

Andreas EICHLER, Freiburg, Markus VOGEL, Heidelberg

Schülervorstellungen zu Begriffen aus dem Bereich Daten und Zufall

Das Einende der Leitidee Daten und Zufall besteht darin, dass jeder Verallgemeinerungswunsch ausgehend von Daten oder kurz die Schülerfrage: „Ist das eigentlich immer so?“ die Wahrscheinlichkeitsrechnung und die Inferenzstatistik mit der rein deskriptiven Datenanalyse verbindet (Eichler & Vogel, 2009). Gibt es aber auch basale Vorstellungen, die gleichermaßen Aspekte der beiden Teilideen Daten *und* Zufall berühren? Diese Frage war der Ausgangspunkt zu einem Forschungsansatz, der im Folgenden nach der Klärung eines theoretischen Rahmens anhand der Ergebnis-Skizze einer Pilotierung diskutiert werden soll.

Bestehende Forschungsansätze zu Daten und/oder Zufall

Es gibt eine Fülle von Arbeiten, die Vorstellungen von Schülerinnen und Schülern zum Zufall (bzw. dem Wahrscheinlichkeitsbegriff) umfassen (vgl. Jones, Langrall & Mooney, 2007). Viele dieser Forschungsansätze stehen in der Tradition von Piaget und Inhelder (1958) und versuchen, Ergebnisse dieser nachhaltigen Forschungsarbeit zu replizieren oder zu modifizieren (z.B. Green, 1979). Entscheidender Gedanke dieser Modifizierungsversuche ist es, die gelingende Einschätzung stochastischer Situationen von Schülerinnen und Schülern nicht allein auf deren Alter, sondern auf die Art der Repräsentation von Aufgaben zu beziehen und – im Gegensatz zu Piaget und Inhelder – Entscheidungen zwischen zwei Zufallsgeneratoren in einer bestimmten Situation vorzulegen.

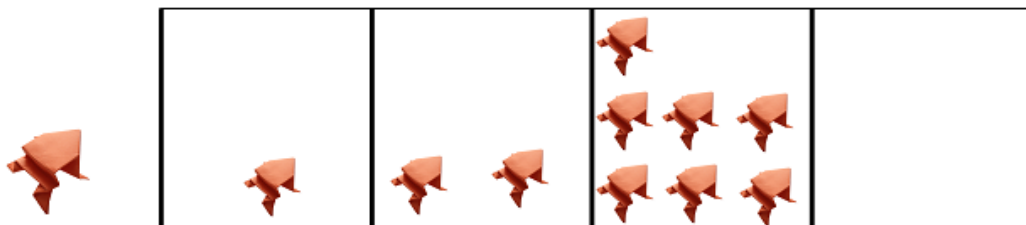
Auf der anderen Seite gibt es wiederum eine Fülle von Arbeiten zum statistischen Denken, die verschiedene Aspekte der deskriptiven Analyse von Daten betreffen (Shaughnessy, 2007). Während ein größerer Anteil dieser Arbeiten ohne offensichtliche Verbindung zum Zufall agieren, besteht dieser, wenn im Zusammenhang mit dem proportionalen Denken eine Verallgemeinerung von wenigen Fällen auf eine größere Menge von Fällen durchdacht werden muss (Jones et al., 2007). Insbesondere dort kommt die Verbindung von Daten und Zufall dann zum Tragen, wenn von einem zufälligen Ereignisses bei wenigen Versuchen bei der Prognose des Ereignisses für viele Versuche der Fokus auf die Variabilität statistischer Daten geschlossen werden muss, die wiederum im Kern den Zufall enthält (Daten = Muster + Variabilität/Zufall; vgl. Eichler & Vogel, 2009).

Für die Beurteilung der Schülerantworten verschiedener Testaufgaben wurden in der Pilotierung allgemein zwei Dimensionen festgelegt, eine Di-

chotome Unterscheidung von richtig und falsch, andererseits eine Unterscheidung der eingeforderten Begründungen nach einem Modell von Biggs und Collis (1982) in der Adaption von Watson und Callingham (2003). In dieser Dimension wurden folgende fünf Kategorien bei der Bearbeitung stochastischer Aufgaben definiert:

0. keine Bearbeitung.
1. Begründung anhand irrelevanter Informationen (praestructural mode).
2. Begründung anhand einer relevanten Information (unistructural mode)
3. Begründung anhand mehrere relevanter Informationen (multistructural mode)
4. Begründung anhand mehrere relevanter Informationen, die mit anderen, nicht in der Aufgabe vorhandenen Informationen vernetzt werden (relational mode)

Bei der folgenden Betrachtung einer Aufgabe, die die vorher genannten Aspekte des proportionalen Denkens sowie der Variabilität aufnimmt, ist zu beachten, dass noch keine Unterscheidung zwischen der Kommunikationsfähigkeit und der stochastisch geprägten Vorstellungen vorgenommen wurde. Die Aufgabe selbst wurde in verschiedenen Altersstufen in gleicher Weise formuliert:



Andrea lässt einen Papierfrosch 10 Mal springen. Der Frosch landet sieben Mal im Feld 3, zwei Mal im Feld 2 und ein Mal im Feld 1.

Die Schülerinnen und Schüler (jeweils etwa 25 einer 4., einer 10., einer 12. und einer 13. Jahrgangsstufe) sollten eine Schätzung der Verteilung für 100 und 1000 Versuche für das Feld 3 sowie bei 1000 Versuche für das Feld 4 machen und eine Begründung für diese Schätzung geben.

Ergebnis-Skizze

Unabhängig von einer altersbedingten Entwicklung der Fähigkeit, Begründungen zu kommunizieren, haben sich auch in einer Alterstufe erhebliche

Unterschiede bei der Bearbeitung der Aufgaben ergeben, wie die folgenden Antworten von Schülerinnen und Schülern einer 10. Jahrgangsstufe zeigen.

Wie oft landet der Frosch bei 100 Versuchen im Feld 3? 25 Mal
 Begründe Deine Antwort kurz:
 Die Wahrscheinlichkeit beträgt 0,25%, dass der Frosch in einem der vier Feldern landet.

Wie oft landet der Frosch bei 100 Versuchen im Feld 3? 70 Mal
 Begründe Deine Antwort kurz:
 100 ist das 10-fache von 10 und somit muss man 7×10 rechnen.

Schätze:
 Wie oft landet der Frosch bei 100 Versuchen im Feld 3? 50 Mal
 Begründe Deine Antwort kurz: Offter als im Feld 1 und 2

Wie oft landet der Frosch bei 1000 Versuchen im Feld 4? 20-10 Mal
 Begründe Deine Antwort kurz:
 Je öfter man den Frosch springen lässt, desto größer ist die W'keit, dass ein Frosch das 4. Feld trifft

Bei der ersten Antwort fließen irrelevante Informationen, die möglicherweise auf ein Primat der Gleichverteilung in stochastischen Situationen hinweist, in die Begründung ein (prästructural). Die zweite Antwort zeigt (auch bezogen auf die nicht gezeigten Antworten) eine strenge proportionale Fortsetzung der Anfangssituation (unistructural), die in der dritten Antwort durch eine erste Beachtung der Variabilität aufgeweicht wird (multisturctural) und schließlich in der letzten Antwort deutlicher zum Tragen kommt (relational).

Betrachtet man alle befragten Schülerinnen und Schüler, so ergibt sich das in der folgenden Tabelle prozentual dargestellte Ergebnis.

	K4	K10	K12	K13	Summe
n.B	0	3	0	0	1
Prä	33	10	13	0	14
Uni	47	41	47	20	41
Multi	20	38	27	40	32
Relational	0	7	13	40	12

Tabelle 1: Prozentuale Zuordnung zu den Kategorien der Erklärungsgüte

Es zeigt sich (wie auch bei anderen Aufgaben) insgesamt im Vergleich der Jahrgangsstufen eine ansteigende Fähigkeit, stochastische Situationen zu beurteilen und diese zu kommunizieren. Dabei ist es allerdings erstaunlich, dass – ohne die Fähigkeit des Kommunizierens von Begründungen zu beachten – der Anteil derjenigen Schülerinnen und Schüler, die eine unstrukturelle Begründungsvariante anbieten, nahezu identisch ist.

Unabhängig von der hier vorgestellten Aufgabe ist allerdings auch auffallend, dass – wenn es eine richtige und falsche Lösung bei den Aufgaben gab – es eine deutliche Korrelation zwischen richtigen Antworten und elaborierten Begründungen gab. Die noch vielen, schwer zu kontrollierenden Variablen sollen in einer Folgeuntersuchung eingeschränkt werden, um den Einfluss einer der als relevant erachteten Variablen (Alter, Schulform, Repräsentation der Aufgabe) auf die basalen Vorstellungen zu Daten und Zufall beurteilen zu können.

Literatur

- Biggs, J. B. & Collis, K. F. (1982). *Evaluating the quality of learning: The SOLO taxonomy*. New York: Academic Press.
- Eichler, A. & Vogel, M. (2009). *Leitidee Daten und Zufall*. Heidelberg: Vieweg.
- Green, D.R. (1979). The chance and probability concepts project. *Teaching Statistics*, 1, 3, 66-71.
- Jones, G.A., Langrall, C.W., & Mooney, E.S. (2007). Research in probability. In F.K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 909-956). Charlotte, USA: Information Age Publishing.
- Piaget, J. & Inhelder, B. (1958). The origin of chance.
- Shaughnessy, M. (2007). Research on statistics learning and reasoning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 957-1010). Charlotte, USA: Information Age Publishing.
- Watson, J. & Callingham (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal*, 2, 3-46.