate – sukzessive zu Doppelbäumen und zu Netzen erweitert werden können. In einer empirischen Studie wurde untersucht, inwieweit bei Bayesianischen Aufgabenstellungen die zusätzliche Präsentation solcher vollständig ausgefüllten Knoten-Ast-Strukturen die Performanz von Schülerinnen und Schülern ($N = 269$) erhöhen kann. Dabei waren zwei gegenläufige Hypothesen denkbar: Zum einen werden durch die sukzessiven Erweiterungen immer mehr statistische Informationen präsentiert, was zur Folge hat, dass sich zunehmend auch Lösungen zu den verschiedensten Aufgabenstellungen unter diesen Zahlen befinden. Dies führt konsequenterweise zu einer Reduzierung des Inferenzgrads in der Schwierigkeitstaxonomie für ausgefüllte Visualisierungen bei Bayesianischen Aufgaben (*Stufe 1*: Identifizieren der relevanten Informationen, *Stufe 2*: einfache additive Verknüpfung, *Stufe 3*: genuine Inferenzbildung), was eine Steigerung der Performanzen bewirken könnte. Zum anderen nimmt bei den sukzessiven Erweiterungen aber auch die Zahl der für die Lösung irrelevanten Informationen zu, was nicht nur die extrinsische Belastung erhöht (Chandler und Sweller 1991), sondern die Präsentation typischer Falschantworten mit sich bringt. Wäre letzterer Aspekt der dominierende, wäre bei Steigerung der Komplexität eine Reduktion der Performanz zu erwarten.

Bei den Knoten-Ast-Strukturen im Häufigkeitsformat zeigt sich die einzige signifikante Zunahme der Lösungswahrscheinlichkeit bei der Erweiterung vom Häufigkeitsbaum zum Häufigkeitsdoppelbaum. Dies ist dadurch zu erklären, dass die beiden absoluten Häufigkeiten, die zur richtigen Lösung ins Verhältnis gesetzt werden müssen, erst ab hier mit einem Ast verbunden sind und so keine Addition von zwei Häufigkeiten mehr nötig ist, sondern ein einfaches Identifizieren der beiden relevanten Häufigkeiten ausreicht (*Stufe 1*). Obwohl den Schülerinnen und Schülern weder der Häufigkeitsdoppelbaum noch das Häufigkeitsnetz bekannt waren, schnitten sie bei diesen beiden Visualisierungen am besten und etwa gleich gut ab.

Dies ist in Bezug auf das Netz besonders hervorzuheben, da dieses durch eine völlig neue Anordnung (im Vergleich zu Baumdiagrammen) in der Tat sehr komplex wirkt. Befürchtungen im Vorfeld, dass dies eventuell zu einer Erhöhung der extrinsischen kognitiven Belastung führen könnte (z. B. Henze und Vehling 2021) und darüber hinaus die vier neuen (unbeschrifteten) Äste gegebenenfalls sogar zu einer falschen Antwort im Sinne der zugehörigen Schnittwahrscheinlichkeit verleiten könnten, sind somit – zumindest für bereits ausgefüllte Diagramme – unbegründet.

Für Knoten-Ast-Strukturen im Wahrscheinlichkeitsformat werden diese Ergebnisse gespiegelt: Auch hier zeigt sich ein signifikanter Performanzzuwachs ausschließlich vom Wahrscheinlichkeitsbaum zum Wahrscheinlichkeitsdoppelbaum, was wieder dadurch zu erklären ist, dass erst ab dem Doppelbaum die Wahrscheinlichkeitsantwort auf eine Bayesianische Aufgabenstellung direkt identifiziert und abgelesen werden kann (Stufe 1). Auch wenn sich hier ebenfalls keine signifikante Reduktion bei der Erweiterung zum Wahrscheinlichkeitsnetz ergibt, geht die Tendenz im Wahrscheinlichkeitsformat in die andere Richtung (deskriptiv fällt es Schülerinnen und Schülern nun etwas schwerer als im Wahrscheinlichkeitsdoppelbaum, die richtige Antwort zu identifizieren). Dennoch sind der bereits vollständig ausgefüllte Wahrscheinlichkeitsdoppelbaum und das ausgefüllte Wahrscheinlichkeitsnetz dem in der Schule regelmäßig verwendeten Wahrscheinlichkeits(einfach-)baum in der vorliegenden Studie überlegen. Allerdings scheitern trotz Wahrscheinlichkeitsdoppelbaum bzw. Wahrscheinlichkeitsnetz 69 % bzw. 77 % der Schülerinnen und Schüler an der Aufgabe, obwohl hier ein einfaches Herauslesen der passenden bedingten Wahrscheinlichkeit genügt hätte. Der mögliche Einwand, das Erfassen der im Diagramm dargestellten Situation und die anschließende Identifikation der gesuchten Zahlen sei für Schülerinnen und Schüler immer leicht – auch in Wahrscheinlichkeitsversionen – ist somit nicht zutreffend.

In der Summe kann festgehalten werden, dass vor allem die Reduktion des Inferenzgrads (in Anlehnung an Curcio 1987) eine Erklärung für die gefundenen Performanzunterschiede liefern kann. Signifikante Hinweise auf eine zu hohe extrinsische kognitive Belastung im Hinblick auf den ausgefüllten Doppelbaum oder das ausgefüllte Netz können in den Daten dagegen nicht gefunden werden.

6.1 Knoten-Ast-Strukturen als Brücke zwischen der Welt der Wahrscheinlichkeiten und der Welt der natürlichen Häufigkeiten

Die Knoten-Ast-Strukturen Baumdiagramm, Doppelbaum und Netz haben gemeinsam, dass Wahrscheinlichkeiten an den Ästen und absolute Häufigkeiten in den Knoten der Strukturen notiert werden können. Durch diese Art der Darstellung tritt die Beziehung zwischen absoluten Häufigkeiten und Wahrscheinlichkeiten besonders in den Vordergrund. Inwiefern dieses Potenzial von Knoten-Ast-Strukturen als Brücke zwischen den verschiedenen Darstellungsarten von Anteilswerten und Wahrscheinlichkeiten fungieren kann, sollte Grundlage zukünftiger Forschung sein. Möglicherweise eignen sich Knoten-Ast-Strukturen besonders gut, um Schülerinnen und Schülern die Verbindung zwischen den verschiedenen Darstellungsarten zu verdeutlichen. Beispielsweise könnten Versuchspersonen mit Netzen konfrontiert werden, die beide Informationsformate (Wahrscheinlichkeiten *und* Häufigkeiten) gleichzeitig

enthalten und untersucht werden, wie diese Art der Darstellung von Versuchspersonen gelesen wird. In diesem Zusammenhang können in künftigen Studien auch verschiedene Design-Prinzipien bei der Diagrammerstellung untersucht werden, z. B. in welche Richtung die Pfeile im Doppelbaum oder Netz zeigen sollten.

6.2 Limitationen und Ausblick

In der vorliegenden Studie wurden vollständig ausgefüllte Visualisierungen gezeigt, sodass weiterhin ungeklärt bleibt, ob nicht bei der eigenständigen Erstellung der jeweiligen Visualisierungen die kognitive Belastung höher als die didaktische Unterstützung ist. Da sich der Doppelbaum im Unterricht leichter als sequentielle Erweiterung des Baumdiagramms einführen lässt und keine Neusortierung der einzelnen Elemente des Diagramms erforderlich ist, könnte in einer entsprechenden Trainingsstudie der Doppelbaum dem Netz überlegen sein. Ein erster Ansatz wie bei einer Bayesianischen Aufgabe das selbstständige Erstellen eines Häufigkeitsnetzes bis hin zur Lösungsfindung mithilfe des Netzes aussehen kann, findet sich in (Binder et al. 2020a), für ein entsprechendes Trainingsprogramm für den Doppelbaum siehe Büchter et al. (2022).

In mathematikdidaktischen Untersuchungen wurden in den vergangenen Jahrzehnten viele unterschiedliche (statische und dynamische) Visualisierungen analysiert, die zur Erreichung verschiedener Ziele im Zusammenhang mit (bedingten Wahrscheinlichkeiten) mehr oder weniger geeignet sind: Baumdiagramm, Vierfeldertafel, Doppelbaum, Netz, Einheitsquadrat uvm. (siehe Abschn. 2.3). Auch ein gleichzeitiges Arbeiten mit verschiedenen Visualisierungen ist hierbei denkbar: So könnte unterrichtlich eine Vierfeldertafel und ein Doppelbaum für eine Aufgabe gezeichnet werden und die Querverbindungen der Visualisierungen herausgearbeitet werden. Eine Implementation *aller* oben genannten Visualisierungen und ein Herausarbeiten aller entsprechender Querverbindungen zwischen diesen Visualisierungen ist allerdings im schulischen Stochastikunterricht wenig zielführend. Ein zentrales Forschungsdesiderat wäre demnach eine konkrete unterrichtliche Empfehlung von der Grundschule bis hin zur gymnasialen Oberstufe, die einige der Visualisierungen zielgerichtet auswählt und in einen fließenden curricularen Ablauf einbettet. In Bezug auf das Häufigkeitsnetz fehlen diesbezüglich aber bislang Erkenntnisse, inwiefern Schülerinnen und Schüler entsprechende Diagramme selbstständig anfertigen können. Die Ergebnisse der vorliegenden Studie legen allerdings nahe, dass das Häufigkeitsnetz grundsätzlich Potenzial für den schulischen Stochastikunterricht hat und bei einem voll ausgefüllten Häufigkeitsnetz die Unterstützung und nicht die extrinsische kognitive Belastung überwiegt.

Additional information Die vorliegende Forschungsarbeit ist Teil des übergeordneten Projekts „Bayesian Reasoning“, ein kooperatives Forschungsprojekt der Universitäten Freiburg, Kassel und Regensburg sowie der Pädagogischen Hochschule Heidelberg und der Ludwig-Maximilians-Universität München zur Untersuchung der Unterstützung des Bayesianischen Denkens: <http://www.bayesianreasoning.de/bayes.html>.

Zusatzmaterial online Zusätzliche Informationen sind in der Online-Version dieses Artikels (<https://doi.org/10.1007/s13138-022-00215-9>) enthalten.

Funding Open Access funding enabled and organized by Projekt DEAL.

Open Access Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden.

Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen.

Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Literatur

- Bea, W. (1995). *Stochastisches Denken*. Frankfurt a.M.: Peter Lang.
- Binder, K., Krauss, S., & Bruckmaier, G. (2015). Effects of visualizing statistical information. An empirical study on tree diagrams and 2×2 tables. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2015.01186>.
- Binder, K., Krauss, S., Bruckmaier, G., & Marienhagen, J. (2018a). Visualizing the Bayesian 2-test case. The effect of tree diagrams on medical decision making. *PLoS one*. <https://doi.org/10.1371/journal.pone.0195029>.
- Binder, K., Krauss, S., & Wassner, C. (2018b). Der Häufigkeitsdoppelbaum als didaktisch hilfreiches Werkzeug von der Unterstufe bis zum Abitur. *Stochastik in der Schule*, 38(1), 2–11.
- Binder, K., Krauss, S., & Wassner, C. (2019). Der Häufigkeitsdoppelbaum – Anteilswerte und bedingte Wahrscheinlichkeiten vorteilhaft visualisieren. *mathematik lehren*, 213, 12–17.
- Binder, K., Krauss, S., & Steib, N. (2020a). Bedingte Wahrscheinlichkeiten und Schnittwahrscheinlichkeiten GLEICHZEITIG visualisieren: Das Häufigkeitsnetz. *Stochastik in der Schule*, 40(2), 2–14.
- Binder, K., Krauss, S., & Wiesner, P. (2020b). A new visualization for probabilistic situations containing two binary events: the frequency net. *Frontiers in psychology*. <https://doi.org/10.3389/fpsyg.2020.00750>.
- Böcherer-Linder, K., & Eichler, A. (2017). The impact of visualizing nested sets. An empirical study on tree diagrams and unit squares. *Frontiers in psychology*. <https://doi.org/10.3389/fpsyg.2016.02026>.
- Brase, G. L. (2008). Pictorial representations in statistical reasoning. *Applied Cognitive Psychology*, 23(3), 369–381. <https://doi.org/10.1002/acp.1460>.
- Bruckmaier, G., Binder, K., Krauss, S., & Kufner, H.-M. (2019). An eye-tracking study of statistical reasoning with tree diagrams and 2×2 tables. *Frontiers in Psychology*, 10, 303. <https://doi.org/10.3389/fpsyg.2019.00632>.
- Büchter, T., Eichler, A., Steib, N., Binder, K., Böcherer-Linder, K., Krauss, S., & Vogel, M. (2022). How to train novices in Bayesian reasoning. *Mathematics*, 10(9), 1558.
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4), 293–332. https://doi.org/10.1207/s1532690xci0804_2.
- Chandler, P., & Sweller, J. (1992). The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, 62(2), 233–246.
- Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, 58(1), 1–73. [https://doi.org/10.1016/0010-0277\(95\)00664-8](https://doi.org/10.1016/0010-0277(95)00664-8).
- Curcio, F. R. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education*, 18(5), 382. <https://doi.org/10.2307/749086>.

- Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: problems and opportunities. In D. Kahneman, P. Slovic, & A. Tversky (Hrsg.), *Judgment under Uncertainty: Heuristics and Biases* (S. 249–267). New York: Cambridge University Press. <https://doi.org/10.1017/CBO9780511>.
- Eichler, A., Böcherer-Linder, K., & Vogel, M. (2020). Different visualizations cause different strategies when dealing with Bayesian situations. *Frontiers in Psychology, 11*, 1897.
- García-Retamero, R., & Hoffrage, U. (2013). Visual representation of statistical information improves diagnostic inferences in doctors and their patients. *Social Science & Medicine, 83*, 27–33. <https://doi.org/10.1016/j.socscimed.2013.01.034>.
- Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: frequency formats. *Psychological Review, 102*(4), 684–704. <https://doi.org/10.1037/0033295X.102.4.684>.
- Henze, N., & Vehling, R. (2021). Im Vordergrund steht das Problem – oder: warum ein Häufigkeitsnetz? *Stochastik in der Schule, 41*(1), 27–32.
- Hoffrage, U., & Gigerenzer, G. (1998). Using natural frequencies to improve diagnostic inferences. *Academic Medicine, 73*(5), 538–540. <https://doi.org/10.1097/00001888-199805000-00024>.
- Hoffrage, U., Lindsey, S., Hertwig, R., & Gigerenzer, G. (2000). Communicating statistical information. *Science, 290*(5500), 2261–2262. <https://doi.org/10.1126/science.290.5500.2261>.
- Kahneman, D., Slovic, P., & Tversky, A. (Hrsg.). (1982). *Judgment under uncertainty: heuristics and biases*. New York: Cambridge University Press.
- Khan, A., Breslav, S., Glueck, M., & Hornbæk, K. (2015). Benefits of visualization in the mammography problem. *International Journal of Human-Computer Studies, 83*, 94–113. <https://doi.org/10.1016/j.ijhcs.2015.07.001>.
- Kolmogoroff, A. (1933). Die elementare Wahrscheinlichkeitsrechnung. In A. Kolmogoroff (Hrsg.), *Grundbegriffe der Wahrscheinlichkeitsrechnung* (S. 1–13). Berlin, Heidelberg: Springer.
- Krauss, S., Weber, P., Binder, K., & Bruckmaier, G. (2020). Natürliche Häufigkeiten als numerische Darstellungsart von Anteilen und Unsicherheit – Forschungsdesiderate und einige Antworten. *Journal für Mathematik-Didaktik, 38*(1), 407. <https://doi.org/10.1007/s13138-019-00156-w>.
- Kultusministerkonferenz (2022). *Bildungsstandards für das Fach Mathematik Erster Schulabschluss (ESA) und Mittlerer Schulabschluss (MSA)*. Bonn: KMK.
- Kunzelmann, A. K., Binder, K., Fischer, M. R., Reincke, M., Braun, L. T., & Schmidmaier, R. (2022). Improving diagnostic efficiency with frequency double-trees and frequency nets in Bayesian reasoning. *MDM Policy & Practice, 7*(1), 238146832210866. <https://doi.org/10.1177/23814683221086623>.
- McDowell, M., & Jacobs, P. (2017). Meta-analysis of the effect of natural frequencies on Bayesian reasoning. *Psychological bulletin, 143*, 1273–1312. <https://doi.org/10.1037/bul0000126>.
- McDowell, M., Galesic, M., & Gigerenzer, G. (2018). Natural frequencies do foster public understanding of medical tests. Comment on Pighin, Gonzalez, Savadori and Giroto (2016). *Medical decision making: an international journal of the Society for Medical Decision Making. https://doi.org/10.1177/0272989X18754508*.
- Micallef, L., Dragicevic, P., & Fekete, J.-D. (2012). Assessing the effect of visualizations on Bayesian reasoning through crowdsourcing. *IEEE Transactions on Visualization and Computer Graphics, 18*(12), 2536–2545. <https://doi.org/10.1109/TVCG.2012.199>.
- Niegemann, H. M. (2008). *Kompendium multimediales Lernen*. X.media.press. Berlin: Springer.
- Operskalski, J. T., & Barbey, A. K. (2016). Risk literacy in medical decision-making. *Science (New York, N.Y.), 352*(6284), 413–414. <https://doi.org/10.1126/science.aaf7966>.
- Saarländisches Ministerium für Bildung und Kultur (2016). Lehrplan Mathematik Gymnasiale Oberstufe G-Kurs. https://www.saarland.de/dokumente/thema_bildung/LP_Ma_GOS_HP_G-Kurs_2016.pdf. Zugegriffen: 24. Apr. 2019.
- Siegrist, M., & Keller, C. (2011). Natural frequencies and Bayesian reasoning. The impact of formal education and problem context. *Journal of Risk Research, 14*(9), 1039–1055. <https://doi.org/10.1080/13669877.2011.571786>.
- Sirota, M., Kostovičová, L., & Juanchich, M. (2014). The effect of iconicity of visual displays on statistical reasoning. Evidence in favor of the null hypothesis. *Psychonomic bulletin & review, 21*(4), 961–968.
- Sloman, S. A., Over, D., Slovak, L., & Stibel, J. M. (2003). Frequency illusions and other fallacies. *Organizational behavior and human decision processes, 91*(2), 296–309. [https://doi.org/10.1016/S0749-5978\(03\)00021-9](https://doi.org/10.1016/S0749-5978(03)00021-9).
- Soto-Andrade, J. (2019). Missing the (Bayesian) wood for the trees? In J. M. Contreras, M. M. Gea, M. M. López-Martín, & E. Molina-Portillo (Hrsg.), *ctas del Tercer Congreso Internacional Virtual de Educación Estadística*. www.ugr.es/local/fqm126/civeest.html.
- Spiegelhalter, D., Pearson, M., & Short, I. (2011). Visualizing uncertainty about the future. *Science, 333*(6048), 1393–1400. <https://doi.org/10.1126/science.1191181>.

- Staatsinstitut für Schulqualität und Bildungsforschung ISB (2021). *LehrplanPlus Gymnasium Bayern, Mathematik, 11. Klasse*.
- Steckelberg, A., Balgenorth, A., Berger, J., & Mühlhauser, I. (2004). Explaining computation of predictive values: 2 x 2 table versus frequency tree. A randomized controlled trial [ISRCTN74278823]. *BMC medical education*, 4, 13.
- Sweller, J., & Chandler, P. (1991). Evidence for cognitive load theory. *Cognition and Instruction*, 8(4), 351–362. https://doi.org/10.1207/s1532690xci0804_5.
- Wassner, C. (2004). *Förderung Bayesianischen Denkens – Kognitionspsychologische Grundlagen und didaktische Analysen*. Hildesheim: Franzbecker.
- Weber, P., Binder, K., & Krauss, S. (2018). Why can only 24% solve Bayesian reasoning problems in natural frequencies: Frequency phobia in spite of probability blindness. *Frontiers in Psychology*, 9, 1833. <https://doi.org/10.3389/fpsyg.2018.01833/pdf>.
- Yamagishi, K. (2003). Facilitating normative judgments of conditional probability. Frequency or nested sets? *Experimental Psychology*, 50(2), 97–106. <https://doi.org/10.1026//1618-3169.50.2.97>.
- Zikmund-Fisher, B. J., Witteman, H. O., Dickson, M., Fuhrel-Forbis, A., Kahn, V. C., Exe, N. L., Valerio, M., Holtzman, L. G., Scherer, L. D., & Fagerlin, A. (2014). Blocks, ovals, or people? Icon type affects risk perceptions and recall of pictographs. *Medical Decision Making*, 34(4), 443–453. <https://doi.org/10.1177/0272989X13511706>.