



# Förderung von Bewertungskompetenz bei Schülerinnen und Schülern in komplexen Problemsituationen Nachhaltiger Entwicklung

Sabina Eggert<sup>1</sup>, Susanne Bögeholz<sup>1</sup>,  
Marcus Hasselhorn<sup>2</sup>, Rainer Watermann<sup>3</sup>

<sup>1</sup>Georg-August-Universität Göttingen

<sup>2</sup>Deutsches Institut für Internationale Pädagogische Forschung (DIPF) Frankfurt

<sup>3</sup>Freie Universität Berlin

08.10.2013 - Bildungspolitisches Forum 2013 - Goethe-Universität, Frankfurt

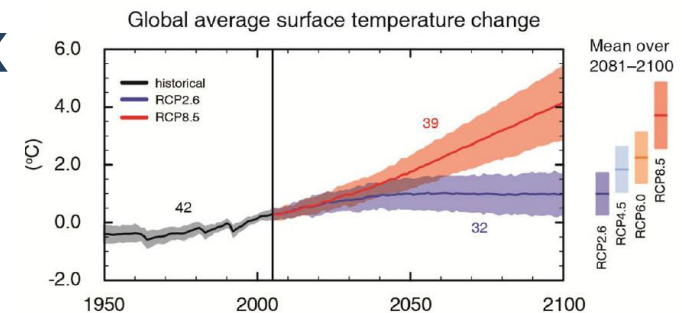
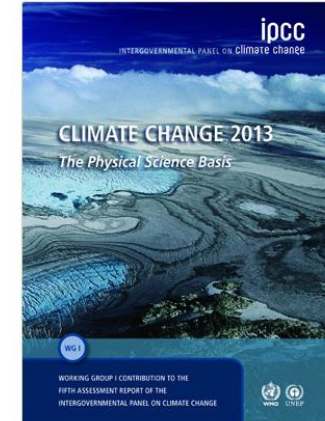
## Überblick

1. Charakteristika komplexer Umweltprobleme (*socioscientific issues*)
2. Kompetenzen von Schülerinnen und Schülern im Umgang mit SSIs
3. Trainingsstudie zur Förderung von Bewertungskompetenz (Vortest-Nachtest Messung)
4. Grenzen und Herausforderungen

# Klimawandel – das globale Umweltproblem des 21. Jh.

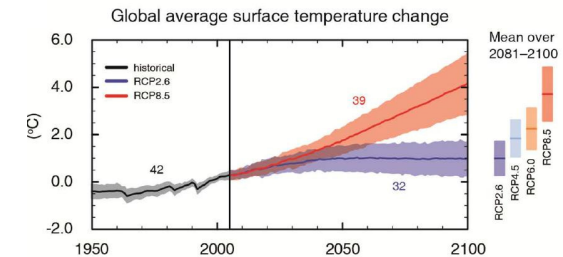
## Charakteristika komplexer Umweltprobleme (*socioscientific issues*)

- > „first-frontier science“ oder „science in the making“
- > faktisch und ethisch komplex
- > keine optimale Lösung, kein „richtig“ oder „falsch“
- > konfligierende, fragile (naturwissenschaftliche) Evidenz



# Schülerkompetenzen im Umgang mit SSIs

- > Problemsituation erkennen und beschreiben
- > Handlungsoptionen entwickeln
  - > Informationen suchen und verarbeiten
- > Handlungsoptionen bewerten
  - > argumentieren, Vor und Nachteile gegeneinander abwägen, Argumente bzw. Entscheidungskriterien gewichten
- > Bewertungs- und Entscheidungsprozesse reflektieren



(IPCC 2013, 33)

um letztendlich informierte Entscheidungen im Sinne einer Nachhaltigen Entwicklung treffen zu können. (Bögeholz 2011; Eggert & Bögeholz, 2006)

→ citizenship education, education for sustainable development, climate literacy

(KMK-BMZ 2008; NRC 2012; QCA 2007; Ratcliffe & Grace, 2003; Sadler et al. 2007; Kolsto, 2006)

# Interventionsstudie zur Förderung von Bewertungskompetenz

- 258 Schülerinnen und Schüler der 7. Jahrgangsstufe, Alter = 12,8 Jahre
- Training A: Kooperatives Lernen (KOOP): n=118;  
Training B: Kooperatives Lernen plus Metakognition (KOOP META):n=140
- 10 Lehrkräfte, 5 pro Variante (zufällig zugeordnet)

Vortest

## Unterrichtseinheit zur Förderung von Bewertungskompetenz



Nachtest

(Eggert, Bögeholz, Watermann & Hasselhorn, 2010)



## Messung Kompetenzzuwachs – Beispiel Diagnoseaufgabe (gekürzt)

Was kann gegen die Überfischung des Dorsch in der Ostsee getan werden?

**Größere Maschenöffnungen**

**Schutzzeiten und -zonen**

**Fangverbot für 3 Jahre**

**Fangquote**

Bild Dorsch: Hans-Petter Fjeld, 2006; Bild Boote: Huber/pixelio; Bild Netze: Netti69/pixelio

## Auswertung der Schülerantworten – Scoring Rubric

„Bei Maßnahme 1 gäbe es für die Konsumenten immer genügend Dorsch zu kaufen, jedoch besteht weiterhin eine Möglichkeit der Überfischung, wodurch die Arbeitsplätze in der Fischerei dann nicht mehr gesichert wären.

Maßnahme 2 hat den Nachteil, dass [...]; jedoch [...]

Maßnahme 3 bringt ein momentan zwar geringes aber nach drei Jahren ausreichendes Ladenangebot mit sich und [...]

Bei Maßnahme 4 ist eine Überfischung ausgeschlossen, jedoch [...]

Wenn man den Dorschbestand in der Ostsee sichern möchte, darf keine Überfischung stattfinden. Außerdem ist es wichtig die Arbeitsplätze zu berücksichtigen.

Wenn diese beiden Punkte berücksichtigt werden, [...]"

## Auswertung der Schülerantworten – Scoring Rubric

1.-4. Item („Maschen“, „Fangverbot“; Schutzzeiten und -zonen“; „Fangquote“)	0: keine Antwort 1: nur Pro- oder nur Kontra Argumente entwickelt 2: Pro- und Kontra Argumente entwickelt
5. Item Gewichten von Argumenten	0: keine Gewichtung von Argumenten 1: Gewichtung auf Basis persönlicher und/oder gesellschaftlicher Werte
6. Reflexion Bewertungsprozess „Ausschlussverfahren“	0: keine Antwort oder teilweise richtige Antwort 1: vollständige Antwort
7. Reflexion Bewertungsprozess „trade-off“	0: keine Antwort oder teilweise richtige Antwort 1: vollständige Antwort
8. Verbesserungsvorschlag „trade-off“	0: kein Vorschlag oder Vorschlag auf Inhaltsebene 1: strategischer Vorschlag



# Auswertungsmethodik

- Modellierung der Daten des Vor- und Nachtests mit dem Rasch Partial Credit Model (Masters, 1982)

- Methodische Herausforderung bei der Vortest-Nachtest Messung:

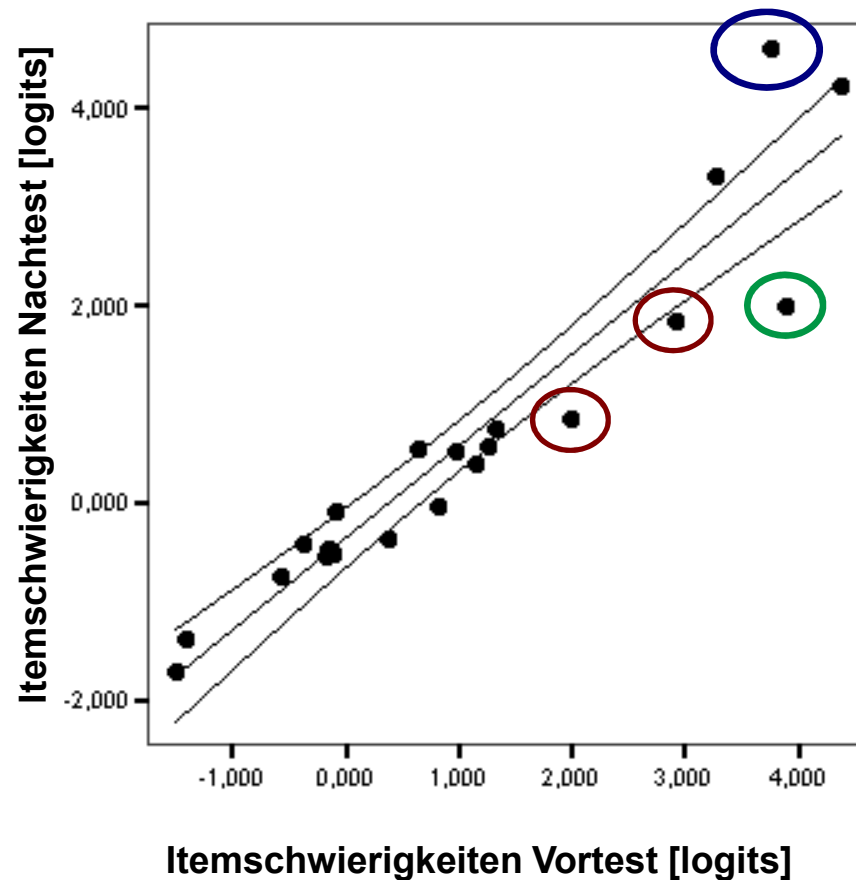
Kann die Bewertungskompetenz der Schülerinnen und Schüler zum Nachtest als Lernzuwachs im Vergleich zum Vortest interpretiert werden?

oder

Misst das Testinstrument dieselbe Kompetenz auch unter veränderten Bedingungen?

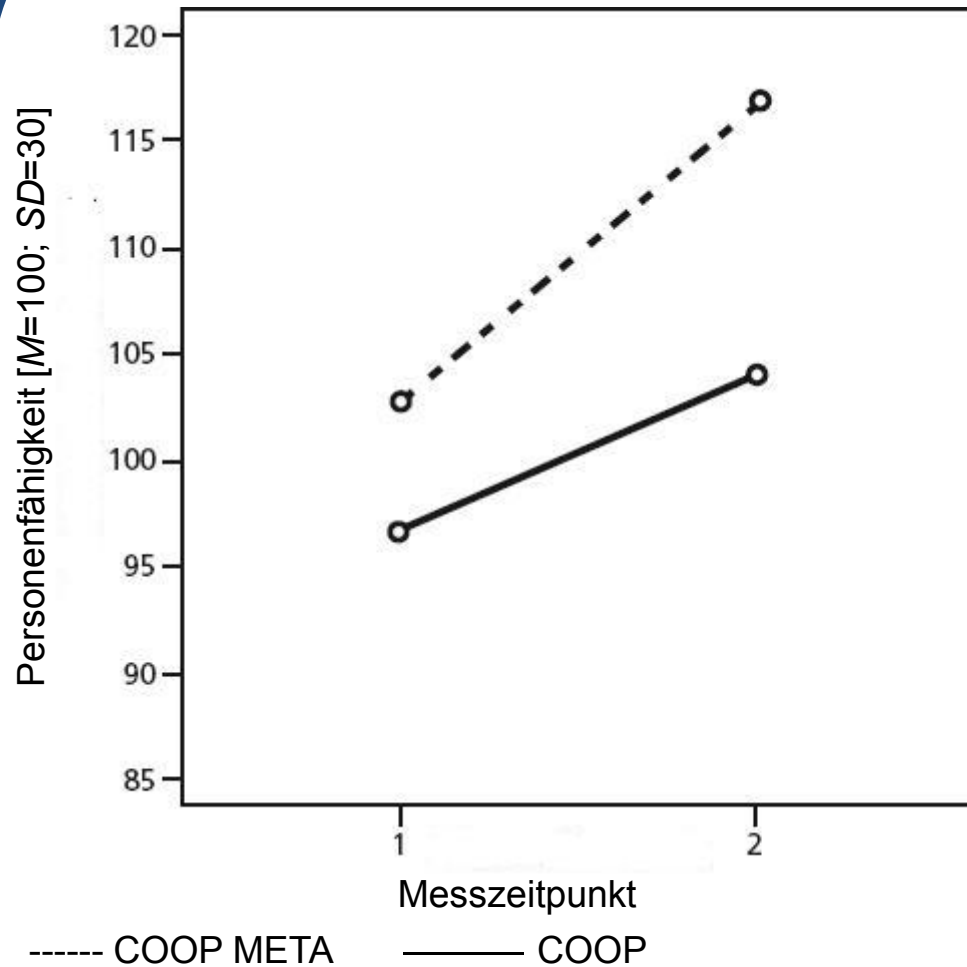
-> differential item functioning (dif)

# Ergebnisse der dif Analysen



- mehrere Items mit mittleren und hohen Logit Differenzen (Paek, 2002 zitiert nach Wilson, 2005)
- einige Items bezüglich der Entwicklung von **Pro und Kontra Argumenten**
- Items bezüglich **Gewichten** von Argumenten  
→ “zu einfach” zum Nachtest (aber generell noch schwierig)
- Item zur **Verbesserung von trade-off Strategien**  
→ “zu schwierig” zum Nachtest

## Ergebnisse des Treatmentvergleichs – KOOP vs. KOOP META



- keine Vortestunterschiede bezüglich Bewertungskompetenz ( $t_{(256)} = -1.632, p=.104$ )

- Haupteffekt Zeit: ( $F_{(1,256)} = 17.240, p<.001$ )

- Haupteffekt Gruppe ( $F_{(1,256)} = 6.136, p<.05$ )

- keine Interaktion: ( $F_{(1,256)} = 1.751, p>.05$ )

- ANCOVA: ( $F_{(1,256)} = 3.598, p=.06$ )

# Zusammenfassung und Diskussion

## Veränderungsmessung

- > Aufgrund zu starker Veränderung der Itemschwierigkeit konnten wichtige Item (-steps) nicht in die Messung des Kompetenzzuwachses der Lernenden aufgenommen werden.
  - > Informationsverlust
- > Nichtbeachtung der *dif* Ergebnisse hätte zur Überschätzung des Kompetenzzuwachses geführt.

## Treatmentvergleich KOOP – KOOP META

- > Zuwachs an Bewertungskompetenz in beiden Varianten
- > KOOP META Gruppe zu beiden Messzeitpunkten besser als KOOP Gruppe
- > differentielle Wirksamkeit tendenziell nachgewiesen 😊
- > Ergebnisse beruhen z.T. auf den Problemen bezüglich der Voraussetzungen für Veränderungsmessung
- > bei ökologisch hoch validem setting

- > Weiterentwicklung des Testinstruments
  - > unterschiedliche Aufgaben im Vor und Nachtest (Ankerdesign)
  
- > Optimierung der Schulung von Lehrkräften
  - > Lehrerinnen und Lehrer fokussierten v.a. auf das Entwickeln von Pro und Kontra Argumenten (sowie Gewichten) mit Hilfe von Bewertungsmatrizen
    - > Bewertungsmatrix als favorisiertes „tool“
  - > vernachlässigten das Reflektieren über sowie das Kritisieren von Bewertungsprozessen
  
- > Chancen und Grenzen von Veränderungsmessungen
  - bisher nur wenige Trainingsstudien, die Veränderungsmessung mit IRT Modellierung einsetzen
  - realistischere Einschätzung der Kompetenzzuwächse von Lernenden
  - wichtige Hinweise auf Umsetzung der Treatments im Unterricht, die über Analysen ausschließlich mit KTT nicht zugänglich gewesen wären



Herzlichen Dank für Ihre  
Aufmerksamkeit!

Kontakt:  
Dr. Sabina Eggert  
[seggert1@gwdg.de](mailto:seggert1@gwdg.de)