

Georg BRUCKMAIER, Basel, Karin BINDER &  
Stefan KRAUSS, Regensburg

## **Strategien beim Lösen statistischer Aufgaben – Eine Eyetracking-Studie zur visuellen Durchmusterung von Baumdiagrammen und Vierfeldertafeln**

### **Einleitung**

Die vorliegende Studie untersucht statistisches Schlussfolgern bei sog. *Bayesianischen Aufgaben*. Damit sind Aufgaben gemeint, bei denen – wie im Fall der Mammographie-Aufgabe – die Prävalenz  $P(K)$  einer Erkrankung mit der Sensitivität  $P_K(T+)$  und der Falsch-Positiv-Rate  $P_{-K}(T+)$  eines Tests kombiniert werden muss, um zu bestimmen, wie wahrscheinlich nach einem positiven Testergebnis  $T+$  nun tatsächlich das Vorliegen der Krankheit  $K$  ist, also  $P_{T+}(K)$ . Derartige Aufgaben und damit einhergehende Entscheidungen sind nicht nur im medizinischen und juristischen Kontext von enormer Bedeutung. Sie können bei geeigneter Einführung auch im schulischen Stochastikunterricht bereits frühzeitig thematisiert werden.

### **Theoretischer Hintergrund**

Wenngleich bereits ausführlich untersucht wurde, welche Faktoren das Verständnis Bayesianischer Aufgaben erleichtern können (u. a. Änderung des Informationsformats, geeignete Visualisierungen; für einen Überblick siehe Binder et al., 2015; McDowell & Jacobs, 2017), sind die zugrundeliegenden und ggf. fehlerhaften Lösungsstrategien (Eichler & Böcherer-Linder, 2018; Gigerenzer & Hoffrage, 1995; Zhu & Gigerenzer, 2006) noch weitestgehend unerforscht. Aus mathematikdidaktischer Perspektive sind diese jedoch interessant, da sie vielversprechende Hinweise liefern können, *wie* und *warum* bestimmte Aufgabenvarianten Schülerinnen und Schülern (nicht) helfen.

Während die Analyse schriftlicher Lösungen und alternative Methoden wie *Lautes Denken* es durchaus ermöglichen, kognitive Prozesse bei der Aufgabenbearbeitung nachzuvollziehen, so bleiben sie doch oft nur oberflächlich oder beeinflussen sogar unmittelbar die Lösungsfindung. Deshalb wurde wiederholt ein breiterer Forschungszugang gefordert, wobei *Eyetracking* wiederholt explizit als Methode genannt wurde (McDowell & Jacobs 2017; Verschaffel et al., 2016). In diesem Sinne setzen wir die Analyse von Blickbewegungen ein, um richtige wie auch falsche Lösungsstrategien beim Bearbeiten von Bayesianischen Aufgaben genauer zu ergründen.

Einige Forschergruppen haben sich bereits ähnlichen Thematiken gewidmet. Beispielsweise untersuchten Lehner und Reiss (2018) an Vierfeldertafeln

mit absoluten Häufigkeiten Zusammenhänge zwischen den Entscheidungsstrategien und den Blickbewegungen ihrer Studienteilnehmer/innen. Sie fanden verschiedene (richtige) Strategien, die unterschiedlich viele der Zellen berücksichtigten, was sich auch in der Anzahl und Dauer der Fixationen auf einzelnen Zellen widerspiegelte. Cohan und Staub (2015) wiederum untersuchten *falsche* Strategien bei Bayesianischen Aufgaben. Sie fanden heraus, dass viele Personen durchgängig bestimmte Strategien anwenden, die fälschlicherweise jedoch nur eine der (typischerweise drei) gegebenen Wahrscheinlichkeiten nutzen (nämlich die Sensitivität). Weiterhin konnten beispielsweise Okan et al. (2016) und Woller-Carter et al. (2012) durch Blickbewegungsanalysen zeigen, dass *Experten* besser in der Lage sind, in Visualisierungen die relevanten Informationen zu finden.

In der vorliegenden Studie geht es darum, bei Bayesianischen Aufgaben (1) verschiedene Informationsformate (Wahrscheinlichkeiten vs. natürliche Häufigkeiten), (2) verschiedene Visualisierungen (Baumdiagramm vs. Vierfeldertafel), und (3) verschiedene Arten von Wahrscheinlichkeiten (Rand-, Schnitt-, bedingte und Bayesianische Wahrscheinlichkeiten bzw. Häufigkeiten) mithilfe der Erfassung von Blickbewegungen zu untersuchen. Daraus ergeben sich insbesondere folgende Fragestellungen:

**F1:** Welche (richtigen und falschen) Strategien können aus den Blickpfaden (*Scanpaths*, und daraus aggregierten *Heatmaps*) der Studienteilnehmer/innen deskriptiv beschrieben werden?

**F2:** Wie beeinflussen das Informationsformat (Wahrscheinlichkeit vs. natürliche Häufigkeit) und die Visualisierung (Baumdiagramm vs. Vierfeldertafel) die Strategien der Studienteilnehmer/innen?

## Methoden

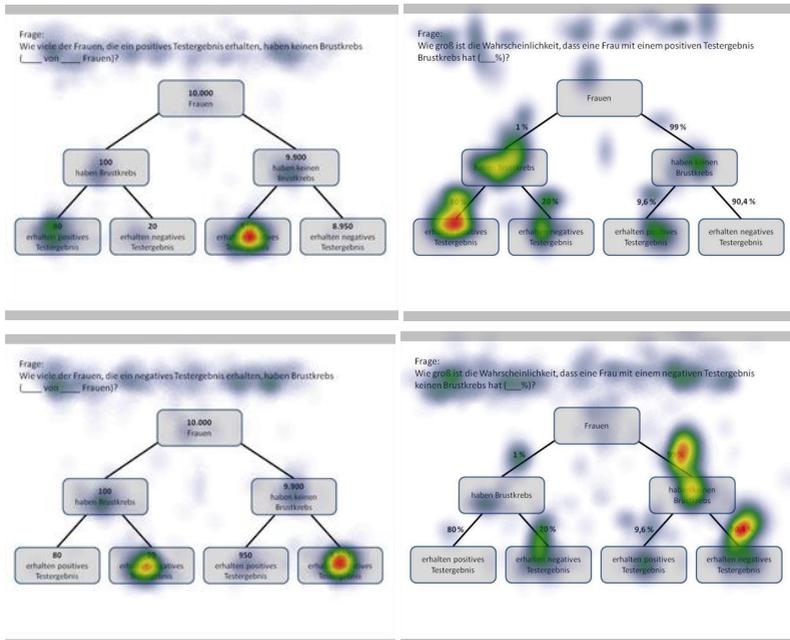
Mit einem mobilen Eyetracker (XRemote 250 mobile von SMI) wurden die Blickbewegungen von 24 Studierenden aufgezeichnet. Diese sollten statistische Aufgaben dabei sowohl möglichst richtig als auch schnell (max. 30 Sek.) lösen. Nach einigen Eingewöhnungsbildern und zwei Beispielaufgaben wurden allen Studienteilnehmer/innen/n insgesamt zwei Mal zehn Aufgaben präsentiert, die jeweils eine Visualisierung (Baumdiagramm *oder* Vierfeldertafel) zusammen mit einer der vier verschiedenen Arten von Wahrscheinlichkeitsfragen (am oberen Bildschirmrand) enthielten (siehe Abbildung 1), jedoch keine zusätzliche textuelle Beschreibung des Problems.

Den Proband/inn/en wurden zuerst zehn Aufgaben mit Baumdiagrammen präsentiert (Kontext: Mammographieproblem), dann zehn Aufgaben mit Vierfeldertafeln (Kontext: Kurswahlproblem) (siehe jeweils Binder, Krauss

& Bruckmaier, 2015). Die jeweils zehn Aufgaben enthielten je fünf Aufgaben in Form von Wahrscheinlichkeiten und fünf (korrespondierende, jedoch nicht identische) Aufgaben in Form von natürlichen Häufigkeiten. Hinsichtlich der Art der Wahrscheinlichkeitsfrage wurden bei beiden Visualisierungen jeweils nach zwei Rand-, zwei Schnitt- und zwei bedingten Wahrscheinlichkeiten bzw. Häufigkeiten gefragt (in randomisierter Reihenfolge) sowie – abschließend – vier Bayesianische Aufgaben gestellt.

### Ergebnisse und Diskussion

Wie nicht anders zu erwarten war, wurden Aufgaben in natürlichen Häufigkeiten deutlich häufiger korrekt gelöst als solche in Wahrscheinlichkeiten (Lösungsrate über alle 20 Aufgaben: 72% zu 47%). Dieses Muster zeigte sich stabil über alle vier Fragearten (Rand-, Schnitt-, bedingte und Bayesianische Wahrscheinlichkeiten bzw. Häufigkeiten), was demnach auch die Anzahl und Art potenzieller Fehlerstrategien beeinflusst. Erste qualitative Auswertungen in Form von *Scanpaths* und *Heatmaps* zeigen bereits interessante Schwerpunkte bei der Durchmusterung der Visualisierungen (siehe Abbildung 1).



**Abb. 1:** Heatmaps über alle vier korrekt gelösten Versionen zu Baumdiagrammen (von links oben nach rechts unten:  $P_{T+}(-K)$ ,  $P_{T+}(K)$ ,  $P_{T-}(K)$ ,  $P_{T-}(-K)$ )

Wie die obigen Beispiele verdeutlichen, zeigen sich systematische qualitative (und quantitative) Unterschiede zwischen den verschiedenen Bayesianischen Aufgaben. Weitere Analysen zeigen, dass verschiedene Fehler (sowie die richtige Lösung) zu ein- und derselben Aufgabenstellung nicht nur in ganz unterschiedlichen Blickpfaden (im Sinne verschiedener Heatmaps) resultieren, sondern sich auch in der Anzahl und Dauer der Fixationen und den sog. *Transitions* zwischen relevanten Regionen (*AOIs*) etc. deutlich unterscheiden. Aus vertieften quantitativen Auswertungen (Bruckmaier, Binder, Krauss & Kufner, eingereicht) lassen sich dabei zum Beispiel Konsequenzen u.a. im Hinblick auf die richtige, didaktisch adäquate Lesart der Visualisierungen ableiten.

## Literatur

- Binder, K., Krauss, S. & Bruckmaier, G. (2015). Effects of visualizing statistical information – An empirical study on tree diagrams and  $2 \times 2$  tables. *Frontiers in Psychology*, 6(1186).
- Bruckmaier, G., Binder, K., Krauss, S. & Kufner, H.-M. (eingereicht). *What eye tracking can tell us on statistical reasoning – An empirical study on tree diagrams and  $2 \times 2$  tables.*
- Cohen, A. L. & Staub, A. (2015). Within-subject consistency and between-subject variability in Bayesian reasoning strategies. *Cognitive Psychology*, 81, 26–47.
- Eichler, A. & Böcherer-Linder, K. (2018). Categorizing errors in Bayesian situations. In M. A. Sorto, A. White & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10)*. Voorburg, The Netherlands: International Statistical Institute.
- Gigerenzer, G. & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, 102(4), 684–704.
- Lehner, M. C. & Reiss, K. (2018). Entscheidungsstrategien an Vierfeldertafeln: Eine Analyse mit Blickbewegungen. *Journal für Mathematik-Didaktik*, 39(1), 147–170.
- McDowell, M. & Jacobs, P. (2017). Meta-Analysis of the Effect of Natural Frequencies on Bayesian Reasoning. *Psychological Bulletin*, 143, 1273–1312.
- Okan, Y., Galesic, M. & Garcia-Retamero, R. (2016). How People with Low and High Graph Literacy Process Health Graphs: Evidence from Eye-tracking. *Journal of Behavioral Decision Making*, 29(2-3), 271–294.
- Verschaffel, L., Lehtinen, E. & Van Dooren, W. (2016). Neuroscientific studies of mathematical thinking and learning: A critical look from a mathematics education viewpoint. *ZDM*, 48(3), 385–391.
- Woller-Carter, M. M., Okan, Y., Cokely, E. T. & Garcia-Retamero, R. (2012). Communicating and distorting risks with graphs: An eye-tracking study. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 56(1), 1723–1727. Sage CA: Los Angeles, CA: SAGE Publications.
- Zhu, L. & Gigerenzer, G. (2006). Children can solve Bayesian problems: the role of representation in mental computation. *Cognition*, 98, 287–308.