

# Modellverstehen im Biologieunterricht

## Evaluation von Diagnosestrategien

Ziel des Projekts ist die Entwicklung eines computerbasierten Instruments zur effizienten Diagnose des individuellen Modellverstehens von Schülerinnen und Schülern. Dies ermöglicht durch direkte Rückmeldung eine differenzierte Förderung von Modellkompetenz.

### Theoretische Grundlagen

Grünkorn et al. (2014) haben auf der Basis von Schüleraussagen ein Kategoriensystem entwickelt, welches für die Teilkompetenz *Eigenschaften von Modellen* Perspektiven beschreibt, die sich auf verschiedene Ähnlichkeits- bzw. Abstraktionsgrade zwischen Modell und Ausgangsobjekt beziehen (Abb. 1). Vor allem der Perspektivwechsel von Niveau II zu Niveau III stellt bei der Vermittlung eine Herausforderung dar (Fleige et al. 2012).

### Forschungsfragen

- Welche Anzahl von Forced Choice (FC)-Aufgabenbearbeitungen ist erforderlich, um eine zuverlässige Zuordnung der Personen in zwei Klassen (Niveau I und II vs Niveau III) zu erreichen?
- Inwieweit ist das mit FC-Aufgaben erfasste Modellverstehen mit dem anderer Antwortformate vergleichbar?

### Methode

Das Kategoriensystem zur Teilkompetenz *Eigenschaften von Modellen* (Abb. 1) wurde für die Entwicklung von 16 FC-Aufgaben genutzt. Davon bearbeiteten Schülerinnen und Schüler ( $N=543$ ; 7.-12. Klasse) in einer ersten Teilstudie je 6-9 Aufgaben. In einer zweiten Teilstudie bearbeiteten Schülerinnen und Schüler ( $N=189$ ; 7.-10. Klasse) je 12 FC-Aufgaben sowie 2 offene Aufgaben. Ergänzend wurden 50 kurze vollstrukturierte Interviews geführt (Abb. 2).

Auf der Basis aller FC-Aufgabenbearbeitungen wurde geprüft, wie wahrscheinlich eine zutreffende Klassifizierung in Abhängigkeit von der Anzahl der Bearbeitungen ist (Bootstrap-Verfahren; Bortz & Döring 2006).

Für die Untersuchung der zweiten Forschungsfrage wurde mit offenen Aufgaben und Interviews trianguliert (Abb. 2). Für jede Person wurde pro Antwortformat der Median der geäußerten Niveaus ermittelt und verglichen (Wilcoxon-Test).

### Ergebnisse und Ausblick

Schülerinnen und Schüler verstehen Modelle vorwiegend als Repräsentationen einer Wirklichkeit im Sinne der Niveaus I und II (Abb. 4; vgl. Grünkorn et al. 2014, Krell 2013).

Die Bootstrap-Simulationen auf der Basis des Medians ergeben, dass bei einer Zuordnung der Personen zu zwei Klassen fünf Aufgabenbearbeitungen ausreichen, um ein angemessen stabiles Diagnoseurteil zu fällen (Abb. 3). Das Bootstrap-Verfahren ermöglicht es, die Effizienz der Diagnosestrategie zu untersuchen, wenn weitere Variablen (z. B. Aufgabenkontext, Klassenstufe, Schulnote) einbezogen werden.

Bei den FC-Aufgaben antworten Schülerinnen und Schüler auf einem signifikant höheren Niveau als bei Aufgaben im offenen Format ( $U = 8.786$ ,  $p < .001$ ,  $r = .660$ ; Abb. 4). Im Vergleich mit geschlossenen Antwortformaten können Aufgaben im offenen Format höhere kognitive Anforderungen an eine Person stellen (Martinez 1999).

Um die Antworten bei den FC-Aufgaben bezogen auf das Modellverstehen interpretieren zu können, ist eine Studie mit Lautem Denken (Sandmann 2014) geplant.

### Eigenschaften von Modellen

Niveau I	Niveau II	Niveau III
Modelle sind Kopien von etwas	Modelle sind idealisierte Repräsentationen von etwas	Modelle sind theoretische Rekonstruktionen von etwas
Modell als Kopie	Modell ist in Teilen eine Kopie	Modell als hypothetische Darstellung
Modell mit großer Ähnlichkeit	Modell als eine mögliche Variante	
Modell entspricht (nicht) subjektiver Vorstellung vom Original	Modell als fokussierte Darstellung	

Abb. 1. Kategoriensystem zur Teilkompetenz *Eigenschaften von Modellen* (nach Grünkorn et al. 2014)

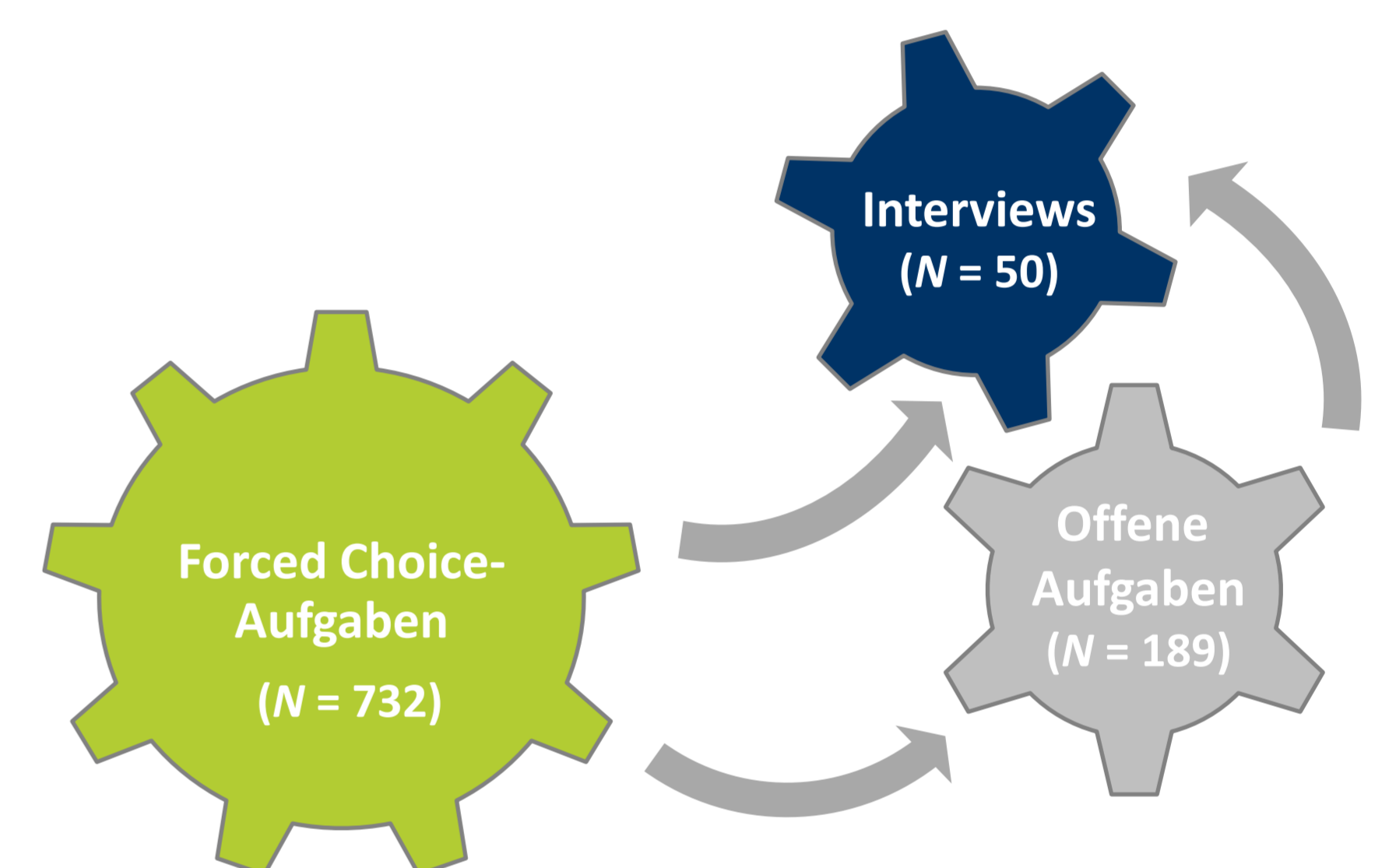


Abb. 2. Zahnradmodell zur Triangulation der Antwortformate

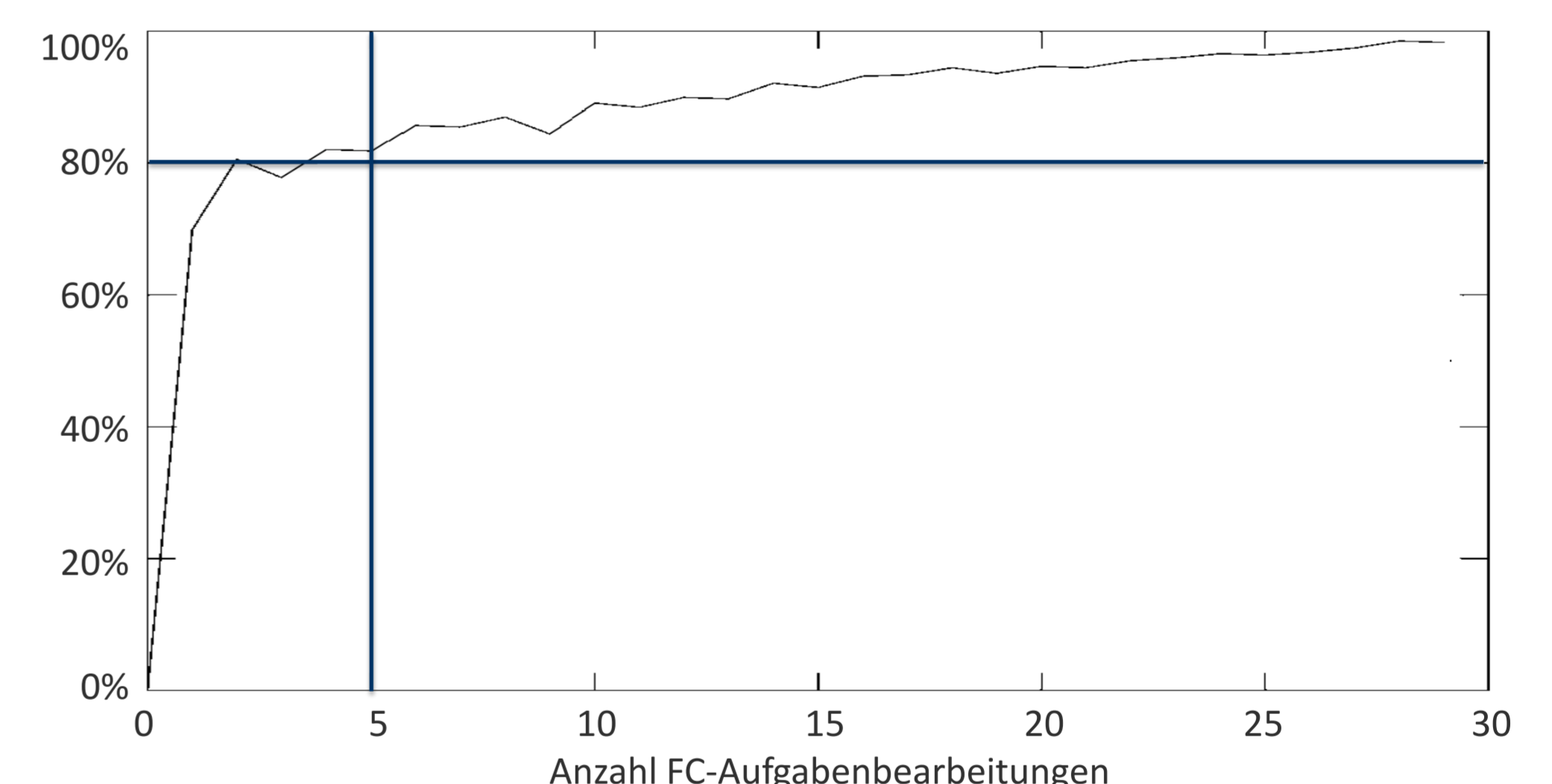


Abb. 3. Zuverlässigkeit des Diagnoseurteils in Abhängigkeit von der Anzahl der FC-Aufgabenbearbeitungen bei der Einordnung einer Person in zwei Klassen auf der Basis des Medians

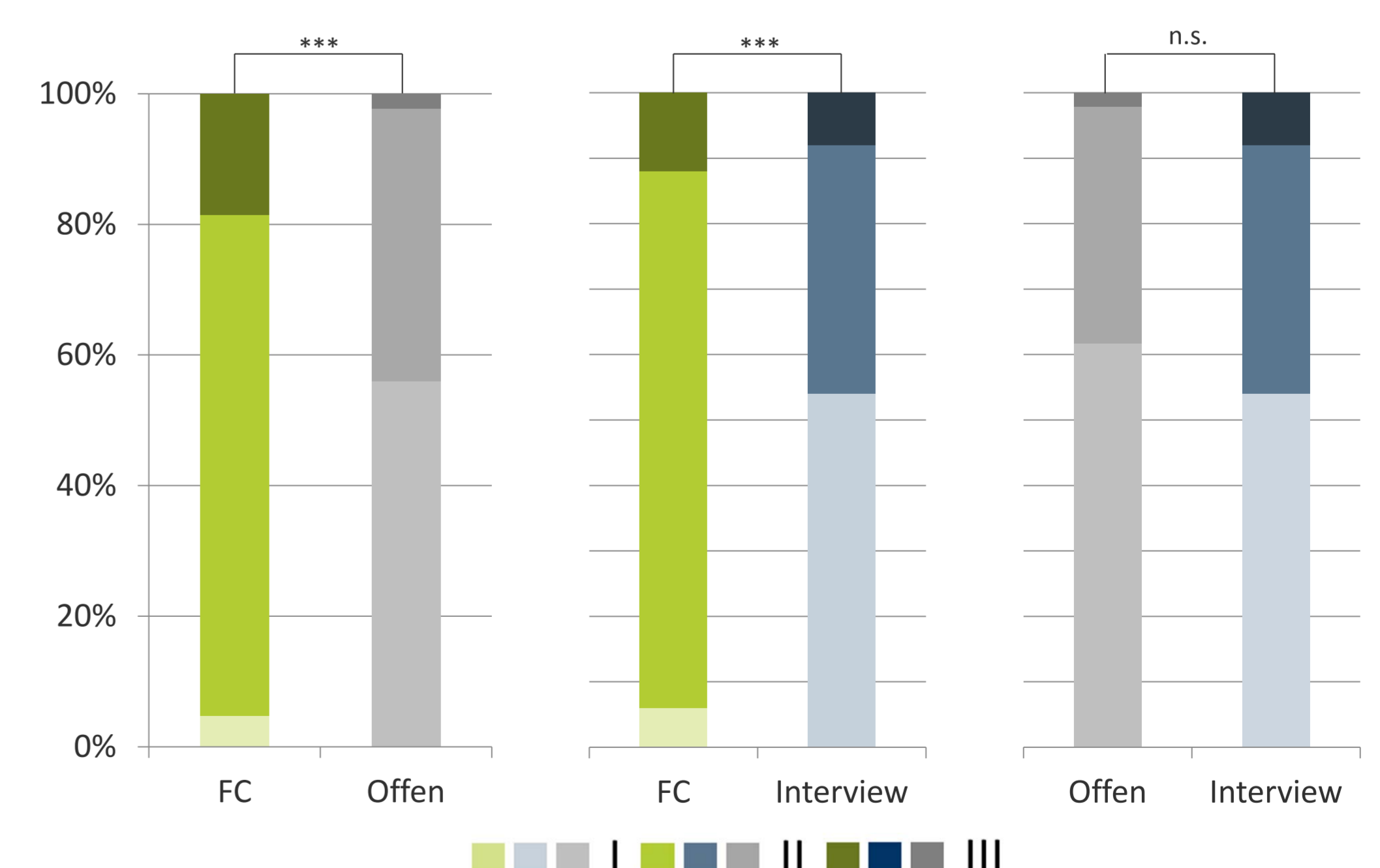


Abb. 4. Häufigkeiten der Mediane der geäußerten Niveaus pro Antwortformat. Vergleich der Antwortformate mittels Wilcoxon-Test (links: FC vs Offen  $N=189$ ; mittig: FC vs Interview  $N=50$ ; rechts: Offen vs Interview  $N=50$ ). n. s.: nicht signifikant, \*\*\*:  $p < .001$





## Literatur

Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaften*.

Heidelberg: Springer.

Fleige, J., Seegers, A., Upmeyer zu Belzen, A., & Krüger, D. (2012). *Modellkompetenz im Biologieunterricht 7-10*.

Donauwörth: Auer.

Grünkorn, J., Upmeyer zu Belzen, A., & Krüger, D. (2014). Assessing students' understandings of biological models and their use in science to evaluate a theoretical framework. *International Journal of Science Education*. doi:10.1080/09500693.2013.873155

Krell, M. (2013). *Wie Schülerinnen und Schüler biologische Modelle verstehen*. Berlin: Logos.

Martinez, M. E. (1999). Cognition and the question of test item format. *Educational Psychologist*, 34, 207–218.

Sandmann, A. (2014). Lautes Denken – die Analyse von Denk-, Lern- und Problemlöseprozesse. In D. Krüger, I. Parchmann & H. Schecker (Hrsg.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (S. 179-188).

Heidelberg: Springer.