

**Interaktion von Erfahrungsgrad und Modelleinsatz
bei Tree-Reading-Prozessen**

D i s s e r t a t i o n

zur Erlangung des akademischen Grades

d o c t o r r e r u m n a t u r a l i u m
(Dr. rer. nat.)

im Fach Biologie

eingereicht an der

Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin
von

M. Sc. Inga Gesine Ubben-Lother, geb. Ubben

Präsidentin der Humboldt-Universität zu Berlin
Prof. Dr. Julia von Blumenthal

Dekan der Lebenswissenschaftlichen Fakultät
Prof. Dr. Dr. Christian Ulrichs

Gutachter/innen: 1. Prof. Dr. Annette Upmeyer zu Belzen
2. Prof. Dr. Sandra Nitz
3. Prof. Dr. Philipp Schmiemann

Tag der mündlichen Prüfung: 12.10.2023

Für Philipp und Pauli

„Da kann noch einiges evolutionär dazwischen passiert sein.“

(NiEr P17, Alt1 methodischer Modelleinsatz)

Inhaltsverzeichnis

| | |
|---|------------|
| ABKÜRZUNGSVERZEICHNIS | IV |
| ABBILDUNGSVERZEICHNIS | V |
| TABELLENVERZEICHNIS | VII |
| ZUSAMMENFASSUNG | 1 |
| ABSTRACT | 4 |
| 1. EINLEITUNG | 5 |
| 2. THEORIE | 9 |
| 2.1. VON DER GEMEINSAMEN ABSTAMMUNG ZU DIAGRAMMEN PHYLOGENETISCHER VERWANDTSCHAFT..... | 9 |
| 2.1.1. DIAGRAMME PHYLOGENETISCHER VERWANDTSCHAFT - TERMINOLOGIE..... | 11 |
| 2.1.2. TREE-THINKING UND TREE-READING | 13 |
| 2.1.3. DIAGRAMME PHYLOGENETISCHER VERWANDTSCHAFT - <i>WISSENSCHAFTLICH ADÄQUATE</i> LESEWEISE | 14 |
| 2.1.4. DIAGRAMME PHYLOGENETISCHER VERWANDTSCHAFT - <i>WISSENSCHAFTLICH INADÄQUATE</i> LESEWEISE..... | 15 |
| 2.2. TREE-READING-FÄHIGKEITEN UND IHRE STRUKTUR | 19 |
| 2.3. DIAGRAMME PHYLOGENETISCHER VERWANDTSCHAFT IM KONTEXT DER MODELLETHEORIE..... | 28 |
| 2.3.1. DIAGRAMME PHYLOGENETISCHER VERWANDTSCHAFT ALS MODELLE <i>VON</i> UND <i>FÜR</i> EVOLUTION | 28 |
| 2.3.2. KOMPETENZMODELLE DER MODELL- UND MODELLIERKOMPETENZ | 30 |
| 2.4. ERFAHRUNGSGRAD..... | 33 |
| 2.5. EYETRACKING - VISUELLE WAHRNEHMUNG VON REPRÄSENTATIONEN | 34 |
| 2.6. DIDAKTISCHE RELEVANZ VON DIAGRAMMEN PHYLOGENETISCHER VERWANDTSCHAFT UND TREE-READING..... | 35 |
| 3. STAND DER FORSCHUNG | 37 |
| 3.1. TREE-READING | 37 |
| 3.1.1. INTERPRETATION VERSCHIEDENER REPRÄSENTATIONSFORMEN UND OBERFLÄCHENMERKMALE | 38 |
| 3.1.2. INTERPRETATION VON <i>MRC</i> A UND RELATIVER VERWANDTSCHAFT..... | 40 |
| 3.1.3. IDENTIFIKATION UND INTERPRETATION VON MERKMALEN UND EVOLUTIVEN MECHANISMEN | 41 |
| 3.1.4. VERGLEICHEN MEHRERER DIAGRAMME PHYLOGENETISCHER VERWANDTSCHAFT..... | 42 |
| 3.2. ERFAHRUNGSGRAD..... | 43 |
| 3.2.1. EINFLUSS DES ERFAHRUNGSGRADES AUF TREE-READING..... | 43 |
| 3.2.2. EINFLUSS DES ERFAHRUNGSGRADES AUF MODELLKOMPETENZ | 43 |
| 3.2.3. EINFLUSS DES ERFAHRUNGSGRADES AUF DIE VISUELLE WAHRNEHMUNG | 44 |
| 4. FORSCHUNGSFRAGEN UND HYPOTHESEN | 46 |
| 5. MATERIAL UND METHODEN | 49 |
| 5.1. METHODISCHE ENTSCHEIDUNGEN | 49 |
| 5.2. OPERATIONALISIERUNG | 51 |

| | | |
|-----------|---|------------|
| 5.3. | PILOTIERUNG <i>BETTSI</i> UND ANALYSE DER ANDEREN KONTROLLVARIABLEN | 53 |
| 5.4. | AUFGABENKONSTRUKTION..... | 54 |
| 5.4.1. | AUSWAHL DER TEILKOMPETENZEN | 54 |
| 5.4.2. | KONSTRUKTION DER AUFGABEN ANHAND DER TEILKOMPETENZEN | 56 |
| 5.4.3. | KONSTRUKTION DER REPRÄSENTATIONEN UND OPTISCHE GESTALTUNG DER AUFGABEN..... | 57 |
| 5.5. | PROBAND:INNEN | 58 |
| 5.6. | VERSUCHSAUFBAU UND -ABLAUF | 58 |
| 5.7. | DATENAUFBEREITUNG..... | 61 |
| 5.8. | QUALITATIVE AUSWERTUNG DER VERBALEN DATEN DURCH KODIERUNG..... | 62 |
| 5.8.1. | ENTWICKLUNG DER KATEGORIEN..... | 63 |
| 5.8.2. | GÜTEKRITERIEN | 64 |
| 5.8.3. | CODEHÄUFIGKEITEN | 66 |
| 5.9. | CHARAKTERISIERUNG UND BESCHREIBUNG VON INDIVIDUELLEN LÖSUNGSPROZESSEN..... | 66 |
| 5.10. | QUALITATIVE AUSWERTUNG DER VISUELLEN DATEN UND TRIANGULATION | 67 |
| 6. | ERGEBNISSE | 68 |
| 6.1. | INDUKTIVE ERWEITERUNG DES KATEGORIENSYSTEMS | 68 |
| 6.2. | PILOTIERUNG <i>BETTSI</i> | 69 |
| 6.3. | STICHPROBENBESCHREIBUNG | 70 |
| 6.4. | INTERCODERRELIABILITÄT..... | 74 |
| 6.5. | ALLGEMEINE CODEHÄUFIGKEITEN IM MEDIALEN MODELLEINSATZ..... | 75 |
| 6.6. | ALLGEMEINE CODEHÄUFIGKEITEN IM METHODISCHEN MODELLEINSATZ..... | 81 |
| 6.7. | MAXIMALE TREE-READING-FÄHIGKEIT UND LÖSUNG DER AUFGABEN..... | 87 |
| 6.8. | INTERINDIVIDUELLER VERGLEICH VON LÖSUNGSPROZESSEN | 90 |
| 6.8.1. | <i>MRCA</i> IDENTIFIZIEREN UND BESCHREIBEN (AUFGABE <i>ZWECK1</i> MEDIALER MODELLEINSATZ)..... | 93 |
| 6.8.2. | DPV VERGLEICHEN (AUFGABE <i>ALT1</i> MEDIALER MODELLEINSATZ) | 96 |
| 6.8.3. | MERKMAL DES <i>MRCA</i> ABLEITEN (AUFGABE <i>ZWECK1</i> METHODISCHER MODELLEINSATZ)..... | 102 |
| 6.8.4. | KONVERGENTE MERKMALSENTSTEHUNG ABLEITEN (AUFGABE <i>ZWECK2</i> METHODISCHER MODELLEINSATZ) | 107 |
| 6.8.5. | VORGEGBENE HYPOTHESE EINEM DPV ZUORDNEN (AUFGABE <i>ALT1</i> METHODISCHER MODELLEINSATZ)..... | 110 |
| 6.8.6. | VORGEGBENE HYPOTHESE EINEM DPV ZUORDNEN UND NEUE HYPOTHESE GENERIEREN (AUFGABE <i>ALT2</i> METHODISCHER MODELLEINSATZ) | 115 |
| 6.9. | TRIANGULATION VON VERBALEN UND VISUELLEN DATEN..... | 117 |
| 6.10. | VERSIONSUNTERSCHIEDE | 121 |
| 6.11. | PRÄVALENZ WISSENSCHAFTLICH INADÄQUATES ABLESEN..... | 123 |
| 6.12. | KONTROLLVARIABLEN | 124 |
| 6.13. | ZUSAMMENFASSUNG DER ERGEBNISSE | 126 |
| 7. | DISKUSSION | 129 |
| 7.1. | DISKUSSIONSGRUNDLAGE | 129 |

| | | |
|------------|---|------------|
| 7.2. | EINFLUSS VON ERFAHRUNGSGRAD UND MODELLEINSATZ AUF TREE-READING-PROZESSE | 130 |
| 7.3. | INTERAKTION VON ERFAHRUNGSGRAD UND MODELLEINSATZ..... | 131 |
| 7.4. | EINORDNUNG IN TREE-READING-STRUKTURIERUNG..... | 132 |
| 7.6. | EINFLUSSFAKTOREN JENSEITS VON ERFAHRUNGSGRAD UND MODELLEINSATZ..... | 135 |
| 7.6.1. | MERKMALE DER REPRÄSENTATIONEN..... | 135 |
| 7.6.2. | AUFGABENREIHENFOLGE | 137 |
| 7.6.3. | FACHSPRACHE UND FACHWISSEN | 138 |
| 7.8. | VISUELLE DATEN | 140 |
| 7.9. | KONTROLLVARIABLEN | 141 |
| 7.10. | METHODENKRITIK | 141 |
| 7.10.1. | VERSUCHSSETTING..... | 141 |
| 7.10.2. | WEITERE EINFLUSSFAKTOREN | 143 |
| 8. | IMPLIKATIONEN | 144 |
| 9. | FAZIT | 146 |
| 10. | AUSBLICK | 147 |
| 11. | DANKSAGUNG | 150 |
| 12. | LITERATURVERZEICHNIS..... | 151 |
| 13. | ANHANG | 160 |
| 13.1. | ANHANG 1: BASIC EVOLUTIONARY TREE-THINKING SKILL INVENTORY (BETTSI)..... | 160 |
| 13.2. | ANHANG 2: AUFGABE ZWECK1 | 168 |
| 13.3. | ANHANG 3: AUFGABE ZWECK2 | 170 |
| 13.4. | ANHANG 4: AUFGABE ALT1..... | 172 |
| 13.5. | ANHANG 5: AUFGABE ALT2..... | 174 |
| 13.6. | ANHANG 6: TRANSKRIPTIONSREGELN NACH DRESING UND PEHL (2018)..... | 176 |
| 13.7. | ANHANG 7: KODIERLEITFADEN | 178 |

Abkürzungsverzeichnis

| | |
|----------------|---|
| DpV | Diagramm phylogenetischer Verwandtschaft |
| <i>MRCA</i> | <i>Most Recent Common Ancestor</i> |
| <i>STREAM</i> | <i>Synthetic Tree-Reading Model</i> |
| TR_2 | <i>Ableiten</i> |
| TR_1a | <i>wissenschaftlich adäquates Ablesen</i> |
| TR_1b | <i>wissenschaftlich inadäquates Ablesen</i> |
| TR_0 | <i>oberflächliches Ablesen</i> |
| HoEr | Person mit hohem Erfahrungsgrad mit DpV |
| NiEr | Person mit niedrigem Erfahrungsgrad mit DpV |
| <i>Zweck1</i> | <i>Aufgabe 1 zur Teilkompetenz Zweck von Modellen</i> |
| <i>Alt1</i> | <i>Aufgabe 1 zur Teilkompetenz Alternative Modelle</i> |
| <i>CTA</i> | <i>Concurrent Think Aloud</i> |
| <i>BETTSI</i> | <i>Basic Evolutionary Tree-Thinking Skills Instrument</i> |
| 3DW-Test | 3D-Würfeltest |
| ZVT | Zahlenverbindungstest |
| Sec | Sekunde |
| DpV1 | linkes DpV |
| DpV2 | rechtes DpV |
| DpV1_v | vergleichendes Beschreiben des linken DpV |
| DpV2_e | Beschreibung des rechten DpV ohne Bezug zum linken DpV |
| n. k. | nicht kodierbar |
| PS | Punkte und Streifen als Merkmal des <i>MRCA</i> |
| Topo | Topografie |
| Spars. | Sparsamkeitsprinzip |
| Konv. | Konvergenz |
| M. <i>MRCA</i> | Merkmal des <i>MRCA</i> |

Abbildungsverzeichnis

| | |
|---|-----|
| ABBILDUNG 1: ANALYSE DER EVOLUTIVEN VERWANDTSCHAFT DER VERSCHIEDENEN SARS-CoV-2-VARIANTEN | 7 |
| ABBILDUNG 2: AUFBAU UND BESTANDTEILE VON DIAGRAMMEN PHYLOGENETISCHER VERWANDTSCHAFT AM BEISPIEL EINES KLADOGRAMMS | 12 |
| ABBILDUNG 3: REVISED SYNTHETIC TREE-READING MODEL (STREAM) | 24 |
| ABBILDUNG 4: BEISPIEL FÜR DIE VISUELLE DARSTELLUNG EINER AUFGABE | 58 |
| ABBILDUNG 5: A) REMOTE EYETRACKER VON SENSOMOTORIC INSTRUMENTS UNTER MONITOR MONTIERT, WEBCAM AUF MONITOR UND TASTATUR DAVOR. B) AUFBAU IN HÖHENVERSTELLBAREM SCHRANK MIT TÜREN ALS SICHT- UND BLENDSCHUTZ. | 60 |
| ABBILDUNG 6: AREAS OF INTEREST (AOI) FÜR ALT1 MEDIALER MODELLEINSATZ ZUR AUSWERTUNG DER FIXATIONEN ÜBER DIE ZEIT. | 62 |
| ABBILDUNG 7: ALLGEMEINE CODEHÄUFIGKEITEN DER TREE-READING-FÄHIGKEITEN IN PROZENT IM MEDIALEN MODELLEINSATZ DER TEILKOMPETENZ ZWECK VON MODELLEN (A) UND ALTERNATIVE MODELLE (B) FÜR HOER UND NIER. | 76 |
| ABBILDUNG 8: ZUORDNUNG DER DPV-BESTANDTEILE ZU TREE-READING-FÄHIGKEITEN IN PROZENT DER TEILKOMPETENZ ZWECK VON MODELLEN IM MEDIALEN MODELLEINSATZ FÜR HOER UND NIER. | 77 |
| ABBILDUNG 9: ZUORDNUNG DER DPV-BESTANDTEILE ZU DEN TREE-READING-FÄHIGKEITEN IN PROZENT DER TEILKOMPETENZ ALTERNATIVE MODELLE IM MEDIALEN MODELLEINSATZ. | 79 |
| ABBILDUNG 10: BEZUG DER AUSSAGEN BEI DER TEILKOMPETENZ ALTERNATIVE MODELLE IM MEDIALEN MODELLEINSATZ. | 80 |
| ABBILDUNG 11: ANZAHL N DER PROBAND:INNEN, BEI DENEN IM MEDIALEN MODELLEINSATZ WISSENSCHAFTLICH NICHT ADÄQUATES ABLESEN (TR_1B) KODIERT WURDE. | 80 |
| ABBILDUNG 12: ALLGEMEINE CODEHÄUFIGKEITEN DER TREE-READING-FÄHIGKEITEN IN PROZENT IM METHODISCHEN MODELLEINSATZ DER TEILKOMPETENZ ZWECK VON MODELLEN (A) UND ALTERNATIVE MODELLE (B) FÜR HOER UND NIER. | 82 |
| ABBILDUNG 13: ZUORDNUNG DER DPV-BESTANDTEILE ZU TREE-READING-FÄHIGKEITEN IN PROZENT DER TEILKOMPETENZ ZWECK VON MODELLEN IM METHODISCHEN MODELLEINSATZ FÜR HOER UND NIER. | 83 |
| ABBILDUNG 14: ZUORDNUNG DER DPV-BESTANDTEILE ZU TREE-READING-FÄHIGKEITEN IN PROZENT DER TEILKOMPETENZ ALTERNATIVE MODELLE IM METHODISCHEN MODELLEINSATZ FÜR HOER UND NIER. | 85 |
| ABBILDUNG 15: BEZUG DER AUSSAGEN BEI DER TEILKOMPETENZ ALTERNATIVE MODELLE IM METHODISCHEN MODELLEINSATZ. | 86 |
| ABBILDUNG 16: ANZAHL N DER PROBAND:INNEN, BEI DENEN IM METHODISCHEN MODELLEINSATZ WISSENSCHAFTLICH NICHT ADÄQUATES ABLESEN (TR_1B) KODIERT WURDE. | 86 |
| ABBILDUNG 17: MAXIMAL ERREICHTE TREE-READING-FÄHIGKEIT IM MEDIALEN MODELLEINSATZ DER EINZELNEN PROBAND:INNEN. | 87 |
| ABBILDUNG 18: MAXIMAL ERREICHTE TREE-READING-FÄHIGKEIT IM METHODISCHEN MODELLEINSATZ DER EINZELNEN PROBAND:INNEN. | 88 |
| ABBILDUNG 19: KLASSIFIZIERUNG DER INDIVIDUELLEN LÖSUNGEN VON HOER UND NIER IM METHODISCHEN MODELLEINSATZ HINSICHTLICH IHRER ENTSPRECHUNG DER WISSENSCHAFTLICHEN KONVENTIONEN. | 89 |
| ABBILDUNG 20: AUFGABE ZWECK1 MEDIALER MODELLEINSATZ: IDENTIFIZIERUNG UND BESCHREIBUNG DES MRCA ZWEIER TAXA. | 93 |
| ABBILDUNG 21: AUFGABE ALT1 MEDIALER MODELLEINSATZ: VERGLEICHEN VON 2 DPV. | 96 |
| ABBILDUNG 22: AUFGABE ALT2 MEDIALER MODELLEINSATZ: VERGLEICHEN VON 2 DPV. | 100 |
| ABBILDUNG 23: AUFGABE ZWECK1 METHODISCHER MODELLEINSATZ: MERKMAL DES MRCA ABLEITEN. | 102 |
| ABBILDUNG 24: AUFGABE ZWECK2 METHODISCHER MODELLEINSATZ: KONVERGENTE MERKMALSENTSTEHUNG ABLEITEN. | 107 |

| | |
|--|-----|
| ABBILDUNG 25: AUFGABE ALT1 METHODISCHER MODELLEINSATZ: VORGEGEBENE HYPOTHESE EINEM DPV ZUORDNEN..... | 110 |
| ABBILDUNG 26: AUFGABE ALT2 METHODISCHER MODELLEINSATZ: VORGEGEBENE HYPOTHESE EINEM DPV ZUORDNEN UND NEUE HYPOTHESE ABLEITEN..... | 115 |
| ABBILDUNG 27: TRIANGULATION VON FIXATIONEN UND VERBALEN BEZÜGEN VON NIER P13 IN AUFGABE ALT1 IM MEDIALEN MODELLEINSATZ..... | 119 |
| ABBILDUNG 28: TRIANGULATION VON FIXATIONEN UND VERBALEN BEZÜGEN VON HOER P14 IN AUFGABE ALT1 IM MEDIALEN MODELLEINSATZ..... | 119 |
| ABBILDUNG 29: TRIANGULATION VON FIXATIONEN UND VERBALEN BEZÜGEN VON NIER P17 IN AUFGABE ALT1 IM MEDIALEN MODELLEINSATZ..... | 120 |
| ABBILDUNG 30: ZUORDNUNG DER IN DEN AUFGABEN ERFORDERLICHEN TREE-READING-FÄHIGKEITEN (WEIßE KÄSTEN) INS REVISED SYNTHETIC TREE-READING MODEL (STREAM)..... | 134 |

Tabellenverzeichnis

| | |
|---|-----|
| TABELLE 1: TREE-READING-FÄHIGKEITEN LAUT VERSCHIEDENER AUTOR:INNEN | 20 |
| TABELLE 2: GEGENÜBERSTELLUNG VON ETCI UND REVISED STREAM | 25 |
| TABELLE 3: KOMPETENZMODELL DER MODELLKOMPETENZ MIT 5 TEILKOMPETENZEN UND 3 NIVEAUS. | 30 |
| TABELLE 4: KOMPETENZMODELL DER MODELLIERKOMPETENZ | 32 |
| TABELLE 5: VOR- UND NACHTEILE VON CONCURRENT THINK ALOUD (CTA) IM VERGLEICH ZU CUED RETROSPECTIVE THINK ALOUD (CRTA). | 50 |
| TABELLE 6: AUFGABENSTELLUNGEN GEMÄß TEILKOMPETENZ, MODELLEINSATZ UND TREE-READING-FÄHIGKEIT..... | 56 |
| TABELLE 7: INTERCODERRELIABILITÄT FÜR DEN MEDIALEN UND METHODISCHEN MODELLEINSATZ DER TEILKOMPETENTEN ZWECK VON MODELLEN UND ALTERNATIVE MODELLE | 65 |
| TABELLE 8: ZUORDNUNG DER 11 BETTSI-ITEMS ZU TREE-READING-FÄHIGKEITEN SOWIE ITEM-SKALA-KORRELATION UND CRONBACHS A..... | 69 |
| TABELLE 9: KONTROLLVARIABLEN „HÄUFIGKEIT ARBEIT MIT DPV“ UND „WISSENSCHAFTLICH ADÄQUATES LESEN VON DPV“ (SELBSTEINSCHÄTZUNG), TREE-READING-FÄHIGKEITEN (BETTSI), RÄUMLICHES DENKEN (3DW) UND KOGNITIVE BEARBEITUNGSGESCHWINDIGKEIT (ZVT)..... | 71 |
| TABELLE 10: KONTROLLVARIABLE MODELLEINSATZ JE ERFAHRUNGSGRAD. | 72 |
| TABELLE 11: VON DER ANALYSE AUSGESCHLOSSENE TRANSKRIPTE JE AUFGABE UND ERFAHRUNGSGRAD..... | 72 |
| TABELLE 12: ANZAHL SINNEINHEITEN UND PROBAND:INNEN (N) JE AUFGABE UND ERFAHRUNGSGRAD..... | 73 |
| TABELLE 13: INTERCODERÜBEREINSTIMMUNG JE MODELLEINSATZ..... | 74 |
| TABELLE 14: INDIVIDUELLE LÖSUNGSPROZESSE BZW. LÖSUNGEN DER AUFGABEN. | 91 |
| TABELLE 15: REPRÄSENTATIVE BEISPIELE FÜR LÖSUNGSPROZESSE..... | 92 |
| TABELLE 16: INDIVIDUELLE LÖSUNGEN BZW. LÖSUNGSPROZESSE GEORDNET NACH VERSUCHSVARIANTEN 1 (V1) UND 2 (V2)..... | 122 |
| TABELLE 17: ABSOLUTE HÄUFIGKEITEN VON SINNEINHEITEN DER TREE-READING-FÄHIGKEIT WISSENSCHAFTLICH NICHT ADÄQUATES ABLESEN (TR_1B) PRO AUFGABE | 123 |
| TABELLE 18: ZUORDNUNG DER KONTROLLVARIABLEN ZU INDIVIDUELLEN LÖSUNGSPROZESSEN BZW. LÖSUNGEN. | 125 |

Zusammenfassung

Diagramme phylogenetischer Verwandtschaft (DpV) modellieren Hypothesen zu evolutiven Verwandtschaftsverhältnissen verschiedener Arten oder Populationen. Das *wissenschaftlich adäquate* Lesen, Interpretieren und Vergleichen von DpV wird als Tree-Reading bezeichnet und ist essentiell für das Evolutionsverständnis. In der Wissenschaft werden DpV als Modelle *von* und *für* evolutive Verwandtschaft verwendet, um Wissen zu kommunizieren (medialer Modelleinsatz) und neue Erkenntnisse zu generieren (methodischer Modelleinsatz). Im Biologieunterricht findet oft nur der mediale Einsatz von Modellen statt. Sowohl im Berliner Rahmenlehrplan als auch in den aktuellen Bildungsstandards Biologie für die Allgemeine Hochschulreife werden DpV zwar explizit erwähnt, die konkrete Umsetzung des Themas obliegt jedoch den Schulen und den einzelnen Lehrkräften und erfolgt erst ab der 10. Jahrgangsstufe.

Während Wissenschaftler:innen mit einschlägigem Forschungsschwerpunkt umfassende Tree-Reading-Fähigkeiten zeigen, attestieren verschiedene Studien Lernenden schwach ausgeprägte Tree-Reading-Fähigkeiten mit überwiegend *oberflächlichen* oder *wissenschaftlich inadäquaten* Aussagen zu DpV. Dabei wird meistens untersucht, *ob* Lernende Tree-Reading-Aufgaben lösen können und wie sich Tree-Reading-Fähigkeiten strukturieren lassen. Es ist jedoch wenig darüber bekannt, *inwiefern* Lernende DpV lesen und interpretieren. Die vorliegende Studie untersucht daher, inwiefern der Erfahrungsgrad mit DpV (hoch versus niedrig) und der Modelleinsatz (medial versus methodisch) bei Tree-Reading-Prozessen interagieren. Zur Charakterisierung von Tree-Reading-Prozessen wurden zwei Erfahrungsgruppen gegenübergestellt: Wissenschaftler:innen mit hohem Erfahrungsgrad mit DpV (HoEr; $n = 10$) und Studierende des *Master of Education* Biologie mit niedrigem Erfahrungsgrad mit DpV (NiEr; $n = 15$). Anhand der Teilkompetenzen *Zweck von Modellen* und *Alternative Modelle* des Kompetenzmodells der Modellkompetenz (Upmeyer zu Belzen & Krüger, 2010) wurden Aufgaben generiert, in denen DpV entweder im medialen oder im methodischen Modelleinsatz eingebettet waren. Für die umfassende Charakterisierung der Tree-Reading-Prozesse wurden sowohl verbale als auch visuelle Daten erhoben, indem die Aufgaben auf einem Bildschirm mit integriertem *Remote-Eyetracker* gezeigt wurden und die Proband:innen zu parallelem Lauten Denken (*Concurrent* bzw. *Congruent Think Aloud*, CTA) aufgefordert wurden.

Mittels einer qualitativen Inhaltsanalyse (Mayring, 2010) der Lösungsprozesse beider Erfahrungsgruppen wurde kodiert, ob Proband:innen anhand der DpV-Bestandteile Informationen *ableiteten* (TR_2) oder sie *wissenschaftlich adäquat* (TR_1a), *wissenschaftlich inadäquat* (TR_1b) oder *oberflächlich* (TR_0) *ablasen*. Bei der Teilkompetenz *Alternative Modelle* wurde zusätzlich kodiert, ob zwei DpV miteinander verglichen oder einzeln beschrieben wurden. Außerdem wurde kodiert, ob eine Aufgabenlösung *den wissenschaftlichen Konventionen entsprach*, davon *abwich* oder es *keine Lösung* gab.

HoEr zeigten sowohl im medialen als auch im methodischen Modelleinsatz fast ausschließlich die für eine Aufgabe erforderlichen globalen und maximalen Tree-Reading-Fähigkeiten und lösten die Aufgaben im methodischen Modelleinsatz *den wissenschaftlichen Konventionen entsprechend*. Bei NiEr ergab sich ein differenzierteres Bild: Im medialen Modelleinsatz zeigten NiEr hinsichtlich der globalen Tree-Reading-Fähigkeiten häufig *oberflächliches Ablesen* (Teilkompetenz *Zweck von Modellen*) bzw. *wissenschaftlich inadäquates Ablesen* (Teilkompetenz *Alternative Modelle*). Dies spiegelte sich auch in den maximal erreichten Tree-Reading-Fähigkeiten in der Teilkompetenz *Zweck von Modellen* wider. Im methodischen Modelleinsatz zeigten NiEr in beiden Teilkompetenzen häufig *von den wissenschaftlichen Konventionen abweichende* Lösungen bzw. ihre Lösungen konnten nicht kodiert werden.

Der interindividuelle Vergleich der Lösungsprozesse über alle Aufgaben hinweg ergab, dass alle HoEr und 2 NiEr keine Abweichungen von den maximal erforderlichen Tree-Reading-Fähigkeiten bzw. *wissenschaftlichen Konventionen* zeigten, 2 weitere NiEr nur im medialen Modelleinsatz. 4 NiEr zeigten Abweichungen im methodischen Modelleinsatz, 6 NiEr sowohl im medialen als auch methodischen Modelleinsatz. Durch die exemplarische Analyse repräsentativer Transkripte konnten häufige Vorgehensweisen in den Lösungsprozessen von HoEr und NiEr illustriert werden wie bspw. das Beschreiben von *MRCA* durch *oberflächliches* Abzählen von Knotenpunkten. Die Triangulation von visuellen und verbalen Daten ergab, dass verbale Bezüge und Fixationen übereinstimmten und kann als weiteres Maß für Validität herangezogen werden.

Aus den Ergebnissen lässt sich schlussfolgern, dass neben dem Modelleinsatz auch Merkmale der Repräsentationen, die Reihenfolge der Aufgaben sowie die Fachsprache einen Einfluss auf die Tree-Reading-Prozesse von NiEr hatten. Es muss jedoch betont werden, dass NiEr sich innerhalb ihrer Erfahrungsgruppe hinsichtlich ihrer Tree-Reading-Prozesse stark unterscheiden und einige anhand der untersuchten Kriterien nicht oder kaum von HoEr zu

unterscheiden sind. Die in der vorliegenden Studie beschriebenen heterogenen Tree-Reading-Prozesse von NiEr lassen sich mit dem *Revised STREAM* (Schramm, Jose, & Schmiemann, 2021a) erklären, das Tree-Reading-Fähigkeiten in fünf Dimensionen einteilt, die von Lernenden in unterschiedlicher Ausprägung beherrscht werden können. Hieraus ergibt sich die Implikation, Tree-Reading-Fähigkeiten von Lernenden individuell zu fördern, um so das Evolutionsverständnis und damit das Verständnis gesellschaftsrelevanter Themen zu stärken. Die vorliegende Studie bietet erstmals umfassende Einblicke in das *Inwiefern* von Tree-Reading und ist damit Ausgangspunkt für tiefergehende Untersuchungen der herausgearbeiteten Aspekte.

Abstract

Phylogenetic trees (PT) model hypotheses about evolutionary relationships among species or populations. Interpreting and comparing PT is defined as tree reading and is essential for understanding evolution. PT as models are used in science to communicate knowledge (medial perspective) and to generate new insights (methodical perspective). In biology classrooms, model use is often limited to the medial perspective. While experienced scientists exhibit scientifically adequate tree reading, learners often show superficial and scientifically inadequate tree reading. Based on verbal and visual data the present study investigates how the degree of experience with PT and the perspective from which models are used interact in tree reading processes. Scientists with a high degree of experience with PT ($n = 10$) and pre-service biology teachers with low degree of experience ($n = 15$) solved tasks with medially and methodically embedded PT, respectively. Pre-service teachers showed heterogeneous tree reading processes, which sometimes did not differ from scientists' processes. This is in line with the *Revised STREAM* (Schramm, Jose, & Schmiemann, 2021) dividing tree reading into five dimensions that learners can master on differing levels. Corresponding to these findings, learners' tree reading skills should be fostered individually to strengthen understanding of evolution and socially relevant issues. The present study provides unique and comprehensive insights into tree reading processes and thus represents a starting point for more profound investigation of the found aspects.

1. Einleitung

Es ist beinahe eine Tradition geworden, eine wissenschaftliche Arbeit mit einem passenden Zitat zu beginnen. Beschäftigt man sich im weitesten Sinne mit dem Thema Evolution, stößt man eher früher als später auf den berühmten Ausspruch von Theodosius Dobzhansky (1973, S. 125): „*Nothing in biology makes sense except in the light of evolution*“. Eine Aussage, die sehr direkt gegen Kreationismus und Intelligent Design gerichtet war und - liest man den Anfang des gleichnamigen Papers (1973) - politischer ist, als man es losgelöst vom Inhalt vermuten mag.¹ Unabhängig von Dobzhanskys Intention passt dieses Zitat eher in eine Einleitung zum Thema Einstellungen zu Evolution.² Daher wird für die hier einzuleitende Arbeit ein Zitat von Metrick und Weitzman (1998, S. 25) gewählt:

„Tree thinking“ represents a prime example of how one picture may be worth a thousand words.“

Doch was haben zwei Ökonomen mit Tree-Thinking zu tun? Metrick und Weitzman erläutern in ihrer Publikation die Frage, inwiefern man ökonomisch und moralisch entscheiden kann, welche Spezies im Sinne der Biodiversität vor der Ausrottung bewahrt werden sollte. Dabei nutzen sie die Arche Noah als Metapher für die Rettung der Biodiversität und das *Evolutionary Library Model* als das der Biodiversität zugrundeliegende Prinzip der gemeinsamen Abstammung und Aufspaltung in verschiedene Spezies. Verschiedene Bibliotheken symbolisieren Kladen, in denen zahlreiche Spezies als Bücher eingeordnet sind:

„The evolutionary library model naturally generates a corresponding evolutionary tree. When a tree structure is present, it seems to induce a way of visualizing and comprehending intuitively relationships among objects that are quite subtle or complicated to describe without the tree. “Tree thinking“ represents a prime example of how one picture may be worth a thousand words.“

(Metrick & Weitzman, 1998, S. 25)

¹ Dobzhansky schildert sehr deutlich, inwiefern er nicht mit der Sicht eines Scheichs über die Entstehung der Natur einverstanden ist (vgl. Dobzhansky, 1973, S. 25).

² Ausführlich untersucht und erläutert in Beniermann (2019). Dobzhanskys Zitat wird hier nicht verwendet.

Metrick und Weitzman nutzen die Baum-Metapher als Versuch einer Grundlage für die Entscheidung, welche Spezies schützenswerter ist als andere. Ihre Intention ist es, dieses komplexe und multikausale Problem herunterzubrechen, um auf sachlicher Ebene und ohne „*desire to have "charismatic megafauna" effects*“ entscheiden zu können (vgl. Metrick & Weitzman, 1998, S. 33). Für die vorliegende Arbeit wurde zunächst die kürzere, zuerst genannte Version des Zitats von Metrick und Weitzman gewählt, da sie den Kern von Diagrammen phylogenetischer Verwandtschaft (DpV; Erläuterung des Begriffs in Kapitel 2.1) illustriert: Die unvorstellbare Komplexität von evolutiver Verwandtschaft, abstrahiert durch Linien, die oberflächlich gesehen einfache Verzweigungsmuster bilden. Sieht man das Zitat im erweiterten Kontext, wird die Relevanz des Themas sichtbar, nämlich das Verständnis von evolutiver Verwandtschaft als Voraussetzung für das Verständnis von Problemen, die unseren Alltag mehr oder weniger direkt betreffen: Pandemien, Artensterben, der Mensch selbst als Evolutionsfaktor oder persistierender pseudowissenschaftlich begründeter Rassismus, um nur einige zu nennen (vgl. Nationale Akademie der Wissenschaften Leopoldina, 2017). Curricula von Schulen und Universitäten in Deutschland spiegeln die Relevanz von Evolution als „das übergreifende, vereinigende Erkenntnisprinzip der Biologie“ (Nationale Akademie der Wissenschaften Leopoldina, 2017, S. 8) jedoch noch immer nicht in ausreichendem Maße wider (vgl. ebd., S. 11).

Ein seit Frühjahr 2020 hochrelevantes Anwendungsgebiet ist die Analyse der evolutiven Verwandtschaft verschiedener SARS-CoV-2-Varianten. Das Projekt *Nextstrain* bspw. nutzt verfügbare Forschungsdaten, um DpV der verschiedenen Varianten sowie deren Verteilung in verschiedenen Regionen zu analysieren (Abbildung 1) und für weitere Analysen zur Verfügung zu stellen (Nextstrain.org).³ Die Analysen können dann z. B. für die Ermittlung des Virusursprungs oder frühzeitige Reaktionen auf Mutationen genutzt werden (Jütte, 2021). Forschende riefen jedoch gleichzeitig dazu auf, vorsichtig mit der Interpretation von solchen Modellierungen zu sein, da die Mutationseigenschaften von SARS-CoV 2 keine geeignete Grundlage bilden, um verlässliche Hypothesen über die Verwandtschaft der einzelnen Varianten zu generieren (z. B. Morel et al., 2021).

³ Nextstrain.org stellt Genom-Analysen zu epidemiologisch Forschungszwecken auch für andere Pathogene wie *Yersinia pestis* oder Ebola bereit.

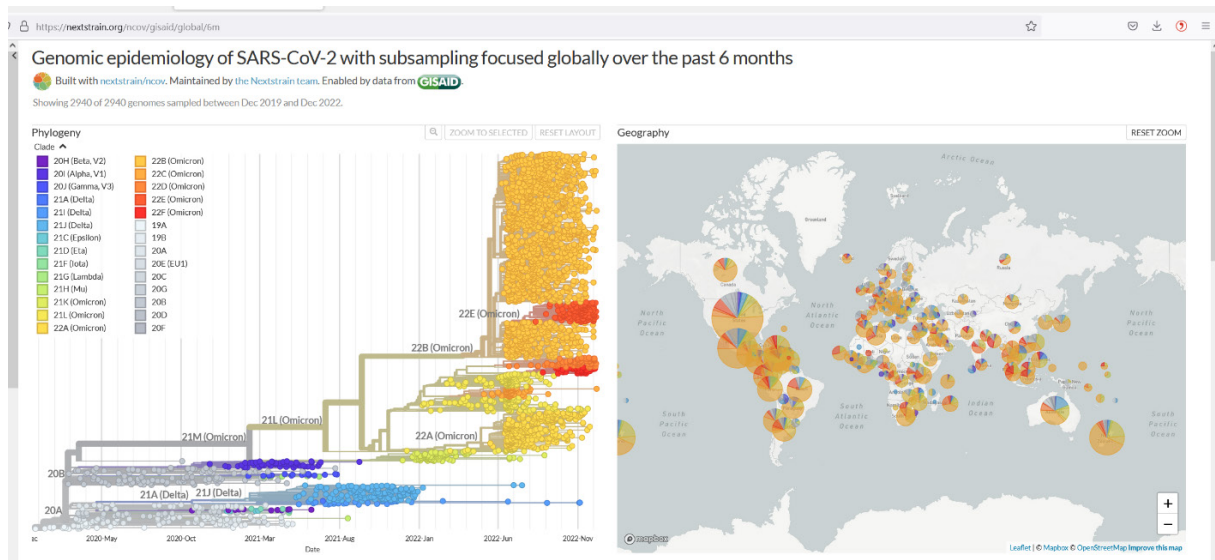


Abbildung 1: Analyse der evolutionen Verwandtschaft der verschiedenen SARS-CoV-2-Varianten als aktuelles Beispiel der Relevanz von DpV. Screenshot von Nextstrain.org.

Das Nutzen von Modellen zur Illustration von Sachverhalten, aber auch zur Erkenntnisgewinnung ist ebenfalls zentral bei Metrick und Weitzman (1998), wodurch das anfangs genannte Zitat einen weiteren Aspekt der vorliegenden Arbeit anspricht und somit doppelt passt. Die beiden nutzen zunächst das *Evolutionary Library Model*, um die Eigenschaften evolutioner Verwandtschaft zu illustrieren und verständlich zu machen. Sie brechen eben dieses Modell herunter auf DpV als abstrakteres Modell für den komplexen Sachverhalt. Darüber hinaus modellieren sie noch die Entscheidung darüber, welche Spezies zu erhalten sind, in Form einer mathematischen Gleichung (vgl. Metrick & Weitzman, 1998). Ebenso kompliziert wie für Außenstehende solche mathematischen Modellierungen wirken, sind DpV als Hypothesen und somit als Modelle *von* und *für* evolutione Verwandtschaft (Mahr, 2008, 2009; Passmore et al., 2014). Für Wissenschaftler:innen mit einschlägigem Forschungsinteresse dienen sie als Arbeitsmittel (Baum et al., 2005; Halverson & Friedrichsen, 2013; O’Hara, 1989), für die meisten Lernenden und Lehrenden sind sie jedoch kaum durchschaubar und werden häufig *wissenschaftlich inadäquat abgelesen* (Baum et al., 2005; Catley, Phillips, & Novick, 2013; Halverson & Friedrichsen, 2013; Meir et al., 2007; Morabito et al., 2010; Novick & Catley, 2013, 2014). So essentiell das Verständnis von Evolution für das Verständnis der Biologie ist (Nationale Akademie der Wissenschaften Leopoldina, 2017; van Dijk & Kattmann, 2010), so ist es das Lesen und Interpretieren von DpV für das Evolutionsverständnis (Baum, DeWitt Smith, & Donovan, 2005; Baum & Smith, 2013; Gregory, 2008; Jördens, Asshoff, & Kullmann, 2011; Meir, Perry, Herron, & Kingsolver, 2007; O’Hara, 1988). Mehrere Studien haben sich bereits damit beschäftigt, *ob* Lernende bestimmte Tree-

Reading-Fragestellungen beantworten können (z. B. Dees, Bussard, & Momsen, 2018; Dees, Freiermuth, & Momsen, 2017; Meir et al., 2007; Novick & Catley, 2013, 2014, 2016), jedoch nicht *inwiefern* sie dies tun. Verschiedene Ansätze klassifizieren Tree-Reading-Fähigkeiten nach Anspruch (Halverson & Friedrichsen, 2013; Kummer, 2017; Kummer, Whipple, Bybee, Adams, & Jensen, 2019; Novick & Catley, 2013, 2016; Schramm, Jose, & Schmiemann, 2021a; Schramm, Schachtschneider, & Schmiemann, 2019), *Prozesse* beim Tree-Reading werden jedoch auch hier nicht beschrieben. Der Umgang mit DpV als Modelle *von* und *für* Evolution wurde darüber hinaus noch nie untersucht, obgleich sie von Wissenschaftler:innen als solche selbstverständlich eingesetzt werden (Baum et al., 2005; Halverson & Friedrichsen, 2013; O'Hara, 1989). Die vorliegende Arbeit hat zum Ziel, den Umgang mit ebendiesen Modellen im medialen und methodischen Modelleinsatz anhand verbaler und visueller Daten zu untersuchen und dabei zu charakterisieren, inwiefern sich Tree-Reading-Prozesse bei Personen mit unterschiedlichem Erfahrungsgrad (hoch versus niedrig) unterscheiden. Daraus sollen Implikationen gewonnen werden, inwiefern Lernende im Umgang mit DpV im Sinne der Modellkompetenz geschult werden können und somit ihr Evolutionsverständnis langfristig verbessert werden kann.

2. Theorie

2.1. Von der gemeinsamen Abstammung zu Diagrammen phylogenetischer Verwandtschaft

Bei einer Recherche zum Thema Evolution bzw. gemeinsamer Abstammung stößt man unweigerlich auf den Namen Charles Darwin. Er gilt als der Begründer der modernen Evolutionstheorie, indem er aus Beobachtungen und bekannten Fakten die wegweisenden Schlüsse ableitete (Purves, Sadava, Orians, & Heller, 2006).⁴ Bis dahin herrschte die Annahme vor, es gäbe höhere und niedrigere Organismen,⁵ die unterschiedliche Ursprünge hätten und niedrigere Organismen lediglich die Vorformen der höheren wie bspw. Säugetiere seien (*Great Chain of Being*).⁶ Die Idee der gemeinsamen Abstammung keimte bereits bei Charles Lyell und Erasmus Darwin, Charles Darwins Großvater, auf, wurde jedoch erst von Charles Darwin in Form der heute weltbekannten Baummetapher modelliert (vgl. Baum & Smith, 2013, S. 9 ff). Die gemeinsame Abstammung aller Organismen setzt dabei evolutive Prozesse voraus (Baum & Smith, 2013, S. 14).

Auf der Basis von Charles Darwins Baummetapher (Darwin, 1837; Notebook B, entry 36) wurden verschiedene Repräsentationen phylogenetischer Verwandtschaft auf Grundlage strenger inhaltlicher Konventionen entwickelt (Kapitel 2.1.1). In der Bezeichnung ebendieser standardisierten Repräsentationsformen herrscht allerdings seit jeher wenig Einigkeit und es koexistieren verschiedene Begrifflichkeiten für dieselbe Repräsentationsform bzw. bestimmte Termini werden für unterschiedliche Repräsentationsformen gebraucht (Ax, 1984). Ax schlägt daher als Oberbegriff für alle Repräsentationsformen phylogenetischer Verwandtschaft den Terminus „Diagramme phylogenetischer Verwandtschaft“ vor, da dieser seiner Ansicht nach am treffendsten beschreibt, „was sie [die graphische Darstellung] nun einmal ist, - ein Diagramm der phylogenetischen Verwandtschaft oder ein phylogenetisches Verwandtschaftsdiagramm“ (Ax, 1984, S. 53). Scholtz und Scholtz (2019) brechen mit dem

⁴ Weniger bekannt aber genauso an der revolutionären Evolutionstheorie beteiligt war Alfred Russell Wallace (vgl. Nationale Akademie der Wissenschaften Leopoldina, 2017).

⁵ In Form des Sozialdarwinismus wird jedoch auch Darwins Theorie der natürlichen Selektion als Argument genutzt, um die Unterdrückung von bestimmten Menschengruppen vermeintlich wissenschaftlich zu begründen (vgl. Lenzen, 2015).

⁶ Aus dieser Annahme ergeben sich ebenfalls *wissenschaftlich inadäquate* Leseweisen bzw. Vorstellungen zum Tree-Reading (Abschnitt 2.1.4).

Vorschlag „dichotome Verzweigungsdiagramme“ die Bezeichnung für Kladogramme sogar auf die rein oberflächliche Repräsentationsform runter.

Neben den nicht trennscharfen Termini in der Wissenschaftswelt fällt ein weiterer Faktor im Alltagsleben ins Gewicht. So ist in englischsprachigen Publikationen meist von *Phylogenetic Trees* die Rede (z. B. Baum & Smith, 2013), in deutschsprachigen Publikationen (z. B. Hammann & Scheffel, 2005; Jördens et al., 2011) aber auch im allgemeinen Sprachgebrauch von (evolutiven) Stammbäumen. Diese Bezeichnung impliziert die historische Baum-Metapher, die aber vermutlich nicht von Darwin intendiert war und sich erst nachträglich als eine Art der Veranschaulichung entwickelte (vgl. Hellström, 2012, S. 236). Aufgrund der ähnlichen Metapher einer Wurzel und ihrer Äste und der Begriffsüberschneidung erfolgt häufig eine Assoziation mit Familienstammbäumen, also die Abstammungsgeschichte einzelner Individuen (Knoop & Müller, 2009, S. 209). Im Rahmenlehrplan Biologie 7-10 in Berlin wird z. B. der Begriff „Stammbaum“ analog für Familienstammbäume für die Analyse von Erbkrankheiten als auch für DpV kommentarlos benutzt (Senatsverwaltung für Bildung, Jugend und Sport Berlin, Ministerium für Bildung, Jugend und Sport des Landes Brandenburg, & Ministerium für Bildung, Wissenschaft und Kultur Mecklenburg-Vorpommern, 2015, S. 36 ff.). Um einen einheitlichen Oberbegriff zu nutzen und zu verdeutlichen, dass es nicht um diese individuelle Abstammung, sondern die phylogenetische Verwandtschaft von Gruppen von Organismen geht, wird in der vorliegenden Arbeit demnach der Begriff *Diagramm phylogenetischer Verwandtschaft* nach Ax (1984) verwendet und im Folgenden mit DpV abgekürzt.

2.1.1. Diagramme phylogenetischer Verwandtschaft - Terminologie

DpV sind Hypothesen zu evolutiven Verwandtschaftsverhältnissen verschiedener Taxa (Baum & Smith, 2013; Eldredge & Cracraft, 1980; Gregory, 2008; Kong & Anderson, 2016; Novick & Catley 2007, 2014). Der für die vorliegende Arbeit gewählte Oberbegriff DpV umfasst zahlreiche spezielle DpV-Formen, die je nach wissenschaftlicher Fragestellung zur evolutiven Verwandtschaft von Organismen unterschiedliche inhaltliche Schwerpunkte setzen. Ein Phylogramm bspw. ordnet Aufspaltungsereignisse anhand einer konkreten Skala ein (absolute Zeit, genetischer Abstand o. ä.).

Kladogramme legen als „dichotomes Verzweigungsdiagramm“ (Scholtz & Scholtz, 2019) den Fokus auf die reine Repräsentation hypothetischer Verwandtschaftsverhältnisse ohne konkrete Angaben zu zeitlichen oder genetischen Abständen (Baum & Smith, 2013).

Stellvertretend für verschiedene DpV-Formen werden im Folgenden Bestandteile und Bedeutung von Kladogrammen beschrieben, da diese in der vorliegenden Arbeit verwendet wurden.⁷ Kladogramme umfassen als Grundelemente oberflächlich betrachtet Linien, die sich, ausgehend von einer gemeinsamen Wurzel, an Knotenpunkten dichotom⁸ aufspalten. Die dichotome Aufspaltung kann u. a. „baumförmig“ (Abbildung 2a & b) oder „leiterförmig“ (Abbildung 2c) dargestellt werden, ohne dass sich die inhaltliche Bedeutung ändert (Baum & Smith, 2013). Inhaltlich repräsentieren die terminalen Linien Taxa, also rezente⁹ Gruppen von Organismen wie Populationen, Gattungen, Arten o. ä. Die sie verbindenden Knotenpunkte stellen hypothetische Aufspaltungsereignisse dar, die zu diesen Taxa geführt haben. Die letzten gemeinsamen Vorfahren bestimmter Taxa, auch *most* oder *more recent common ancestors (MRCA)*, werden je nach Konvention direkt am Knotenpunkt (z. B. Baum & Smith, 2013) oder in den Linien verortet, die von der Wurzel aus gesehen darauf zuführen (Morabito, Catley, & Novick, 2010; Abbildung 2a). Alle Taxa, die von einem bestimmten MRCA abstammen, bilden mit ihm zusammen eine Klade bzw. ein Monophylum (z. B. U, L und I inkl. MRCA; Abbildung 2b). Ein Paraphylum wiederum umfasst den MRCA mit nicht allen von ihm abstammenden Taxa (z. B. nur I und U in Abbildung 2).

⁷ Für nähere Informationen zu weiteren DpV-Repräsentationsformen siehe z. B. Baum und Smith (2013), Gregory (2008), Losos (2014) und Weigand (2015).

⁸ Wenn evolutive Aufspaltungsereignisse noch nicht eindeutig geklärt sind, kann dies in Kladogrammen mithilfe von Polytomien, also einer Aufspaltung in mehr als 2 Linien, dargestellt werden (Baum & Smith, 2013).

⁹ Auch fossile Taxa können als terminale Gruppen dargestellt werden: „*They are treated as living forms that have undergone no evolution in millions of years since they were entombed in rock.*“ (Baum & Smith, 2013, S. 41).

Die Taxa, deren Verwandtschaftsverhältnisse Gegenstand eines DpV sind, bilden die *Ingroup* (Taxa A, I, L, U in Abbildung 2). Um das DpV zu „wurzeln“, wird ihr eine *Outgroup* (Taxon P in Abbildung 2) entgegengestellt. Die Wurzel zeigt also das hypothetische initiale Aufspaltungsereignis von *Ingroup* und *Outgroup*. Die relative Zeitachse verläuft von dieser Wurzel (Vergangenheit) bis hin zu den Taxa (Gegenwart). Die Entstehung oder der Verlust von Merkmalen wird entlang der Entwicklungslinien von Taxa bzw. Kladen verortet (Abbildung 2a).

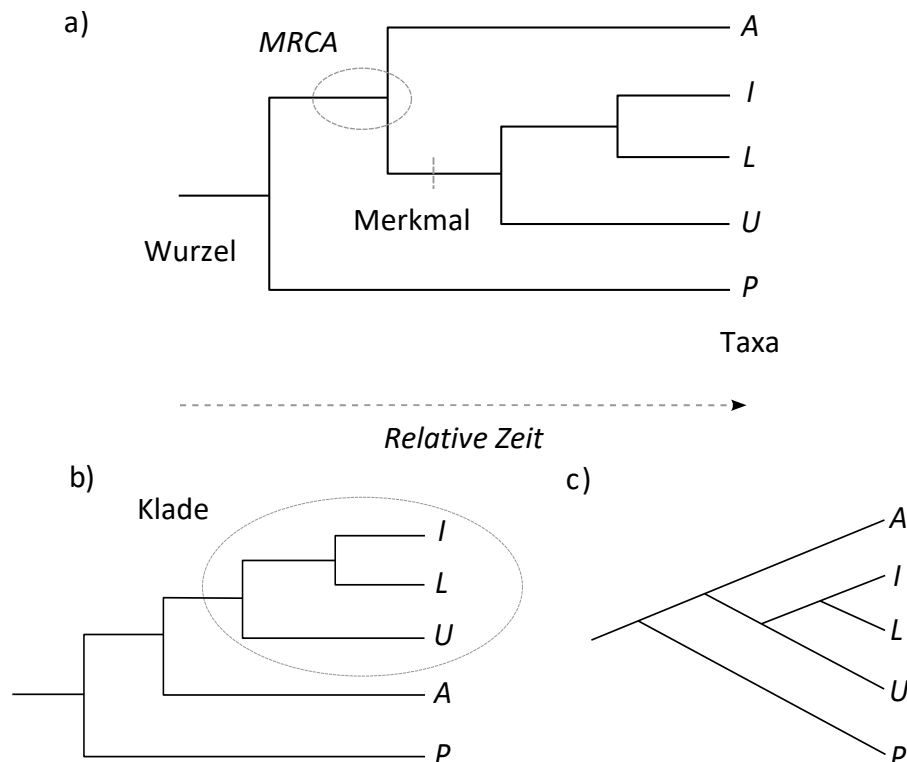


Abbildung 2: Aufbau und Bestandteile von Diagrammen phylogenetischer Verwandtschaft am Beispiel eines Kladogramms der Taxa P, A, U, L und I. a) Baumförmiges DpV, die relative Zeitachse verläuft von der Wurzel in Richtung Taxa. MRCA = Most Recent Common Ancestor, Merkmal = Entstehung eines neuen Merkmals entlang der Entwicklungslinie. b) und c) stellen 2 alternative Repräsentationsformen der Verwandtschaftsverhältnisse aus a) dar: In b) wurden die Taxa am Knotenpunkt der Klade A+I+L+U rotiert; in c) ist das DpV leiterförmig dargestellt. In b) ist die Klade mit den Taxa U, L und I und ihrem MRCA markiert.

Anzestrals Merkmale, also Merkmale des MRCA, werden als Plesiomorphie bezeichnet. Das Gegenteil davon ist ein abgeleitetes Merkmal, eine Apomorphie. Begründet eine Apomorphie eine neue Klade, wird sie als Synapomorphie bezeichnet. Teilen mehrere paraphyletische Taxa eine Plesiomorphie, wird diese Sympleiomorphie genannt. Besitzt nur ein einziges Taxon ein abgeleitetes Merkmal, handelt es sich um eine Autapomorphie (Baum & Smith, 2013). Das Sparsamkeitsprinzip (auch Parsimonie) wird angewendet, um Hypothesen über evolutive Verwandtschaftsbeziehungen mit der kleinsten Anzahl an Verlust- bzw.

Entstehungsereignissen von Merkmalen zu erklären (Baum & Smith, 2013). In der vorliegenden Arbeit wird für alle Formen der Apomorphien der Begriff „Merkmal“ genutzt, da dieser auch für die leichtere Verständlichkeit in den Aufgaben eingesetzt wurde.

2.1.2. Tree-Thinking und Tree-Reading

Der Begriff Tree-Thinking wurde 1988 von O’Hara als Gegenstück zum *population thinking* geprägt, indem er betonte, dass Fragen nach evolutiven Prozessen nur verstanden werden können, wenn man den gesamten evolutiven Verlauf einer Stammlinie betrachtet, statt sich auf die zu untersuchenden Taxa bzw. Kladen zu beschränken (O’Hara, 1988, 1997). Verschiedene Definitionen des Begriffs unterscheiden sich hinsichtlich ihrer Denkebene. So definieren Catley und Novick Tree-Thinking als „*a habit of mind that uses the history of life on Earth as its first line of evidence while providing students with a hierarchical view of the natural world.*“ (Catley & Novick, 2008, S. 977), später konkreter als „*ability to understand and reason with the information depicted in cladograms*“ (Novick & Catley, 2014, S. 271). Außerdem wird Tree-Thinking konkreter als Fähigkeit definiert, einerseits Informationen aus DpV ablesen und diese interpretieren zu können und andererseits generierte Informationen in Form von DpV pointiert zu repräsentieren (Baum & Smith, 2013; Halverson, 2011) sowie zahlreiche DpV unabhängig von ihrer Repräsentationsform miteinander zu vergleichen (Halverson & Friedrichsen, 2013). Hierbei beeinflusst Wissen über wissenschaftliche Methoden der Kladistik Tree-Thinking positiv (Baum & Smith, 2013). Halverson unterteilt Tree-Thinking in die Komponenten Tree-Reading (TR) und Tree-Building, wobei Tree-Reading das Interpretieren und Vergleichen mehrerer DpV und Tree-Building das Erstellen von DpV umfasst (Halverson, 2011). Sie postuliert, dass Tree-Reading als Basis für das kognitiv anspruchsvollere Tree-Building dient und das Beherrschen von beiden notwendig ist, um ein *Expert Tree Thinker* zu sein (Halverson, 2011). Auch Kummer (2017) bestätigt, dass Tree-Thinking insgesamt besser beherrscht wird, wenn zuerst Tree-Reading erlernt wird. Eddy und Kolleg:innen (2013) stellen dieser Aussage jedoch entgegen, dass Tree-Building-Aktivitäten das initiale Erlernen von Tree-Reading fördern können. Abgesehen von einigen Autor:innen (z. B. Eddy, Crowe, Wenderoth, & Freeman, 2013; Halverson, Pires, & Abell, 2011; Schramm et al. 2019, 2021a) wird nicht explizit von Tree-Reading gesprochen, sondern der Begriff Tree-Thinking sowohl für das Interpretieren als auch das Konstruieren von DpV genannt.

2.1.3. Diagramme phylogenetischer Verwandtschaft - *Wissenschaftlich adäquate* Leseweise

Die *wissenschaftlich adäquate* Leseweise von DpV wird wie in Kapitel 2.1.1 anhand von Kladogrammen als Stellvertreter für DpV erläutert. Kladogramme modellieren die Topologie, also das Verzweigungsmuster der evolutiven Verwandtschaft der enthaltenen Taxa. Um Aussagen über evolutive Verwandtschaft machen zu können, kann zunächst die kleinste Verwandtschaftseinheit des Kladogramms betrachtet werden, das Schwestertaxon. Dieses umfasst 2 Taxa, die vom selben *MRCA* abstammen und diesen aufgrund des dichotomen Verzweigungsmusters mit keinen weiteren Taxa teilen (Abbildung 2, Taxa *I* und *L*), diese beiden Taxa bilden eine Klade. Das Schwestertaxon zur Klade *I+L* ist wiederum *U*, gemeinsam mit ihrem *MRCA* bilden *I+L+U* abermals eine Klade (Abbildung 2b). So können alle Taxa des Kladogramms in Kladen eingeteilt werden, die jeweils ein geschlossenes, monophyletisches Verwandtschaftssystem zeigen (*nested taxonomy* bzw. *nested clades*). Dies zeichnet sich dadurch aus, dass man den *MRCA* und alle von ihm abstammenden Taxa durch einen einzigen „Schnitt“ vom restlichen DpV abtrennen könnte. Um Aussagen über Gemeinsamkeiten und Unterschiede hinsichtlich der Topologie mehrerer DpV zu treffen, kann ein Vergleich der jeweiligen Kladen herangezogen werden (Baum & Smith, 2013).

Die Topologie ist grundsätzlich von der oberflächlichen Darstellung von DpV zu unterscheiden. Unabhängig davon, ob ein DpV baumförmig oder leiterförmig dargestellt wird oder in welcher Reihenfolge die Taxa angeordnet sind, sind zur Einteilung in Kladen nur die Knotenpunkte zu betrachten. Die von ihnen ausgehenden Äste bzw. Taxa können beliebig um diesen Knotenpunkt rotiert werden (Abbildung 2b; Gregory, 2008; Baum & Smith, 2013). Die Knotenpunkte bzw. die durch sie repräsentierten *MRCA* bzw. Aufspaltungsereignisse stellen wiederum den Bezugspunkt für den Vergleich relativer Verwandtschaftsverhältnisse dar. Da Taxon *I* und *L* von einem jüngeren *MRCA* abstammen, sind sie näher miteinander verwandt als *I* und *A* (Abbildung 2a).

Die Entstehung eines neuen Merkmals (Synapomorphie) begründet die Aufspaltung von Entwicklungslinien. Da der genaue Zeitpunkt einer Merkmalsentstehung nicht festgelegt werden kann bzw. Kladogramme lediglich den relativen Zeitverlauf darstellen, werden Merkmale entlang der Entwicklungslinien angegeben, deren Abspaltung durch sie begründet ist und nicht direkt an den Knotenpunkten.

Anhand dieser Angaben kann aus dem DpV effizient abgelesen werden, welche Taxa bzw. Kladen welche Merkmale aufweisen: „*The tree is a much more efficient way to store such information than trying to recall the full list of traits that occur in a particular organism*“ (Baum & Smith, 2013, S. 4). Konvergente Merkmalsentstehungen können so ebenfalls visualisiert werden (Baum & Smith, 2013).

2.1.4. Diagramme phylogenetischer Verwandtschaft - *Wissenschaftlich inadäquate* Leseweise

Während Wissenschaftler:innen DpV *wissenschaftlich adäquat* interpretieren können, um aus ihnen Schlussfolgerungen über evolutive Verwandtschaftsverhältnisse zu ziehen, fällt Lernenden wie Schüler:innen und Studierenden Tree-Reading häufig schwer bzw. sie zeigen keine *wissenschaftlich adäquaten* Tree-Reading-Fähigkeiten bzw. Tree-Reading-Performanzen (z. B. Baum et al., 2005; Catley, Phillips, Novick, 2013; Halverson & Friedrichsen, 2013; Meir et al., 2007; Morabito et al., 2010; Novick & Catley, 2013, 2014; Kapitel 3). Die Bezeichnungen für diese *wissenschaftlich inadäquaten* Tree-Reading-Performanzen reichen von *misconceptions* (z. B. Gregory, 2008) über *pitfalls* (Schramm & Schmiemann, 2019) zu *alternative strategies* (Blacquiere, Fawaz & Hoese, 2020). In deutschsprachigen Publikationen findet man vor allem im Kontext Evolution Begriffe wie Fehlvorstellungen (z. B. Beniermann, 2019) oder Schülervorstellungen (z. B. Brennecke, 2014). Laut Gropengießer und Marohn (2018) sollte stattdessen der neutralere Begriff „Alternative Vorstellungen“ genutzt werden.¹⁰ Da in der vorliegenden Studie der Fokus nicht auf Vorstellungen von Studierenden liegt und die Methodik eine Diagnose dieser nicht gewährleistet, wird als Abgrenzung zum *wissenschaftlich adäquaten Ablesen*, also dem inhaltlichen Ablesen und Interpretieren gemäß wissenschaftlicher Konventionen, die Bezeichnung *wissenschaftlich inadäquates Ablesen* entgegengestellt.

In der Tree-Reading- bzw. Tree-Thinking-Literatur findet meist keine Abgrenzung zwischen einer grundlegenden Vorstellung und *wissenschaftlich inadäquatem Ablesen* von DpV aufgrund von mangelndem Fachwissen statt. In der folgenden Auflistung wird daher nicht spezifiziert, wann es sich um eine tief verankerte alternative Vorstellung handelt und wann *wissenschaftlich inadäquates Ablesen* ohne zugrundeliegende Vorstellung gezeigt wird.

¹⁰ Graf und Hamdorf (2011) argumentieren, dass der Begriff Fehlvorstellungen genutzt werden kann, wenn „aus der Sicht der Fachwissenschaft argumentiert wird“ (Graf & Hamdorf, 2011, S. 26).

Stattdessen wird dargestellt, inwiefern Lernende laut Literatur häufig von den wissenschaftlichen Konventionen des Tree-Reading abweichen, wenn sie Tree-Reading-Aufgaben lösen sollen:

Merkmale

- **Veränderungen treten nur an Knotenpunkten auf** (Crisp & Cook, 2005; Gregory, 2008): Lernende formulieren, dass zwischen den Aufspaltungsereignissen keine weiteren Veränderungen stattgefunden haben. Veränderungen geschehen jedoch stetig, werden aber nur angegeben, wenn sie für den Fokus des DpV relevant sind bzw. ein Aufspaltungsereignis begründen.
- **Durchgehender Ast bedeutet keine Veränderungen** (Gregory, 2008; Novick & Catley, 2007): Weist ein Ast keine oder nur wenige Knotenpunkte auf, formulieren Lernende, dass in dieser Entwicklungslinie keinerlei Aufspaltungsereignisse stattgefunden haben bzw. keinerlei Veränderung bis in die Gegenwart stattgefunden hat. Je nach Fokus eines DpV werden jedoch nur Aufspaltungsereignisse bzw. Merkmale angegeben, die für die jeweilige Hypothese relevant sind.
- **Höhere und niedrigere Spezies bzw. lineare Entwicklung** (Crisp & Cook, 2005; Gregory, 2008; Krell & Cranston, 2004): Häufig werden Organismen von Lernenden als höher bzw. niedriger entwickelt bezeichnet als andere Organismen. Dies geschieht besonders häufig, wenn der Mensch als Taxon Teil eines zu interpretierenden DpV ist. Evolution ist jedoch nicht zielgerichtet oder linear und es gibt keine wissenschaftliche Begründung, Eigenschaften oder Merkmale von Organismen als höher oder niedriger entwickelt einzustufen.
- **Outgroup ist primitiv, weniger entwickelt als die anderen Taxa** (Omland et al., 2008): Vereint alle vorhergehenden *wissenschaftlich inadäquaten* Leseweisen der Kategorie „Merkmal“. Hier betrachten Lernende die *Outgroup* als weniger entwickelt als andere Taxa, da ihre Entwicklungslinie keine weiteren Aufspaltungsereignisse aufweist.

Verwandtschaft und MRCA

- **Outgroup ist MRCA aller anderen Taxa** (Gregory, 2008). Lernende bezeichnen die *Outgroup* als Vorfahre der anderen Taxa. Sie ist jedoch ebenfalls ein rezentes Taxon, das möglichst wenig mit den anderen Taxa verwandt ist, um den Baum zu „wurzeln“. Sie hat einen *MRCA* mit allen anderen Taxa und ist nicht selbst der *MRCA*.

- **Ermittlung bzw. Vergleich der relativen Verwandtschaft durch Knotenzählen** (Blacquiere & Hoese, 2016; Gregory, 2008; Halverson, 2011; Halverson & Friedrichsen, 2013; Meir et al., 2007): Um zu bestimmen, wie nah verschiedene Taxa miteinander verwandt sind, zählen Lernende die Knotenpunkte, die die Taxa gemeinsam haben und vergleichen diese mit der Anzahl gemeinsamer Knotenpunkte anderer Taxa. Der Grad der Verwandtschaft wird jedoch ermittelt, in dem der *MRCA* der betreffenden Taxa bestimmt wird. Je jünger dieser ist, desto näher verwandt sind die Taxa.
- **Lage der Taxa zueinander** (Baum et al., 2005; Blacquiere & Hoese, 2016; Gregory, 2008; Halverson, 2011; Halverson & Friedrichsen, 2013; Meir et al., 2007; Nee, 2005; O'Hara, 1988): Lernende definieren Verwandtschaftsverhältnisse verschiedener Taxa anhand der oberflächlichen Lage der Taxa im DpV anstatt anhand von *MRCA*.
- **Physische Ähnlichkeit als Kriterium für evolutive Verwandtschaft** (Baum et al., 2005; Gregory, 2008): Lernende ermitteln Verwandtschaftsverhältnisse verschiedener Taxa anhand physischer Ähnlichkeiten und nicht aufgrund der *MRCA* bzw. Kladen.
- **Rezente Taxa stammen voneinander ab** (Gregory, 2008): Lernende bezeichnen ein rezentes Taxon als Vorfahre eines anderen rezenten Taxons. Taxa stammen jedoch immer von (hypothetischen) *MRCA* ab und nicht voneinander.
- **Artenstehung durch Anagenese** (Gregory, 2008): Lernende postulieren, dass eine Spezies in eine andere übergeht und dadurch verschwindet. Neue Arten entstehen jedoch durch Aufspaltungsereignisse.
- **Einteilung der Taxa nach Vorwissen über Lebensräumen (Morabito et al., 2014) bzw. Ökologie** (Blacquiere & Hoese, 2016; Halverson, 2011; Halverson & Friedrichsen, 2013; Schramm et al., 2019): Lernende gruppieren Taxa anhand ihrer Lebensräume und nicht aufgrund der dargestellten Verwandtschaftsverhältnisse.
- **Einteilung in allgemeingebäuchliche Gruppen** (Atran, 1998): Lernende gruppieren Taxa anhand geläufiger Kategorien (Landsäugetiere, Vögel, "Reptilien", Amphibien, Fische, Insekten) und nicht aufgrund der dargestellten Verwandtschaftsverhältnisse.

Anthropozentrismus

- **Hauptast (zum Menschen) mit Seitenästen** (Gregory, 2008; Halverson, 2011; Halverson & Friedrichsen, 2013; Schramm & Schmiemann, 2019): Vor allem bei leiterförmigen DpV bezeichnen Lernende einen oberflächlich durchgehender Ast als „Hauptast“, von dem „Seiten-“ oder „Nebenäste“ abgehen. Das Taxon am Ende des „Hauptasts“ bezeichnen sie als „Hauptart“, alle anderen Taxa als sich von dieser „Hauptart“ abspaltend. Einerseits verläuft Evolution jedoch nicht zielgerichtet und andererseits können DpV an den Knotenpunkten beliebig rotiert werden, sodass der vermeintliche „Hauptast“ durch Rotation aufgelöst bzw. balanciert werden kann.

Zeitverlauf

- **Ältere und jüngere Abstammungslinien, gleichgesetzt mit Zeitpunkt des ersten Auftauchens eines Taxons** (Gregory, 2008): Lernende stufen Taxa, die aus einem jünger datierten *MRCA* hervorgegangen sind, als jünger ein als Taxa von früher datierten *MRCA*. Erdgeschichtlich haben sich jedoch alle rezenten Taxa gleich lang entwickelt, auch wenn das eine Klade oder ein Taxon begründende Aufspaltungsereignis erst später stattgefunden hat.
- **Zeitachse verläuft entlang der Taxa** (Gregory, 2008): Lernende verorten die Zeitachse entlang der Taxa, meist von links nach rechts. Die relative Zeitachse verläuft jedoch von der Wurzel als initialem Aufspaltungsereignis hin zu den Taxa.

Repräsentationsform

- **Vergleich von DpV anhand ihrer oberflächlichen Topologie** (Halverson, 2011; Halverson & Friedrichsen, 2013): Lernende vergleichen DpV anhand ihrer oberflächlichen Erscheinung. Ein inhaltlicher Vergleich der evolutiven Verwandtschaftsbeziehungen findet nicht statt.
- **Vergleich der Repräsentationsform** (Halverson, 2011): Lernende vergleichen DpV aufgrund ihrer Oberflächenmerkmale. Ein inhaltlicher Vergleich der evolutiven Verwandtschaftsbeziehungen findet nicht statt.

2.2. Tree-Reading-Fähigkeiten und ihre Struktur

Die vorliegende Arbeit bezieht sich auf das Interpretieren und Vergleichen mehrerer DpV und somit auf Tree-Reading gemäß der Definition von Halverson (2011; Kapitel 2.1.2). Je nach Studie wird dies ebenfalls als Tree-Reading, aber auch oft allgemeiner als Tree-Thinking bezeichnet. Im Folgenden wird daher immer dann der Begriff Tree-Reading genutzt, wenn es eindeutig um das Interpretieren und Vergleichen von DpV geht.

In der biologiedidaktischen Literatur werden verschiedene Tree-Reading-Fähigkeiten genannt (Baum & Smith, 2013; Blacquiere & Hoese, 2016; Halverson, 2011; Halverson & Friedrichsen, 2013; Kummer et al., 2019; Meir et al., 2007; Novick & Catley, 2016; Novick, Stull, & Catley, 2016; Schramm et al. 2019), die sowohl fachwissenschaftlich als auch biologiedidaktisch relevant sind.¹¹ Vergleicht man diese Veröffentlichungen, können die in Tabelle 1 aufgezeigten Tree-Reading-Fähigkeiten identifiziert werden. Das Verständnis der Prinzipien des *MRCA* (Blacquiere & Hoese, 2016; Novick & Catley, 2013) und der verschachtelten Kladen (*nested clades*; Novick & Catley, 2013) werden als essentiellste Tree-Reading-Fähigkeiten angesehen. Hierbei ist der Vergleich mehrerer DpV anspruchsvoller als die Interpretation eines einzelnen DpV (Halverson, 2011; Halverson & Friedrichsen, 2013; Novick & Catley, 2016; Schramm et al., 2019; Schramm et al., 2021a) und erfordert einen hohen Grad an Erfahrung mit DpV (Halverson & Friedrichsen, 2013).

¹¹ Die angegebenen Publikationen zu Tree-Reading enthalten meist auch Tree-Building-Fähigkeiten, die jedoch nicht primär relevant für die vorliegende Arbeit sind und entsprechend nicht genannt werden.

Tabelle 1: Tree-Reading-Fähigkeiten laut verschiedener Autor:innen unterteilt in die Interpretation einzelner DpV, dem Vergleich verschiedener DpV und übergreifenden Tree-Reading-/ Tree-Thinking-Fähigkeiten.

| Identifikation und Interpretation von... | Quelle |
|---|---|
| DpV-Bestandteilen | |
| Ästen/ Linien | Halverson, 2011; Schramm et al., 2019; Kummer et al., 2019 |
| Knotenpunkten als <i>MRCA</i> | Blacquiere & Hoese, 2016; Halverson, 2011; Halverson & Friedrichsen, 2013; Kummer et al., 2019; Morabito et al., 2010; Schramm et al., 2019 |
| Zeitverlauf | Halverson, 2011; Schramm et al., 2019 |
| Leserichtung | Novick, Stull, & Catley, 2012 |
| Kladen | Baum & Smith, 2013; Halverson, 2011; Halverson & Friedrichsen, 2013; Kummer et al., 2019; Novick & Catley, 2016; Schramm et al., 2019 |
| Verschachtelten Kladen (<i>nested clades</i>) | Novick & Catley, 2013; Novick & Catley, 2016; Baum & Smith 2013 |
| Relativer Verwandtschaft | Baum & Smith, 2013; Blacquiere & Hoese, 2016; Kummer et al., 2019; Novick & Catley, 2016; Schramm et al., 2019 |
| Rotation | Baum & Smith, 2013; Halverson, 2011; Kummer et al., 2019; Novick & Catley, 2016; Schramm et al., 2019 |
| Merkmale | |
| eines Taxons | Halverson & Friedrichsen, 2013; Kummer et al., 2019; Meir et al., 2007; Novick & Catley, 2016; Schramm et al., 2019 |
| mehrerer Taxa (mit einem bestimmten Merkmal) | Halverson & Friedrichsen, 2013; Kummer et al., 2019; Novick & Catley, 2016; Schramm et al., 2019; |
| des <i>MRCA</i> (anzestral) | Kummer et al., 2019; Meir et al., 2007 |
| Evolutiven Mechanismen | |
| Sparsamkeitsprinzip (Parsimonie) | Kummer et al., 2019 |
| Konvergenz | Kummer et al., 2019; Morabito et al., 2010; Novick & Catley, 2016 |
| Vergleich von... | |
| Mehreren DpV | Halverson, 2011; Schramm et al., 2019 |
| Kladen in verschiedenen DpV | Halverson, 2011; Halverson & Friedrichsen, 2013 |
| Mehreren DpV durch Rotation | Kummer et al., 2019; Schramm et al., 2019 |
| Unabhängig von Repräsentationsform | Halverson & Friedrichsen, 2013; Kummer et al., 2019; Schramm et al., 2019 |

Übergreifende Tree-Reading-/ Tree-Thinking-Fähigkeiten

| | |
|--|--|
| passendes DpV aus mehreren auswählen | Halverson & Friedrichsen, 2013; Schramm et al., 2019 |
| erörtern, darlegen vorhersagen | Baum & Smith, 2013; Halverson & Friedrichsen, 2013; Schramm et al., 2019 |
| schlussfolgern | Baum & Smith, 2013; Halverson, 2011; Novick & Catley, 2016 |
| neue mentale Modelle von DpV modellieren | Schramm et al., 2019 |

Tabelle 1 wurde mit dem Ziel erstellt, alle publizierten Tree-Reading-Fähigkeiten als Grundlage für die Analysewerkzeuge der vorliegenden Arbeit zusammenzufassen, die sich auf die Interpretation oder den Vergleich bestimmter DpV-Bestandteile beziehen.¹² Betrachtet man jedoch die einzelnen Publikationen, beschreiben die Autor:innen je nach Fokus unterschiedlich viele Tree-Reading-Fähigkeiten. So behandeln Blacquiere und Hoesé (2016) ausschließlich evolutive Verwandtschaft als fundamentalen Faktor für das Verständnis von DpV und Evolution. Im Gegensatz dazu beschreiben Novick und Catley zunächst 5 nicht-hierarchische Tree-Reading-Fähigkeiten (2013) und erweitern diese später auf 11 (2016):¹³

- Merkmale identifizieren
- Taxa mit gleichem Merkmal identifizieren
- Kladen identifizieren
- Verschachtelte Kladen identifizieren
- Evolutive Verwandtschaft: *Resolved Structure*
- Evolutive Verwandtschaft: Polytomie
- Schlussfolgerungen (*Inference*) zu evolutiver Verwandtschaft
- zeitliche Reihenfolge von Merkmalen
- Konvergenz
- *Subsets of the Tree of Life*
- Rotation

Während manche Autor:innen Tree-Reading-Fähigkeiten auf die Interpretation und den Vergleich konkreter DpV-Bestandteile bzw. Mechanismen beziehen (z. B. Blacquiere &

¹² Die Publikationen von Schramm et al. (2019) und Kummer et al. (2019) waren zum Zeitpunkt der Erstellung der Analysewerkzeuge noch nicht erschienen und wurden der Vollständigkeit halber Tabelle 1 hinzugefügt. Außerdem werden sie in der Diskussion der vorliegenden Arbeit herangezogen.

¹³ Die genaue Definition der einzelnen Tree-Reading-Fähigkeiten ist der Publikation von Novick & Catley (2016) zu entnehmen.

Hoese, 2016; Novick & Catley, 2013, 2016), postulieren Halverson und Friedrichsen (2013) 7 aufeinander aufbauende, an Schwierigkeit zunehmende Level,¹⁴ anhand derer der Umgang mit DpV-Bestandteilen, Merkmalen und Mechanismen übergreifend klassifiziert werden soll. Diese 7 Level sind angelehnt an die 5 Fähigkeiten zur Erlangung von Repräsentationskompetenz von Kozma und Russell (2005). Der hierarchische Aufbau der 7 Level wurde bisher noch nicht empirisch überprüft.

1. *No Use of Representation*: Studierende nutzen Vorwissen statt DpV und vergleichen DpV nicht miteinander.
2. *Superficial Use of Representation*: Studierende nutzen DpV-Bestandteile ausschließlich oberflächlich zum Interpretieren und Vergleichen von DpV.
3. *Simplified Use of Representation*: Studierende nutzen wissenschaftlich nicht adäquat den „Hauptast“ und die Astlänge zum Interpretieren und Vergleichen von DpV.
4. *Symbolic Use of Representation*: Studierende zeigen grundsätzliches Verständnis der inhaltlichen Bedeutung der DpV-Bestandteile, nutzen aber vorwiegend Knotenpunkte wissenschaftlich nicht adäquat (Zählen der Knotenpunkte zwischen Taxa, Position und Anzahl der Knotenpunkte) zum Interpretieren und Vergleichen von DpV.
5. *Conceptual Use of Representation*: Studierende können DpV an Knotenpunkten mental rotieren, sehen DpV aber nicht im evolutiven Kontext. Sie vergleichen DpV anhand der Topologie und abhängig von der Repräsentationsform.
6. *Scientific Use of Representation*: Studierende interpretieren und vergleichen DpV wissenschaftlich adäquat anhand von Kladen, MRCA und Merkmalen unabhängig von der Repräsentationsform.
7. *Expert Use of Representation*: Umfassende Interpretation, Vergleich und Anwendung von zahlreichen DpV zur Lösung von Fragestellungen, Erklärung von Phänomenen und zur Voraussage. Verschiedene DpV werden zur Überprüfung von Hypothesen herangezogen. Dieses Level kann nicht von Studierenden erreicht werden, sondern bleibt Expert:innen vorbehalten. In diesem Level wird nicht mehr nach Tree-Reading und Tree-Building unterschieden.

¹⁴ Neben den 7 Tree-Reading-Leveln werden auch 7 entsprechende Tree-Building-Niveaus beschrieben (siehe Halverson und Friedrichsen, 2013).

Die Vielfalt der bisherigen Ansätze, Tree-Reading-Fähigkeiten zu beschreiben und einzuteilen, bewegten Schramm und Kolleg:innen dazu, das *Synthetic Tree-Reading Model (STREAM)* zu entwickeln (2019) und zu evaluieren (2021a). Auf Grundlage der oben erläuterten Studien von Novick und Catley (2016) und Halverson und Friedrichsen (2013) sowie Tree-Reading-Fähigkeiten anderer Autor:innen modellierten sie zunächst 6 hierarchische Level für Tree-Reading-Fähigkeiten auf theoretischer Grundlage (Schramm et al., 2019):

0. *Naïve Handling*: Studierende können DpV nicht *wissenschaftlich adäquat* interpretieren und nutzen uninformativ DpV-Bestandteile und alternative Vorstellungen.
1. *Identifying Structures*: Studierende können DpV-Bestandteile (z. B. Knotenpunkte) interpretieren und können Fragen zur DpV-Struktur beantworten.
2. *Handling Apomorphies*: Studierende können Fragen zu Merkmalen beantworten und Taxa entsprechend ordnen.
3. *Identifying Relationships*: Studierende können Klagen und Verwandtschaftsverhältnisse bestimmen.
4. *Comparing Trees*: Studierende können Verwandtschaftsverhältnisse zwischen verschiedenen DpV vergleichen.
5. *Arguing and Inferring*: Studierende können DpV nutzen, um Schlussfolgerungen und Vorhersagen zu treffen, die über das Gezeigte hinausgehen.

Um die angenommene Hierarchie zu überprüfen, erstellten Schramm und Kolleg:innen (2021a) zunächst einen Tree-Reading-Fragebogen aus bereits publizierten, angepassten Tree-Reading-Items, die den einzelnen Leveln entsprechen. Die Antworten von 455 Studierenden sowie die Daten aus einer zusätzlichen Think-Aloud-Studie mit 8 Studierenden flossen in die Analyse ein. Mithilfe der Item-Response-Theorie wurden die Items hinsichtlich ihrer Schwierigkeit und Levelzugehörigkeit analysiert. Hier konnte eine hierarchische Ordnung der Level nicht eindeutig bestätigt werden, da die Items innerhalb der meisten Level eine stark variierende Schwierigkeit aufwiesen und es hier zwischen den Leveln häufig zu Überlappungen kam. Es wurden zwar ansteigende Schwierigkeiten von Level 2 zu 3 und 4 festgestellt, diese seien laut der Autor:innen jedoch wahrscheinlich durch die zunehmende inhaltliche Komplexität und Leseschwierigkeit der Items begründet. Zur weiteren Differenzierung der Level wurden Dimensionsanalysen durchgeführt. Hierbei erwies sich ein fünfdimensionales

Modell als am wahrscheinlichsten, bei dem die Dimensionen nicht hierarchisch, sondern als distinkte Tree-Reading-Fähigkeiten zu betrachten sind. Diese Dimensionen ergeben sich aus den vormaligen Leveln 1 bis 4, wobei das Level 3 (*Identify Relationships*) zusätzlich in 2 separate Dimensionen unterteilt ist: *Identify Relationships* und *Determining monophyletic Groups* (Abbildung 3). Schramm und Kolleg:innen schlussfolgern abschließend, dass das Level 0 - *Naïve Handling* als *Base Level* bestehen bleibt, weil bei den hier einzuordnenden Studierenden keine bzw. kaum Tree-Reading-Fähigkeiten vorliegen. Im *Learner's Level* liegen bei Studierenden unterschiedliche Ausprägungen der 5 Tree-Reading-Fähigkeiten bzw. Dimension vor. Es wird also nicht wie bei Halverson und Friedrichsen postuliert eine Tree-Reading-Fähigkeit nach der anderen erlernt, sondern eher alle Tree-Reading-Fähigkeiten parallel in unterschiedlichem Maße. Werden alle 5 Tree-Reading-Fähigkeiten bzw. Dimensionen beherrscht und angewendet, können Personen dem *Expert's Level - Arguing and Inferring* zugeordnet werden (Abbildung 3; vgl. Schramm et al., 2021a).

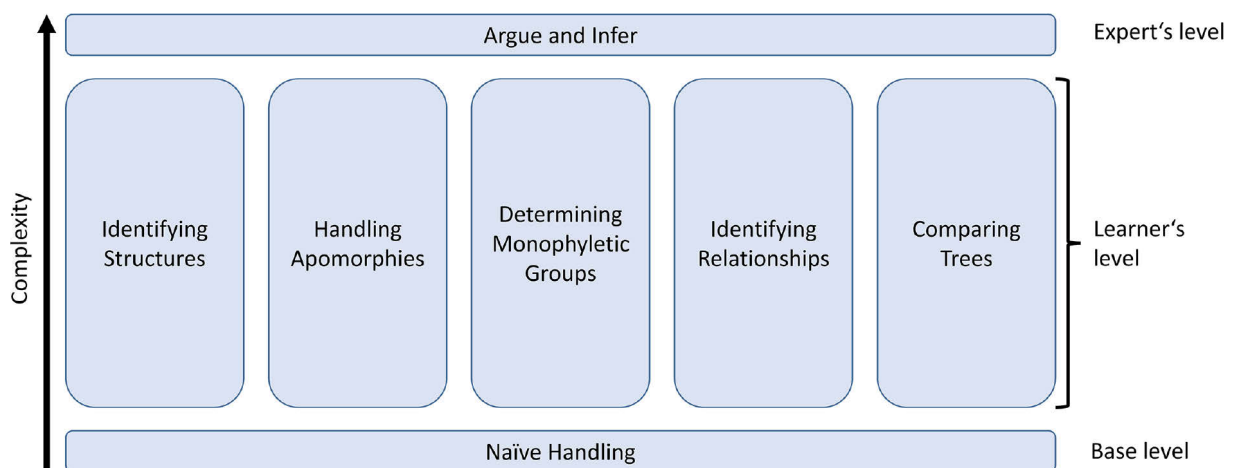


Abbildung 3: Revised Synthetic Tree-Reading Model (STREAM) von Schramm et al. (2021a, S. 9). Tree-Reading-Fähigkeiten können als Dimensionen unabhängig voneinander auf unterschiedlichen Leveln beherrscht bzw. erlernt werden.

Schramm und Kolleg:innen (2021a) kontrastieren ihr *Revised STREAM* mit dem *Evolutionary Tree Concept Inventory (ETCI)* (Kummer, 2017; Kummer et al., 2019; Tabelle 2) das *Learning Outcomes*, also Tree-Reading-Fähigkeiten, die Studierende beherrschen sollten, erst in 3 (Gemeinsame Abstammung, Merkmale, allgemeines Tree-Thinking; Kummer, 2017) und später anhand einer Faktorenanalyse empirisch in 5 Faktoren einteilt (Kummer et al., 2019):

Tabelle 2: Gegenüberstellung von ETCI und Revised STREAM in Anlehnung an Schramm et al. (2021a, S. 10).

| Learning Outcomes des ETCI (Kummer et al. 2017, 2019) | Dimension aus Revised STREAM (Schramm et al., 2021) |
|--|--|
| 1. Studierende können evolutive Verwandtschaft zwischen Taxa vergleichen. | <i>Identifying Relationships</i> |
| 2. Studierende können unterscheiden, ob DpV unterschiedliche Rotationen oder unterschiedliche evolutive Verwandtschaftsverhältnisse darstellen. | <i>Comparing Trees</i> |
| 3. Studierende können anhand theoretischer Aspekte von DpV die Evolution von Gruppen und Merkmalen anhand gemeinsamer Abstammung und des Sparsamkeitsprinzips bestimmen. | <i>Argue and Infer</i> |
| a. Homologie und Analogie identifizieren. | <i>Argue and Infer</i> |
| b. Merkmale anhand des Sparsamkeitsprinzips analysieren. | <i>Argue and Infer</i> |
| c. Monophyla, Paraphyla und Polyphyla unterscheiden. | <i>Determining monophyletic Groups</i> |
| d. DpV-Bestandteile identifizieren. | <i>Identifying Structures</i> |
| 4. Studierende verstehen, wie Merkmale der MRCA weitergegeben werden, wenn diese im DpV angegeben sind. | <i>Handling Apomorphies</i> |
| a. Synapomorphien identifizieren. | <i>Handling Apomorphies</i> |
| b. Abgeleitete bzw. anzestrale Merkmale identifizieren. | <i>Handling Apomorphies</i> |
| c. Merkmale eines Taxons anhand eines DpV bestimmen. | <i>Handling Apomorphies</i> |
| 5. Studierende verstehen Evolution als kontinuierlichen und nicht-teleologischen Prozess. | <i>Argue and Infer</i> |
| a. Komplexität von DpV ist kein Maß für primitivere oder weiterentwickelte Organismen. | <i>Argue and Infer</i> |
| b. Alle extanten Populationen entwickeln sich stets weiter und haben sich permanent entwickelt. | <i>Argue and Infer</i> |

Laut Schramm und Kolleg:innen umfassen das *Revised STREAM* und das *ETCI* überwiegend die gleichen Aspekte, legen allerdings andere Schwerpunkte. So stellt das *ETCI* ein *Concept Inventory* für Tree-Thinking im weiteren Sinne dar, während das *Revised STREAM* konkret Tree-Reading modelliert (vgl. Schramm et al., 2021a).

Die genannten Strukturierungsansätze haben unterschiedliche Schwerpunkte und Ziele. So charakterisieren Halverson und Friedrichsen (2013) Level, die von Studierenden beim Erlernen von Tree-Reading durchlaufen werden, die also z. B. zum Erfassen des Lernfortschritts genutzt werden könnten. Novick und Catley (2013, 2016) nicht-hierarchische Tree-Reading-Fähigkeiten auf, die nötig sind, um Tree-Reading zu beherrschen (vgl. Schramm et al., 2019). Zum Zeitpunkt der Entwicklung der Items und Analyseinstrumente der vorliegenden Arbeit lagen weder das *Revised STREAM* (Schramm et al., 2021a) noch das *ETCI* (Kummer et al., 2019) vor. Eine empirische Überprüfung der 7 Level von Halverson und Friedrichsen (2013) erfolgte bisher noch nicht. Um Tree-Reading-Prozesse in der vorliegenden Arbeit dennoch klassifizieren zu können, wurden als Grundlage die 3 Stufen des Kompetenzmodells zum Umgang mit Diagrammen herangezogen, das sich sowohl auf das Interpretieren als auch das Erstellen von Diagrammen bezieht (Nitz, Meister, Schwanewedel, & Upmeier zu Belzen, 2018):

1. *Abbildend*: Interpretation von Diagrammen ausschließlich aufgrund von Oberflächenmerkmalen ohne Berücksichtigung der Diagrammkonventionen.
2. *Syntaktisch*: Interpretation von Diagrammen gemäß der Diagrammkonventionen ohne Verknüpfung mit zugrundeliegendem Phänomen.
3. *Semantisch-reflexiv*: Erklären und Vorhersagen über den Diagramminhalt hinaus.

Bezogen auf DpV ergeben sich daraus übergeordnete Tree-Reading-Fähigkeiten, anhand derer DpV-Bestandteile, Merkmale und evolutive Mechanismen auf unterschiedlichen Stufen interpretiert und verglichen werden können:

0. *Oberflächliches Ablesen* von Informationen aus einem oder mehreren DpV ohne Berücksichtigung der wissenschaftlichen Konventionen bzw. des Inhalts.
1. *Ablesen* von Informationen aus einem oder mehreren DpV auf syntaktischer Ebene
 - a. *wissenschaftlich adäquat*: anhand wissenschaftlicher Konventionen
 - b. *wissenschaftlich inadäquat*: ohne Berücksichtigung wissenschaftlicher Konventionen bzw. anhand alternativer Leseweisen und/ oder Vorstellungen
2. *Ableiten*, erklären und Vorhersagen¹⁵ von Informationen, die nicht direkt im DpV dargestellt sind.

¹⁵ Da DpV immer evolutive Ereignisse aus der Vergangenheit modellieren, erfolgt hier eine retrospektive Erklärung dieser Ereignisse und somit keine Vorhersage im eigentlichen Sinne. Siehe Abschnitt 2.3.2 zur Einordnung in das Kompetenzmodell der Modellkompetenz bzw. Modellierkompetenz.

Ableiten (TR_2) wird hierbei als anspruchsvollste Tree-Reading-Fähigkeit eingestuft, da über den Inhalt des DpV hinaus Schlüsse gezogen werden. *Wissenschaftlich adäquates Ablesen* (TR_1a) erfolgt mit direktem Bezug zum Inhalt des DpV und ist somit potenziell weniger anspruchsvoll. Dem untergeordnet ist *oberflächliches Ablesen* (TR_0), da hier nicht der Bezug zum Inhalt erfolgt. Eine Sonderstellung nimmt *wissenschaftlich inadäquates Ablesen* (TR_1b) ein. Hier erfolgt ein Bezug zum Inhalt, weshalb es wie *wissenschaftlich adäquates Ablesen* (TR_1a) auf der syntaktischen Ebene angesiedelt ist. Der Bezug beruht jedoch auf fehlendem oder fehlerhaftem Fachwissen bzw. alternativen Vorstellungen. Dadurch ergibt sich auf fachwissenschaftlicher Ebene ein geringerer Anspruch für *wissenschaftlich inadäquates Ablesen* (TR_1b).

Hierbei ist zu betonen, dass die Bewertung des Anspruches der Tree-Reading-Fähigkeiten in der vorliegenden Arbeit auf einer theoretischen Annahme beruht, da erst durch Schramm und Kolleg:innen (2021a) eine empirisch geprüfte Struktur mit Dimensionen und Leveln vorgestellt wurde.

2.3. Diagramme phylogenetischer Verwandtschaft im Kontext der Modelletheorie

2.3.1. Diagramme phylogenetischer Verwandtschaft als Modelle *von* und *für* Evolution

Ausgehend vom Begriff „Diagramme phylogenetischer Verwandtschaft“ lassen sich DpV zunächst als strukturierende Diagramme (Kattmann, 2013) betrachten. Halverson und Friedrichsen (2013) betten sie in den Kontext der Repräsentationskompetenz nach Kozma und Russell (2005) ein. Ein wichtiger Aspekt des Tree-Reading ist hier die Erkenntnis, dass DpV trotz unterschiedlicher Repräsentationsformen, bspw. durch Rotation an Knotenpunkten oder durch baumförmige oder leiterförmige Darstellung, die gleichen evolutiven Verwandtschaftsverhältnisse zeigen können (z. B. Halverson & Friedrichsen, 2013, Kummer et al., 2019; Novick & Catley, 2016, Schramm et al., 2019) und somit auf das gleiche zugrundeliegende Modell zurückgehen (Passmore, Gouvea, & Giere, 2014). Die Verwendung verschiedener Repräsentationsformen des gleichen Modells kann Reasoning-Prozesse fördern, da wesentliche Charakteristika des Modells hervorgehoben werden können. Auch Kommunikation anhand verschiedener Repräsentationsformen kann Erkenntnisse über das Modell an sich generieren (Nersessian, 2002).

Der Modellbegriff ist nach wie vor nicht einheitlich definiert (vgl. Upmeier zu Belzen & Krüger, 2010; Krüger, D., Kauertz, A., & Upmeier zu Belzen, 2018; Upmeier zu Belzen, van Driel & Krüger, 2018).¹⁶ Die vorliegende Arbeit nutzt als Grundlage die Definition nach Mahr (2008), laut der ein Modell erst durch eine Person zu einem Modell erklärt wird. Darauf aufbauend formulieren Krüger und Kolleg:innen:

„Modelle lassen sich demnach im Urteil von Personen als Repräsentationen von Originalen bzw. Phänomenen oder Systemen einer Erfahrungswelt auffassen, die diese Phänomene für einen bestimmten Zweck und zeitlich befristet beschreiben oder erlauben, über die Erfahrungswelt Hypothesen aufzustellen und diese zu untersuchen.“
(Krüger, Kauertz, & Upmeier zu Belzen, 2018, S. 142)

Wird ein Modellobjekt als Repräsentation eines Originals erstellt, liegt laut Mahr ein Modell *von* etwas vor (Herstellungsperspektive). Dieses Modellobjekt selbst kann nun als Modell *für*

¹⁶ Für die ausführliche Erläuterung der verschiedenen Modellbegriffe siehe ebd.

etwas genutzt werden (Anwendungsperspektive; Mahr, 2008, 2009; Passmore et al., 2014), also bspw. zur Aufstellung von Hypothesen über das Original. Im ersten Fall wird das Modell also als Medium eingesetzt, im zweiten als Methode (Gilbert, 1991; Mahr, 2009).

Bezogen auf DpV ist das zu modellierende Original die Evolution an sich bzw. evolutive Verwandtschaftsverhältnisse. Angesichts der enormen Zeitspannen und Komplexität der Entwicklung des Lebens auf der Erde formulieren Knoop und Müller (2009) ein scheinbar drastisches Urteil über Modelle, das sie dann relativieren: *„Alle Modelle sind falsch – es sind eben nur Modelle, und die Realität können, sollen und werden sie niemals exakt abbilden. Dennoch sind manche Modelle nützlicher als andere, und man sollte meinen, dass es grundsätzlich wünschenswert ist, wenn die Substitutionsmodelle der biologischen Realität so nah wie möglich kommen, also möglichst komplex sind.“* (Knoop & Müller, 2009, S. 278). Sie beziehen sich dabei auf verschiedene Möglichkeiten, evolutive Verwandtschaft retrospektiv zu modellieren. Es sind also nicht die DpV an sich gemeint, die grundsätzlich falsch seien, sondern die Simulationsmodelle, anhand derer z. B. DpV modelliert werden und die je nach Fragestellung und Datenlage mehr oder weniger gut zur Hypothesengenerierung geeignet sind. Gleichzeitig betonen sie die Schwierigkeit für Wissenschaftler:innen, die passende Modelliermethode zu wählen, um Vergangenes zu rekonstruieren: *„Woher weiß man eigentlich, welche der im Buch vorgestellten Rekonstruktionsmethoden man verwenden soll? Gar nicht (zumindest nicht ohne weiteres). Sonst würden wir hier nur „die“ beste vorstellen und empfehlen, den Rest zu ignorieren. So einfach ist das Ganze aber nun einmal nicht, [...]“* (Knoop & Müller, 2009, S. 277).¹⁷ Wissenschaftler:innen nutzen also verschiedene Methoden, um evolutive Verwandtschaft als Hypothesen in Form von DpV zu modellieren (Eldredge & Cracraft, 1980). Das DpV als Modellobjekt wird medial genutzt, um über verschiedene Hypothesen zu evolutiven Verwandtschaftsverhältnissen zu kommunizieren. In der Anwendungsperspektive setzen Wissenschaftler:innen DpV methodisch ein, um neue Erkenntnisse über das Original zu erlangen und neue Hypothesen zu generieren.

¹⁷ Für ausführliche Informationen zur Modellierung von evolutiven Verwandtschaftsverhältnissen siehe Baum und Smith (2013) und Knoop und Müller (2009).

2.3.2. Kompetenzmodelle der Modell- und Modellierkompetenz

Ausgehend vom Kompetenzbegriff von Weinert (2001) definieren Upmeier zu Belzen und Krüger (2010) die Gesamtheit der verschiedenen Aspekte des Umgangs mit Modellen als Modellkompetenz:

„Modellkompetenz umfasst die Fähigkeiten, mit Modellen zweckbezogen Erkenntnisse gewinnen zu können und über Modelle mit Bezug auf ihren Zweck urteilen zu können, die Fähigkeiten, über den Prozess der Erkenntnisgewinnung durch Modelle und Modellierungen in der Biologie zu reflektieren sowie die Bereitschaft, diese Fähigkeiten in problemhaltigen Situationen anzuwenden.“

(Upmeier zu Belzen und Krüger, 2010, S. 49)

Daraus resultierend entwickelten sie das Kompetenzmodell der Modellkompetenz, in dem die verschiedenen Aspekte des Umgangs mit Modellen in 5 Teilkompetenzen mit 3 Niveaus aufgegliedert sind (Tabelle 3; Krüger, Kauertz, & Upmeier zu Belzen, 2018; Upmeier zu Belzen & Krüger, 2010; Upmeier zu Belzen & Krüger, 2019).

Tabelle 3: Kompetenzmodell der Modellkompetenz mit 5 Teilkompetenzen und 3 Niveaus. Niveau I und II entsprechen dem medialen Modelleinsatz, Niveau III dem methodischen. In Anlehnung an Krüger, Kauertz und Upmeier zu Belzen (2018), Upmeier zu Belzen & Krüger (2010), Upmeier zu Belzen & Krüger (2019).

| Modelleinsatz | Medial | | Methodisch |
|-----------------------------------|---|--|--|
| Teilkompetenzen | Niveau I | Niveau II | Niveau III |
| Eigenschaften von Modellen | Modelle sind Kopien von etwas | Modelle sind idealisierte Repräsentationen von etwas | Modelle sind theoretische Rekonstruktionen von etwas |
| Alternative Modelle | Unterschiedliche Konstruktion führt zu unterschiedlichen Modellobjekten | Ausgangsobjekt ermöglicht Herstellung unterschiedlicher Modelle | Verschiedene Hypothesen führen zu unterschiedlichen Modellen |
| Zweck von Modellen | Modellobjekt zur Beschreibung von etwas einsetzen | Bekannte Zusammenhänge von Variablen im Ausgangsobjekt erklären | Zusammenhänge von Variablen für zukünftige neue Erkenntnisse voraussagen |
| Testen von Modellen | Modellobjekt überprüfen | Parallelisieren mit dem Ausgangsobjekt; Modell von etwas testen | Überprüfen von Hypothesen bei der Anwendung; Modell für etwas testen |
| Ändern von Modellen | Mängel am Modellobjekt beheben | Modell als Modell von etwas durch neue Erkenntnisse oder zusätzliche Perspektiven revidieren | Modell für etwas aufgrund falsifizierter Hypothesen revidieren |

Übergreifend können Niveau I und II dem medialen Modelleinsatz zugeordnet werden. Niveau I spiegelt die Fähigkeit wider, sich fast ausschließlich auf das Modellobjekt selbst zu beziehen, nur bei den Teilkompetenzen *Eigenschaften* und *Zweck von Modellen* findet eine Verknüpfung zum modellierten Phänomen statt, wenn das Modellobjekt als Kopie von etwas oder zur Beschreibung von etwas genutzt wird. Niveau II entspricht der Fähigkeit, Modellobjekte als Repräsentation des Originals zu begreifen und sie medial zur Erklärung von Bekanntem zu nutzen. Niveau III stellt die Anwendungsperspektive dar, also die Fähigkeit, ein Modell methodisch zu nutzen, um neue Erkenntnisse über das Original zu erlangen bzw. Voraussagen zu treffen (vgl. Krüger et al., 2018; Upmeier zu Belzen & Krüger, 2019). Modellkompetenz ist wie andere Kompetenzen erlern- und vermittelbar sowie kontextspezifisch (Hartig & Klieme, 2006).

Wie im vorherigen Kapitel 2.3.1 beschrieben, nutzen Wissenschaftler:innen DpV als Modelle sowohl medial als auch methodisch und bewegen sich so in den verschiedenen Teilkompetenzen, indem sie z. B. alternative DpV miteinander vergleichen, neue Erkenntnisse mithilfe der DpV generieren, sie testen und ggf. ändern. Hierbei muss jedoch bedacht werden, dass im Gegensatz zu anderen zu modellierenden Phänomenen evolutive Verwandtschaft nicht in der Zukunft liegt und somit keine zukünftigen Ereignisse vorausgesagt, sondern vergangene Ereignisse rekonstruiert werden.

Das der vorliegenden Arbeit zugrundeliegende Kompetenzmodell der Modellkompetenz wurde um ein Niveau zum abduktiven Schließen ergänzt und zum Kompetenzmodell der Modellierkompetenz erweitert (Krüger und Upmeier zu Belzen, 2021). So wird auf Niveau III nun zwischen abduktivem Schließen (Niveau IIIa) bei der Modellherstellung und dem Prüfen deduzierter Hypothesen bei der Modellanwendung (Niveau IIIb, ursprünglich Niveau III) unterschieden, die beide erkenntniserweiternd sind (Tabelle 4). Wurde der Begriff des Erklärens im ursprünglichen Kompetenzmodell in Niveau II lediglich medial genutzt, wird er hier in Niveau IIIa auch methodisch interpretiert, indem bei der Modellherstellung durch Abduktion Erklärungen für ein Phänomen gesucht werden und es zum Erkenntnisgewinn kommt. Im Gegensatz zur Anwendung des Modells zum hypothetisch-deduktiven Erkenntnisgewinn kann abduktives Schließen somit retrospektiv eingesetzt werden, z. B. zur Erklärung evolutiver Phänomene wie den unterschiedlichen Schnabelformen von Darwinfinken. Dies geschieht, indem bestehende Theorien genutzt werden oder neue

Theorien innovativ erzeugt werden und somit ein Modell im Sinne einer erkenntnisgenerierenden Erklärung erstellt wird (vgl. Krüger & Upmeyer zu Belzen, 2021, S. 4). Baum und Smith (2013) beschreiben, dass DpV dazu dienen, evolutive Veränderungen retrospektiv zu rekonstruieren, also abduktiv von Merkmalen rezenter Spezies auf deren Entstehungsgeschichte zu schließen. Dabei werden durch immer präzisere Berechnungsverfahren verschiedene DpV als Modelle miteinander verglichen, um die sparsamste Erklärung für die Entstehung eines Merkmals und somit die Verwandtschaftsbeziehungen relevanter Taxa zu ermitteln. So können wie in Kapitel 1 beschrieben bspw. Erkenntnisse über Krankheitserreger gewonnen werden, um Ausbrüche erklären und bekämpfen zu können.

Tabelle 4: Kompetenzmodell der Modellierkompetenz mit 5 Teilkompetenzen und 4 Niveaustufen (Krüger & Upmeyer zu Belzen, 2021, S. 7).

| Teilkompetenzen | Niveau I | Niveau II | Niveau IIIa | Niveau IIIb |
|-----------------------------------|--|--|---|---|
| | Exklusive Betrachtung des Modellobjekts | Sachbezogen verständnis-erzeugende Erklärung des Phänomens | Abduktiv schließende Erklärung des Phänomens | Hypothetisch-deduktive Untersuchung des Phänomens |
| Eigenschaften von Modellen | Entspricht möglichst vollständig dem Phänomen | Repräsentiert verständlich und abstrahiert das Phänomen | Repräsentiert beste Erklärung des Phänomens | Repräsentiert eine theoretische Rekonstruktion des Phänomens |
| Alternative Modelle | Existieren wegen unterschiedlicher Objektmerkmale | Existieren wegen der Komplexität des Phänomens unterschiedlicher Modelle | Existieren wegen unterschiedlicher theoretischer Erklärungen | Existieren wegen konkurrierender theoretischer Vorstellungen |
| Zweck von Modellen | Dient dem Beschreiben und Veranschaulichen des Phänomens | Dient dem verständlichen Erklären des Phänomens | Liefert die beste theoretische Erklärung des Phänomens | Erlaubt mit Hypothesen das Voraussagen empirischer Ergebnisse zum Phänomen |
| Testen von Modellen | Erfolgt durch Untersuchen der Robustheit oder Funktionalität | Erfolgt durch Überprüfen der Passung zwischen Phänomen und Modell | Erfolgt durch theoretisches Überprüfen der Erklärung auf Konsistenz | Erfolgt durch empirisches Überprüfen von Hypothesen in der Erfahrungswelt |
| Ändern von Modellen | Führt zum Beheben von Mängeln | Führt zum Revidieren bei fehlender Passung durch neue Erkenntnisse | Führt zum Revidieren der Erklärungen bei fehlender Konsistenz | Führt zum Integrieren von neuen Erkenntnissen aus falsifizierten Hypothesen |

2.4. Erfahrungsgrad

Expert:innen zeigen eine „*consistently superior performance on a specified set of representative tasks for a domain*“ (Ericsson & Lehmann, 1996, S. 277). Im Gegensatz dazu stehen Noviz:innen, die eine bestimmte Aufgabe mangels benötigtem Wissen bzw. Erfahrung nicht ausführen können. Abgesehen von der Performanz kann Expertise auch über den Grad der sowohl qualitativen als auch quantitativen Erfahrung als Kontinuum zwischen Noviz:innen und Expert:innen gesehen werden. Zusätzlich kann auch eine herausragende Leistung innerhalb eines Fachgebiets als Expertise benannt werden (vgl. van Gog, 2012). In zahlreichen Studien finden Vergleiche zwischen Noviz:innen und Expert:innen statt, um die Performanz der Noviz:innen anhand der Performanz der Expert:innen einzuordnen (z. B. Grosslight, Unger, & Jay, 1991; Jarodzka, Scheiter, Gerjets, & van Gog, 2010).

Expert:innen im Umgang mit DpV sind Personen, die sowohl Tree-Reading als auch Tree-Building beherrschen (Halverson, 2011). Halverson und Friedrichsen (2013) formulieren dies im Kontext der Repräsentationskompetenz im Umgang mit DpV im siebten Level „*Expert Use of Representation*“ als umfassende Interpretation, Vergleich und Anwendung von zahlreichen DpV zur Lösung von Fragestellungen, Erklärung von Phänomenen und zur Voraussage. Verschiedene DpV werden zur Überprüfung von Hypothesen herangezogen. Diese Definition hebt sich von den vorhergehenden 6 Leveln ab, da hier nicht mehr nach Tree-Reading und Tree-Building unterschieden wird, sondern Expert:innen Tree-Thinking als Gesamtheit beherrschen. Außerdem wird postuliert, dass Studierende dieses Level nicht erreichen können (Halverson & Friedrichsen, 2013; Kapitel 2.2).

Schramm und Kolleg:innen (2021a) formulieren ebenfalls ein *Expert's Level*, dieses kann jedoch auch von Studierenden erreicht werden, wenn sie alle 5 definierten Tree-Reading-Fähigkeiten (Abbildung 3) beherrschen: „*Experts use their knowledge of all skills and other concepts to interpret trees and infer information.*“ (Schramm et al., 2021a, S. 8).

Die vorliegende Arbeit nutzt als Kriterium zur Festlegung der Vergleichsgruppen den angenommenen Erfahrungsgrad mit DpV. Als Definition dient hier die Annahme, dass Wissenschaftler:innen aus dem Feld der Evolutionsbiologie einen deutlich höheren Erfahrungsgrad mit DpV vorweisen als Lehramtsstudierende der Biologie, die Tree-Reading im Rahmen ihres Studiums bis zu einem gewissen Grad erlernen.

Da die für Expert:innen angenommene höhere Performanz nicht als Kriterium zur Festlegung der Vergleichsgruppen diente, wird in der vorliegenden Arbeit nicht der Begriff des

Expert:innen-Noviz:innen-Vergleichs verwendet. Stattdessen werden zur passenden Beschreibung der Vergleichsgruppen die Bezeichnungen „Personen mit hohem Erfahrungsgrad“ (HoEr) und „Personen mit niedrigem Erfahrungsgrad“ (NiEr) herangezogen. Aufgrund des angenommenen Erfahrungsgrades besteht die Annahme, dass HoEr Tree-Reading-Fähigkeiten zeigen, die den oben beschriebenen Expert:innenleveln entsprechen, die von NiEr nicht bzw. nur selten erreicht werden.

2.5. Eyetracking - Visuelle Wahrnehmung von Repräsentationen

Anhand von DpV werden komplexe Informationen über evolutive Verwandtschaftsverhältnisse stark abstrahiert modelliert. Neben der Erhebung verbaler Daten (z. B. Halverson & Friedrichsen, 2013) können visuelle Aspekte des Tree-Reading wie z. B. die Leserichtung (Novick et al., 2012) anhand von Eyetracking untersucht werden. Die Untersuchung von Blickbewegungen beruht auf der Annahme, dass die Aufmerksamkeit einer Person auf dem Bereich liegt, der von ihr angeschaut wird (*Eye-Mind-Assumption*; Just & Carpenter, 1980). Dieser fixierte Bereich wird dabei unmittelbar kognitiv verarbeitet (*Immediacy Assumption*; Just & Carpenter, 1980). Zusammen mit verbalen Daten bilden visuelle Daten diese kognitiven Prozesse umfassend und sich gegenseitig ergänzend ab (Stieff, Hegarty, & Deslongchamps, 2011). Bearbeiten NiEr eine Aufgabenstellung anhand einer Repräsentation, fixieren sie eher auffällige aber aufgabenirrelevante Bereiche. Im Gegensatz dazu fixieren HoEr die relevanten Bereiche schneller und kürzer (Gegenfurtner, Lehtinen, & Säljö, 2011). Sie kommen zum selben Schluss, nutzen jedoch heterogene Lösungswege (Jarodzka et al, 2010). Die Autorin der vorliegenden Studie beschreibt in 2 Publikationen ausführlich die Grundlagen der Eyetracking-Methode sowie deren Einsatz bei der Untersuchung von Repräsentations- und Modellierungskompetenz in Kombination mit verbalen Methoden (Ubben, Nitz, Daniel, & Upmeier zu Belzen, 2018; Ubben, Salisbury, & Daniel, 2019).

2.6. Didaktische Relevanz von Diagrammen phylogenetischer Verwandtschaft und Tree-Reading

DpV modellieren die gemeinsame Abstammung von Organismen als Grundprinzip der Evolution auf stark konventionalisierte und abstrahierte Weise. Als biologiespezifische Modelle *von* und *für* evolutive Verwandtschaft (Mahr, 2008, 2009) sind sie essentiell für das Evolutionsverständnis (Baum et al., 2005; Baum & Smith, 2013; Gregory, 2008; Jördens et al., 2011; Meir et al., 2007; O'Hara, 1988). Sie bieten außerhalb der Wissenschaft eine wertvolle Möglichkeit, um Tree-Thinking und somit das grundsätzliche Evolutionsverständnis Lernender zu fördern (Baum et al., 2005) und können so gezielt im Biologieunterricht eingesetzt werden (Novick & Catley, 2014).

Das Thema Evolution ist laut des Berliner Rahmenlehrplans das erste Mal am Ende der 10. Jahrgangsstufe für die „Vernetzung und punktuelle Vertiefung des bereits erworbenen Wissens über Ökologie, Genetik und Evolution“ (Senatsverwaltung für Bildung, Jugend und Sport Berlin, Ministerium für Bildung, Jugend und Sport des Landes Brandenburg, & Ministerium für Bildung, Wissenschaft und Kultur Mecklenburg-Vorpommern, 2015, S. 38) vorgesehen und dient damit als Rahmung für bisher Gelerntes. In der Oberstufe war Evolution bislang im letzten Kurshalbjahr der Qualifikationsphase geplant und als „Evolution und Zukunftsfragen“ ebenfalls dazu gedacht, bisher Gelerntes einzuordnen, zu vernetzen und ein „Verständnis des Lebens“ zu erlangen. DpV wurden nur als Stichwort „Stammbaum“ genannt und nicht weiter ausgeführt, die Tiefe der Einbettung von DpV wird somit der Umsetzung im schulinternen Curriculum überlassen (vgl. Senatsverwaltung für Bildung, Jugend und Sport Berlin, Ministerium für Bildung, Jugend und Sport des Landes Brandenburg, & Ministerium für Bildung, Wissenschaft und Kultur Mecklenburg-Vorpommern, 2015, S. 24). Der aktuell gültige Berliner Rahmenlehrplan für die gymnasiale Oberstufe (Senatsverwaltung für Bildung, Jugend und Sport Berlin, & Ministerium für Bildung, Jugend und Sport des Landes Brandenburg, 2021) sieht sowohl in der Einführungsphase der Jahrgangsstufe 11 das Themenfeld „Vielfalt und Entwicklung der Wirbeltiere“ als auch im 4. Kurshalbjahr der Qualifikationsphase das Themenfeld „Vielfalt und Entwicklung des Lebens“ vor. Die Zielsetzungen der Themenfelder sind gesellschaftsrelevanter als in der Vorgängerversion von 2015 und setzen Evolution in den Kontext des Verständnisses von Biodiversität, aber auch Diversity in der Gesellschaft. In beiden Themenfeldern werden DpV konkret benannt. Im Themenfeld „Vielfalt und Entwicklung der Wirbeltiere“ werden sowohl „Stammbäume rezenter Wirbeltiere“ als auch

„Erstellen von Stammbäumen“ festgelegt, aber auch der „Vergleich von Stammbäumen, die auf verschiedenen Grundlagen erstellt werden“ (Senatsverwaltung für Bildung, Jugend und Sport Berlin, & Ministerium für Bildung, Jugend und Sport des Landes Brandenburg, 2021, S. 24).

In den einheitlichen Prüfungsanforderungen für das Abitur waren DpV bisher in den Themenbereich „Stammesgeschichte, Verwandtschaftsbeziehungen, Systematik“ eingeordnet (KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland), 2004b, S. 7). Laut der nationalen Bildungsstandards sollen Schüler:innen die Analyse der „[...] *stammesgeschichtliche[n] Verwandtschaft* [...] *durch kriteriengeleitetes Vergleichen, [...]*“ umsetzen können (KMK, 2004a, S. 14). In den seit 2020 geltenden Bildungsstandards Biologie für die Allgemeine Hochschulreife werden DpV im Zusammenhang von ursprünglichen und abgeleiteten Merkmalen und der menschlichen Evolution im Inhaltsbereich „Vielfalt des Lebens“ aufgeführt, jedoch nicht näher spezifiziert (KMK, 2020, S. 22).

Die Nationale Akademie der Wissenschaften Leopoldina widmete 2017 eine Stellungnahme dem Thema „Evolutionsbiologische Bildung in Schule und Hochschule“ (2017) und attestierte dem Evolutionsunterricht an deutschen Schulen erhebliche Mängel, die bereits in der Ausbildung von Lehrkräften begründet liegen. Sie fordert nachdrücklich, Evolution aus ihrem in Deutschland historisch bedingten Nischendasein zu holen und als Rahmung und roten Faden für den gesamten Biologieunterricht zu etablieren.¹⁸ Dies sollte bei den Lernenden bereits in der Grundschule ansetzen und auch einen größten Stellenwert in der Ausbildung von Lehrkräften bekommen, damit diese wiederum das Thema Evolution sinnvoll unterrichten können (Nationale Akademie der Wissenschaften Leopoldina, 2017).¹⁹ Wie anfangs erwähnt, gilt dies insbesondere für TR, da diese Fähigkeit die Grundlage für das Evolutionsverständnis bildet. Die aktuell gültigen Rahmenlehrpläne bzw. Bildungsstandards bieten hierfür inzwischen mehr Verpflichtung als vor dem Jahr 2020, jedoch obliegt es weiterhin jeder Schule und insbesondere jeder Lehrkraft, diese Anforderungen den wissenschaftlichen Konventionen entsprechend umzusetzen.

¹⁸ Van Dijk und Kattmann forderten dies ebenfalls bereits 2010.

¹⁹ Die Nationale Akademie der Wissenschaften Leopoldina (2017) stellt diesbezüglich Handlungsempfehlungen für die Integration des Thema Evolution vor, die der Stellungnahme zu entnehmen sind.

3. Stand der Forschung

3.1. Tree-Reading

Tree-Thinking wird als unabdingbar für das Evolutionsverständnis eingeordnet (Baum et al., 2005; O'Hara, 1988; Omland et al., 2008). Auch wenn bereits Kinder nach einer kurzen Einführung dazu in der Lage sind, DpV mit wenigen Taxa hinsichtlich einfacher Fragestellungen zu interpretieren (Ainsworth & Saffer, 2010, 2013), werden Schüler:innen (Catley et al., 2013; Novick & Catley, 2014) und Studierenden in den meisten Studien schwache oder - mit stärkerem fachlichen Hintergrund - zumindest mittelmäßige Tree-Reading-Fähigkeiten attestiert (Baum et al., 2005; Catley, Phillips, & Novick, 2013; Halverson & Friedrichsen, 2013; Meir et al., 2007; Morabito et al., 2010; Novick & Catley, 2013, 2014). Diese lassen sich zwar durch Interventionen verbessern, bleiben jedoch häufig weiterhin auf niedrigem Niveau (Phillips, Novick, Catley, & Funk, 2012; Novick & Catley, 2016). Studierende sind dazu in der Lage, Fragen zu Tree-Reading zunächst korrekt zu beantworten. Lässt man sie jedoch ihre Antwort erläutern, nutzen sie häufig nicht *den wissenschaftlichen Konventionen entsprechende* Begründungen (Catley, Phillips, Novick, 2013; Kummer et al., 2019) wie z. B. Oberflächenmerkmale (Reihenfolge der Taxa, Länge der Äste) und zeigen *wissenschaftlich inadäquates Ablesen* (Gregory, 2008, Halverson & Friedrichsen, 2013). Trotz der hohen Relevanz des Themas und der offenkundigen Herausforderungen beschäftigen sich nur verhältnismäßig wenige Autor:innen damit. Für das Unterrichten von Tree-Reading in Schulen und Universitäten gibt es mehrere biologiedidaktische Artikel, die die *wissenschaftlich adäquate* Leseweise von DpV behandeln (z. B. Baum & Offner, 2008; Kong et al., 2016; Nee, 2005; Omland & Cook, 2008; Schramm & Schmiemann, 2019; Weigand, 2015) bzw. Aktivitäten zur Förderung von Tree-Reading und Tree-Building im Unterricht vorschlagen (Goldsmith, 2003; Singer, Hagen, & Sheehy, 2001; Blacquiere et al., 2020; Halverson 2010; Wilson, Long, Momsen, & Bray Speth, 2020; Stenlund, Schönborn, & Tibell, 2021). Diese sind jedoch meist erfahrungs- bzw. theoriebasiert. Mit Blick auf empirische Studien, die sich konkret mit Tree-Reading beschäftigen, zeigen sich mehrere thematische Schwerpunkte: der Einfluss verschiedener Repräsentationsformen von DpV auf Tree-Reading (Kapitel 3.1.1), die Evaluation verschiedener Tree-Reading-Fähigkeiten (bezogen auf DpV-Bestandteile; Kapitel 3.1.2, 3.1.3, 3.1.4) und das Vorkommen von *wissenschaftlich inadäquatem Ablesen* als Gegensatz zum *wissenschaftlich adäquaten* Tree-Reading (Kapitel 2.1.4).

3.1.1. Interpretation verschiedener Repräsentationsformen und Oberflächenmerkmale

DpV werden überwiegend in der Baum- oder Leiterform dargestellt (Kapitel 2.1.1). In einer nicht repräsentativen Lehrbuchstudie von Catley und Novick (2008) wurde das Vorkommen verschiedener Repräsentationsformen von DpV in US-amerikanischen Schul- und Universitätslehrbüchern untersucht. 2021 wurde diese Studie von Schramm und Kolleg:innen wiederaufgelegt und um deutsche Universitätslehrbücher erweitert. Sie kamen zu dem Ergebnis, dass im Gegensatz zur Vorgängerstudie von Catley und Novick (2008) inzwischen die Baumform im Vergleich zu leiterförmigen und anderen DpV überwiegt, bei US-Lehrbüchern mit 87 % der Abbildungen ($N = 792$) sogar noch mehr als bei deutschen mit 75 % ($N = 405$). Außerdem verläuft die Zeitachse bei 54 % der DpV in US-Büchern von links nach rechts, in deutschen Büchern sogar zu knapp 69 %. Ein Zeitverlauf von unten nach oben kommt bei 41 % der DpV in US- und bei knapp 29 % in deutschen Lehrbüchern vor. Merkmale wurden lediglich in 30 % bzw. 28 % der DpV in US- bzw. deutschen Lehrbüchern eingezeichnet. Sowohl in US- als auch in deutschen Schulbüchern wurden überwiegend DpV mit 5 bis 10 bzw. 6 bis 10 Taxa gefunden. Diese Taxa können in 55 % bzw. 60 % der DpV in US- bzw. deutschen Lehrbüchern dem Tierreich zugeordnet werden. Auch wenn in verhältnismäßig wenigen DpV Menschen als Taxon zu finden sind, werden sie doch in der Hälfte dieser Fälle in eine scheinbar wichtige Position innerhalb des DpV gesetzt, z. B. oben rechts. Insgesamt schlussfolgern Schramm und Kolleg:innen, dass im Vergleich zur Vorgängerstudie von Catley und Novick (2008) weniger missverständliche Repräsentationsformen bzw. DpV-Eigenschaften in US- und deutschen Lehrbüchern zu finden sind (Schramm, Jose, & Schmiemann, 2021b).

Eine nicht repräsentative Analyse von Lehrbüchern der Berliner Oberstufe zeigt ebenfalls, dass baumförmige DpV überwiegen. Im Kontext der Modellierkompetenz zeigt sich, dass DpV, die in Aufgaben eingebunden sind, vor allem der Teilkompetenz *Zweck von Modellen* (30 von 47) bzw. *Alternative Modelle* (16 von 47) zuzuordnen sind. Innerhalb der jeweiligen Teilkompetenz entsprechen die Aufgaben wiederum ausschließlich Niveau I und II (*Zweck von Modellen*) bzw. überwiegend (*Alternative Modelle*, 11 von 16; Viebig, 2020). Nur 8 % der DpV in ausgewählten Biologiebüchern für die Oberstufe sind methodisch eingebettet (Ubben, Nitz, Rousseau, & Upmeier zu Belzen, 2015).

Während Wissenschaftler:innen zahlreiche DpV unabhängig von der Repräsentationsform (Halverson & Friedrichsen, 2013) interpretieren und vergleichen, um Erkenntnisse zu

generieren (Baum et al., 2005; Halverson & Friedrichsen, 2013; O'Hara, 1989), spielt die Auswahl der im vorherigen Kapitel beschriebenen Repräsentationsformen von DpV eine große Rolle bei den Tree-Reading-Fähigkeiten von Lernenden. So können je nach Repräsentationsform Tree-Reading-Fähigkeiten von Studierenden stark variieren (Halverson & Friedrichsen, 2013).

Die Baumform kann von Laien (Museumsbesucher:innen; MacDonald, 2014; Novick et al., 2014) und Biologiestudierenden mit unterschiedlich starkem Fachwissen (Novick & Catley, 2007; Novick, Shade, & Catley, 2011; Dees et al., 2017; Dees et al., 2018) sowohl besser interpretiert als auch in andere Repräsentationsformen übersetzt werden als die Leiterform. Novick, Catley und Kolleg:innen (Novick & Catley, 2007; Novick, Catley, & Funk, 2012; Novick & Fuselier, 2019) begründen dies mit dem Gestaltgesetz der guten Fortsetzung, dem zufolge im Leiterformat von Studierenden eine durchgehende Linie gesehen wird statt mehrerer hierarchisch geordneter Knotenpunkte. Auch nach einer Intervention bleibt das Leiterformat weiterhin schwieriger zu interpretieren als das Baumformat (Novick, Catley, & Funk, 2011). Unabhängig von der Orientierung leiterförmiger DpV lesen Studierende sie stets von links nach rechts, was zu unterschiedlich gutem Verständnis führt und folgen mit dem Blick der vermeintlichen „Hauptlinie“. Liegt die Wurzel des DpV auf der linken Seite (Zeitverlauf von links nach rechts), können Studierende es weniger gut in die Baumform übersetzen. Ist oben links eine Klade zu finden (Zeitverlauf von unten nach oben), führt dies zu besserem Verständnis (Novick et al., 2012). Bei baumförmigen DpV hat die Orientierung keinen Einfluss auf die Tree-Reading-Performanz von Studierenden (Phillips et al., 2012).

Der Zeitverlauf - ein kritischer Aspekt zum Verständnis von evolutiver Verwandtschaft - wird von einigen Studierenden (39 % mit wenig Fachwissen; bzw. 21 % mit mehr Fachwissen) nicht grundsätzlich von der Wurzel des DpV zu den Taxa, sondern immer in der gewohnten Leserichtung von links nach rechts vermutet. Bei DpV mit einem Zeitverlauf von unten nach oben wird daher trotzdem das Taxon, das ganz links dargestellt ist, als Vorfahre aller anderen Taxa gesehen (Meir et al., 2007). Werden Merkmale in DpV angegeben, können Studierende häufig die DpV besser interpretieren als ohne die Angabe (Novick, Catley, & Funk, 2010).

3.1.2. Interpretation von *MRCA* und relativer Verwandtschaft

MRCA bzw. die Verwandtschaftsbeziehungen zwischen Taxa, die anhand von *MRCA* modelliert werden, sind der zentrale Bestandteil von DpV und insbesondere von Kladogrammen (Blacquiere & Hoese, 2016; Catley & Novick, 2013; Morabito et al., 2010). Dees und Kolleg:innen (2018) postulieren, dass vor einer thematischen Einführung nur wenige Studierende in der Lage sind, evolutive Verwandtschaftsverhältnisse korrekt zu identifizieren (max. 15 %) bzw. ihre Antwort wissenschaftlich adäquat zu begründen (max. 8 %). Auch nach einer Intervention sind nur jeweils ca. 50 % der Studierenden dazu fähig.

Sollen DpV als grafisch sehr reduzierte Repräsentationen hinsichtlich evolutiver Verwandtschaftsbeziehungen interpretiert werden, zeigen sich bei Lernenden zahlreiche Varianten von *wissenschaftlich inadäquatem Ablesen* (Auflistung in Kapitel 2.1.4). Eine davon ist das Abzählen von Knotenpunkten beim Vergleichen von Verwandtschaftsbeziehungen (Blacquiere & Hoese, 2016; Catley et al., 2013; Dees et al., 2017; Gregory, 2008; ; Halverson, 2011; Halverson & Friedrichsen, 2013; Meir et al., 2007; Phillips et al., 2012), die auch nach einer thematischen Einführung bestehen bleibt (Dees et al., 2018; Halverson, 2011; Novick & Catley, 2012; Phillips et al., 2012). Meir und Kolleg:innen (2007) zählen das Knotenzählen zu den 4 bedeutendsten und häufigsten Formen von *wissenschaftlich inadäquatem Ablesen* und berichten, dass in einem Multiple-Choice-Test bis zu 38 % der Studierenden unabhängig von der Stärke ihres fachlichen Hintergrunds diese Vorgehensweise zeigen. Sowohl bei leiter- als auch baumförmigen DpV konnten 22 % der Antworten von Studierenden dem Knotenzählen zugeordnet werden (Dees et al., 2017).

Auch die oberflächliche Platzierung von Taxa beeinflusst TR. So wird von mindestens einem Viertel (geringes und hohes Fachwissen; Meir et al., 2007) bzw. einem Drittel der Studierenden (Halverson, 2011) diese *Tip Proximity* als Kriterium des Verwandtschaftsgrades herangezogen. Zusätzlich kann auch die Gruppierung von Taxa in Kladen zu *wissenschaftlich inadäquatem Ablesen* führen: Befinden sich Taxa in einer Klade, die häufig *wissenschaftlich inadäquat* aufgrund eines gemeinsamen Lebensraumes o. ä. als näher verwandt zueinander gesehen werden, verstärkt sich das Gruppieren aufgrund des gemeinsamen Lebensraumes. Werden sie jedoch in verschiedenen Kladen gruppiert, wird Verwandtschaft weniger oft mit dem gemeinsamen Lebensraum begründet. Dies wird auf das Gestaltgesetz der Nähe zurückgeführt und kann bei der Konstruktion von DpV für Unterrichtszwecke berücksichtigt werden (Liu & Novick, 2021; Novick & Fuselier, 2019).

Wie generell bei der Interpretation von Repräsentationen (Cook, 2006) ist beim Tree-Reading das Vorwissen bzw. Alltagswissen über dargestellte Taxa ein kritischer Einflussfaktor, der insbesondere bei geringem fachlichem Hintergrund eine wichtige Rolle spielt (Novick & Catley, 2013, 2014). So nutzen Zehntklässler:innen zur Bestimmung von Verwandtschaftsverhältnissen überwiegend allein ihr Vorwissen über dargestellte Taxa und ignorieren das dazugehörige DpV. Auch Studierende nutzen eher ihr Vorwissen bei ihnen bekannten Taxa als die im DpV dargestellten Informationen (Halverson & Friedrichsen, 2013; Novick & Catley, 2013), jedoch insgesamt seltener als Zehntklässler:innen (Novick & Catley, 2013). Organismen werden z. B. anhand ihrer physischen Ähnlichkeit (Gregory, 2008; Halverson et al., 2011) oder eines ähnlichen Lebensraumes (Novick & Catley, 2014) als näher miteinander verwandt eingeordnet als im DpV modelliert ist. Morabito und Kolleg:innen (2008) ließen Studierende aus 2 DpV dasjenige auswählen, das am besten zu vorher behandelten Verwandtschaftsverhältnissen passt und die Auswahl begründen. Abgesehen davon, dass nur 17,3 % der Studierenden das korrekte DpV wählten, begründeten nur 8,1 % der Studierenden mit stärkerem Fachhintergrund und nur 1,5 % der Studierenden mit schwächerem Fachhintergrund ihre Wahl *wissenschaftlich adäquat* mit dem MRCA. Die vorherrschenden Begründungen waren in beiden Gruppen bestimmte gemeinsame Merkmale oder die physische Ähnlichkeit.

Zeigt ein DpV Verwandtschaftsverhältnisse, die dem Alltagswissen von Studierenden widersprechen (*anomalous data*), mindert auch dies die Tree-Reading-Performanz (Catley, Phillips, & Novick, 2013). Im Gegensatz dazu zeigen Schüler:innen und Studierende bessere Tree-Reading-Performanzen bezüglich der Verwandtschaftsbestimmung, wenn DpV ihnen unbekannte oder fiktive Taxa enthalten (Novick & Catley, 2014; Kummer et al., 2019). Außerdem kann das generelle Verständnis von Hierarchien Studierenden auch beim Bestimmen der Verwandtschaftsverhältnisse helfen, wenn sie dieses auf DpV anwenden (Catley, Phillips, & Novick, 2013).

3.1.3. Identifikation und Interpretation von Merkmalen und evolutiven Mechanismen

Werden Merkmale in einem DpV angegeben, kann dies Studierenden helfen, auch leiterförmige DpV besser zu interpretieren, da das Gestaltgesetz der guten Fortsetzung umgangen wird. Dieser Effekt kommt sowohl Studierenden mit geringem als auch mit starkem

Fachhintergrund zugute (Novick et al., 2010). Geht es in einer Aufgabe um den *Verlust* von Merkmalen, wird diese unabhängig vom Kontext deutlich seltener korrekt von Studierenden gelöst und es kommt deutlich öfter zu *wissenschaftlich inadäquatem Ablesen*, als wenn es um *Merkmalsentstehung* geht (Nehm & Ha, 2011). Soll das Verwandtschaftsverhältnis zwischen Taxa verglichen werden, wird jedoch häufig auch *wissenschaftlich inadäquat* die Anzahl von neu entstandenen Merkmalen zwischen Taxa als Kriterium herangezogen (in 50 % der Antworten bei leiter- und 45 % bei baumförmigen DpV; Dees et al., 2017). Diese Form von *wissenschaftlich inadäquatem Ablesen* bleibt auch nach einer Intervention bestehen (Dees et al., 2018). Studierende mit geringem Fachhintergrund sind häufig nicht dazu in der Lage, Merkmale rezenter Taxa (67 %) und insbesondere Merkmale von *MRCA* (91 %) abzuleiten. Auch Studierende mit stärkerem Fachhintergrund zeigen hier häufig Schwierigkeiten (55 bzw. 71 %; Meir et al., 2007).

Liegt eine konvergente Entwicklung von Merkmalen vor, wird nur von 68 % der Studierenden mit stärkerem und nur 54 % mit schwächerem fachlichen Hintergrund erkannt, dass es sich nicht um eine Synapomorphie handelt. Soll hierfür eine verbale Begründung gegeben werden, nennen nur 19,4 % der Studierenden mit starkem und nur 4,6 % mit schwachem fachlichen Hintergrund konkret eine konvergente Entstehung des Merkmals. Die meisten Studierenden beziehen sich darauf, dass das jeweilige Merkmal bei anderen Mitgliedern einer Klade nicht vorliegt (53,2 bzw. 41,5 %).

3.1.4. Vergleichen mehrerer Diagramme phylogenetischer Verwandtschaft

Wissenschaftler:innen vergleichen mehrere DpV unabhängig von ihrer Repräsentationsform, um Hypothesen zu prüfen, Studierenden fällt dies jedoch oft schwer (Halverson, 2011). Ihnen ist - auch mit stärkerem Fachhintergrund - häufig nicht bewusst, dass ein DpV an sämtlichen Knotenpunkten frei rotierbar ist, ohne seinen Inhalt, die Verwandtschaftsbeziehungen zwischen Taxa, zu verändern (Halverson, 2011; Novick & Catley, 2007). Durch das fehlende mentale Rotieren der DpV fällt so auch der Vergleich mehrerer DpV schwer bzw. geschieht nur anhand oberflächlicher Merkmale. Auch unterschiedliche Repräsentationsformen können dazu führen, dass Studierende DpV nicht adäquat miteinander vergleichen können. Studierende ($N = 27$) beziehen sich beim Vergleichen von DpV hauptsächlich auf die oberflächliche Topologie (41 %) bzw. die Repräsentationsform (26 %) und weniger *wissenschaftlich adäquat* auf die Kladen (19 %). Jedoch lässt sich das Vergleichen anhand von

Kladen durch eine gezielte Intervention drastisch verbessern (von 19 auf 78 % der Studierenden; Halverson, 2011). Studierende beziehen sich beim Vergleichen von DpV abgesehen von der Topologie außerdem häufig auf oberflächliche Merkmale wie Astlänge oder die Anordnung der Taxa (Halverson, 2011).

3.2. Erfahrungsgrad

3.2.1. Einfluss des Erfahrungsgrades auf Tree-Reading

Bei evolutiven Problemstellungen ergibt sich hinsichtlich der Lösungsprozesse ein äußerst heterogenes Bild bei Studierenden mit Hauptfach Biologie, die unabhängig von ihrem formalen fachlichen Hintergrund sämtliche Abstufungen zwischen naiven bis hin zu von Wissenschaftler:innen nicht unterscheidbaren Vorgehensweisen zeigen. Während Wissenschaftler:innen kohärente Lösungsansätze in verschiedenen Aufgaben zeigen, wechseln Studierende zwischen verschiedenen Lösungsansätzen und Formen *wissenschaftlich inadäquaten Ablebens*, was auf eine starke Abhängigkeit der Tree-Reading-Performanz von der jeweiligen Aufgabe hinweist (Nehm & Ridgway, 2011). Viele empirische Studien zu Tree-Reading vergleichen Studierende mit unterschiedlich starkem Fachhintergrund (Kapitel 3). Dabei zeigt sich, dass die Performanz der Studierenden mit dem Fachhintergrund zusammenhängt, jedoch auch bei starkem Fachhintergrund oft auf einem mittleren bis schwachen Niveau bleibt (z. B. Catley, Philips, Novick, 2013; Halverson & Friedrichsen, 2013; Meir et al., 2007; Morabito et al., 2010; Novick & Catley, 2013, 2014). Halverson und Friedrichsen (2013) postulieren, dass es ausschließlich Wissenschaftler:innen mit viel Erfahrung vorbehalten ist, über ein wissenschaftliches Niveau (*Level 6: Scientific Use of Representations*) hinaus Expertise im Tree-Thinking zu erreichen (*Level 7: Expert Use of Representations*; Kapitel 2.2).

3.2.2. Einfluss des Erfahrungsgrades auf Modellkompetenz

Fragt man Wissenschaftler:innen danach, wie sie Modelle sehen und wofür sie sie nutzen, nennen sie sowohl den medialen als auch den methodischen Modelleinsatz (Trier, Krüger, & Upmeier zu Belzen, 2014). Aus einer Interviewstudie mit 2 Wissenschaftler:innen geht hervor, dass sie Modelle beispielsweise nutzen, um anwendungsorientiert Aussagen zur Wirksamkeit oder wirtschaftlichem Nutzen eines neuen Medikaments zu machen. Bezogen

auf Grundlagenforschung werden DpV genutzt, um die Ergebnisse zu kommunizieren und um evolutive Verwandtschaftsverhältnisse abzuleiten (France, Compton, & Gilbert, 2011).

Im Gegensatz dazu erkennen Schüler:innen häufig nicht den Mehrwert eines Modells als Instrument zur Erkenntnisgewinnung (Trier et al., 2014), sondern kennen Modelle aus dem MINT-Unterricht überwiegend als Anschauungsobjekt (Borrmann, Reinhardt, Krell, & Krüger, 2014; Krell und Krüger, 2016; van Driel & Verloop, 2002; vgl. Upmeier zu Belzen & Krüger, 2019). Schüler:innen haben also nicht genug Möglichkeiten, um Erfahrungen im methodischen Modelleinsatz zu sammeln (Treagust, Chittleborough, & Mamiala, 2002). Ordnet man Aussagen von Schüler:innen zum Modellverständnis den 3 Niveaustufen des Kompetenzmodells der Modellkompetenz zu (Upmeier zu Belzen & Krüger, 2010), spiegelt sich dies deutlich wider: So werden Schüler:innenaussagen in den Teilkompetenzen *Eigenschaften von Modellen* und *Alternative Modelle* zum Großteil Niveau I zugeordnet (69 % bzw. 44 %), weniger häufig Niveau II (17 % bzw. 31 %) und selten Niveau III (4 % bzw. 8 %). Bei den Teilkompetenzen *Testen* und *Ändern von Modellen* kommt es am häufigsten zu Aussagen auf Niveau II (68 % bzw. 68 %), danach auf Niveau I (30 % bzw. 31 %) und selten auf Niveau III (8 % bzw. 1 %). Im Gegensatz dazu werden bei der Teilkompetenz *Zweck von Modellen* neben Niveau II (43 %) ähnlich viele Aussagen Niveau I und III zugeordnet (25 % bzw. 23 %; Grünkorn, Upmeier zu Belzen & Krüger, 2014).

France (2019) argumentiert, dass Schüler:innen von einer Außenperspektive auf Modelle blicken, während Wissenschaftler:innen Modelle entwickeln, anwenden und hinterfragen und diese somit eng mit deren Arbeit verknüpft sind. Daher kann die Modellkompetenz von Wissenschaftler:innen nicht mit der von Schüler:innen verglichen oder mit den gleichen Instrumenten gemessen werden (France, 2019).

3.2.3. Einfluss des Erfahrungsgrades auf die visuelle Wahrnehmung

Unterschiede zwischen Personen mit unterschiedlichem Erfahrungsgrad finden sich in zahlreichen Kontexten wieder und schlagen sich sowohl in verbalen als auch in visuellen Prozessen nieder. So kommen Personen mit hohem Erfahrungsgrad bei Problemlöseaufgaben zur gleichen Lösung, zeigen jedoch heterogene verbale und visuelle Lösungsprozesse (Jarodzka et al., 2010). Dies wird mit dem unterschiedlichen fachlichen Hintergrund bzw. Forschungsschwerpunkt begründet. Personen mit hohem Erfahrungsgrad fixieren relevante Bereiche von Repräsentationen schneller und kürzer während Personen mit niedrigem

Erfahrungsgrad eher aufgabenirrelevante, auffällige Bereiche fixieren (z. B. Metaanalyse von Gegenfurtner et al., 2011).

Insgesamt wurden bisher in der didaktischen Forschung wenige Eyetracking-Studien durchgeführt, die auch meist eine geringe Stichprobe umfassten und eher explorativ gestaltet waren (vgl. Lenzer & Nehring, 2022, S. 146). Lediglich eine Studie befasste sich dabei bisher mit Tree-Reading-Prozessen, indem untersucht wurde, in welche Richtung Studierende bzw. 2 erfahrene Professor:innen leiterförmige DpV lesen, während sie diese ins Baumformat übertragen. Dabei stellte sich heraus, dass Studierende und Professor:innen DpV von links nach rechts lesen und der „Hauptlinie“ folgen, die zwar visuell auffällig ist, jedoch inhaltlich keine Aussage hat. Tiefergehende Analysen der visuellen Daten wie z. B. Betrachtung bestimmter Bereiche wurden nicht beschrieben (Novick, Stull & Catley, 2012).

4. Forschungsfragen und Hypothesen

Wie in Kapitel 3 erläutert, stellt das *wissenschaftlich adäquate* Interpretieren und Vergleichen von DpV Lernende vor verschiedene Herausforderungen. Dies steht in einem starken Kontrast zu Wissenschaftler:innen, die dazu in der Lage sind, DpV unabhängig von der Repräsentationsform *wissenschaftlich adäquat* zu nutzen. Es finden sich also große Unterschiede in einem Gebiet, das essentiell für das Verständnis des Evolutionskonzeptes als grundlegendes Prinzip der Biologie ist. Die meisten Studien teilen Lösungen von Tree-Reading-Aufgaben in *richtig* und *falsch* ein und nutzen kurze verbale Begründungen der Proband:innen, um zu beurteilen, ob sie *wissenschaftlich adäquat* oder *inadäquat* geantwortet haben (z. B. Dees et al., 2017; Dees et al., 2018; Meir et al., 2007; Novick & Catley, 2013, 2014, 2016). Einige Studien haben zum Ziel, verschiedene Tree-Reading-Fähigkeiten zu identifizieren, nach Anspruch zu ordnen und miteinander in Beziehung zu setzen (Halverson & Friedrichsen, 2013; Kummer et al. 2017, 2019; Novick & Catley, 2013, 2016; Schramm et. al., 2019; Schramm et. al., 2021a). Ausschließlich Studien von Halverson (2009, 2011) bzw. Halverson & Friedrichsen (2013) charakterisieren den Umgang mit DpV als Repräsentationen basierend auf umfassenden verbalen Daten und unterscheiden zwischen oberflächlichem und inhaltlichem TR.

Während also häufig untersucht wurde, *ob* Lernende Aufgaben auf eine bestimmte Weise lösen, ist wenig über das *inwiefern* bekannt. Die vorliegende Studie hat ausgehend davon das Ziel, Tree-Reading-Prozesse sichtbar zu machen, indem individuelle Lösungsprozesse von Proband:innen über mehrere Aufgaben hinweg untersucht werden. Hier liegt der Fokus auf der Interaktion mit dem DpV als Modell. Hinausgehend über die übliche Unterscheidung in *wissenschaftlich adäquat* und *inadäquat* wird in der vorliegenden Studie differenziert, mit welcher Komplexität der Umgang mit dem DpV erfolgt (siehe Stufen in Kapitel 2.2). Um Tree-Reading-Prozesse von Studierenden mit niedrigem Erfahrungsgrad mit DpV (NiEr) charakterisieren und einordnen zu können, werden sie denen von Wissenschaftler:innen mit hohem Erfahrungsgrad mit DpV (HoEr) entgegengestellt. Dies führt zu **Forschungsfrage 1**:

*Inwiefern beeinflusst der Erfahrungsgrad mit DpV (hoch versus niedrig)
Tree-Reading-Prozesse?*

Entsprechend der Ergebnisse vorausgehender Studien (Kapitel 3.1 und 3.2.1) ergeben sich dabei folgende Hypothesen zur Überprüfung:

Hypothese 1a: *Personen mit hohem Erfahrungsgrad lesen und leiten Informationen aus DpV wissenschaftlich adäquat ab.*

Hypothese 1b: *Personen mit niedrigem Erfahrungsgrad nutzen häufig oberflächliche DpV-Merkmale und lesen wissenschaftlich inadäquat ab.*

Wie in Kapitel 3.1.1 dargelegt, gibt es einige Studien zum Einfluss verschiedener Repräsentationsformen von DpV auf TR, die sich auf die Orientierung und Darstellungsweise (baum- vs. leiterförmig) beschränken. Halverson und Friedrichsen (2013; Kapitel 2.2) vertiefen diesen Ansatz, indem sie ihre Strukturierung von Tree-Reading-Fähigkeiten an den Stufen der Repräsentationskompetenz (Kozma & Russell, 2005) orientieren und so unterscheiden, mit welcher Komplexität DpV beim Tree-Reading genutzt werden und wie sich Tree-Reading-Fähigkeiten von Studierenden über die Zeit entwickeln. Es ist jedoch nicht ersichtlich, nach welchen Kriterien die untersuchten Tree-Reading-Aufgaben erstellt wurden, wie komplex sie waren und welche Repräsentationsformen von DpV genutzt wurden.

Eine Hypothese als Modell für evolutive Verwandtschaftsverhältnisse kann in unterschiedlichen Repräsentationsformen dargestellt werden (Passmore et al., 2014). Die verschiedenen Repräsentationsformen bieten hier die Möglichkeit, über das Modell an sich nachzudenken und zu kommunizieren (Nersessian, 2002). DpV als Modelle werden in der Wissenschaft damit sowohl medial zur Weitergabe von Wissen über evolutive Verwandtschaftsverhältnisse eingesetzt als auch methodisch, um neue Erkenntnisse zu generieren (Kapitel 2.3). Jedoch wurden DpV bisher noch nicht unter dem Aspekt der Modellkompetenz (Upmeier zu Belzen & Krüger, 2013) bzw. Modellierkompetenz (Krüger & Upmeier zu Belzen, 2021) untersucht. Als erster Schritt zur Untersuchung von DpV als Modelle von und für Evolution ergibt sich **Forschungsfrage 2:**

Inwiefern beeinflusst der Modelleinsatz (medial versus methodisch) Tree-Reading-Prozesse?

Aufgrund der Struktur des Kompetenzmodells der Modellkompetenz (Kapitel 2.3.2) ist zunächst folgende Hypothese zu überprüfen:

Hypothese 2: *Der methodische Modelleinsatz erfordert elaboriertere Lösungsprozesse.*

Während in der Wissenschaft Modelle sowie DpV im Speziellen sowohl medial als auch methodisch eingesetzt werden, findet in der Schule überwiegend ein medialer Umgang mit Modellen statt (Borrmann et al, 2014; van Driel & Verloop, 2002; Krell und Krüger, 2016; vgl. Upmeier zu Belzen und Krüger, 2019), sodass Lernende wenig Erfahrung im erkenntnisgenerierenden Umgang mit Modellen haben (Treagust, Chittleborough, & Mamiala, 2002; Kapitel 3.2.2). Auch Tree-Reading-Fähigkeiten sind bei Lernenden schwächer ausgeprägt als bei Wissenschaftler:innen (Kapitel 3.2.1). Erstmals wird in der vorliegenden Studie die Interaktion von Modellkompetenz und Erfahrungsgrad am Beispiel von DpV untersucht. Als Kombination der 2 vorausgehenden Forschungsfragen ergibt sich hier **Forschungsfrage 3** mit den zu prüfenden Hypothesen:

Inwiefern interagieren Erfahrungsgrad und Modelleinsatz bei Tree-Reading-Prozessen?

Hypothese 3a: *Personen mit hohem Erfahrungsgrad zeigen bei beiden Modelleinsätzen die wissenschaftlich erforderlichen Tree-Reading-Fähigkeiten.*

Hypothese 3b: *Personen mit niedrigem Erfahrungsgrad zeigen bei methodischem Modelleinsatz eingeschränkte Tree-Reading-Fähigkeiten.*

5. Material und Methoden

5.1. Methodische Entscheidungen

Ziel der vorliegenden Arbeit ist die Untersuchung von Tree-Reading-Prozessen in Abhängigkeit vom Erfahrungsgrad mit DpV und dem Modelleinsatz. Die meisten Studien nutzen ausschließlich verbale Daten zur Untersuchung von Tree-Reading-Prozessen (z. B. Halverson & Friedrichsen, 2013). Beim Lesen und Interpretieren von DpV als stark abstrahierte Repräsentationen sind visuelle Wahrnehmungsprozesse zentral, die jedoch bisher kaum untersucht wurden (Novick et al., 2012). In der vorliegenden Arbeit werden verbale und visuelle Datenquellen kombiniert, da dies tiefgehende Einblicke in kognitive Prozesse während der Informationsverarbeitung erlaubt (Stieff et al., 2011). Dieser Ansatz der Datentriangulation hat sich in der Lehr- und Lernforschung zunehmend etabliert (vgl. Ubben et al., 2019, S. 106-107; Klein, Graulich, Kuhn & Schindler, 2022) und wird auch in der Expertiseforschung angewendet (z. B. Gegenfurtner et al., 2011; Jarodzka, Scheiter, Gerjets, & van Gog, 2010).

Visuelle Daten liefern Auskunft darüber, welche Bereiche einer Repräsentation Proband:innen wie lange und in welcher Reihenfolge fixieren. Aufschlüsse über die zugrunde liegenden kognitiven Prozesse können durch die Analyse verbaler Äußerungen der Proband:innen gewonnen werden. Stieff und Kolleg:innen (2011) beschreiben den Mehrwert dieser Datentriangulation für Forschungsdesigns, die den Umgang mit mehreren Repräsentationen vorsehen. Bezogen auf die vorliegende Studie betrifft dies vor allem die Teilkompetenz *Alternative Modelle*. So können einerseits Datenlücken wie bspw. fehlende sprachliche Bezüge gefüllt werden. Außerdem kann festgestellt werden, ob Proband:innen für ihre Aufgabenbearbeitung mehr Repräsentationen bzw. Bereiche nutzen als sie verbal erwähnen (Stieff et al., 2011). Die Kombination von verbalen und visuellen Daten kann durch das Laute Denken umgesetzt werden, das während (*Concurrent* bzw. *Congruent Think-Aloud*, CTA; Ericsson & Simon 1993; van Someren, Barnard, & Sandberg, 1994) oder nach der Bearbeitung einer Aufgabe anhand einer grafischen Rekapitulation der Blickbewegungen stattfindet (*cued Retrospective Think-Aloud*, cRTA; van Gog, Paas, & van Merriënboer, 2005). CTA liefert mehr „action infos“ über das „why“ und „how“ der Aufgabenbearbeitung (van Gog et al., 2005) und „levels of processing involved in studying certain learning materials (Holmqvist et al., 2011, S. 103) während *cued RTA* mehr Schlüsse über metakognitive Prozesse zulässt (van Gog et al.,

2005). Für die vorliegende Studie wurde *CTA* ausgewählt, um die unmittelbaren kognitiven Prozesse während der Aufgabenbearbeitung zu untersuchen. In Tabelle 5 werden die damit verbundenen Vor- und Nachteile gegenübergestellt, die jeweils im Vergleich mit *cued RTA* zu sehen sind.

Tabelle 5: Vor- und Nachteile von Concurrent Think Aloud (CTA) im Vergleich zu cued Retrospective Think Aloud (cRTA). Zahlen geben Quelle an: (1) Holmqvist et al. (2011, S. 103); (2) Bojko (2013).

| Vorteile CTA gegenüber cRTA | Nachteile CTA gegenüber cRTA |
|---|--|
| Verbale und visuelle Daten synchron (1) | Ggf. Einfluss auf Dauer und Art der Fixationen (1; 2) |
| Unmittelbar, keine retrospektive Korrektur der Aussage (1, 2) | Erhöhter <i>Work Load</i> (2) |
| Keine Vorstrukturierung der Aussage (1) | Langsame Performanz (1) |
| Keine „ausgedachten“ Aussagen (2) | Ggf. negativer Einfluss auf Güte der Performanz (1; 2) |
| Zeitsparend, da keine weiteren Durchgänge nötig (1, 2) | |
| Kein Vergessen oder Auslassen von Aussagen (1; 2) | |

Neben der besseren Darstellbarkeit von detaillierten Lösungsprozessen kann *CTA* auch bei diesem speziellen Versuchssetting als Hilfe für die Proband:innen angesehen werden. Bedingt durch das Vorgehen beim Remote-Eyetracking (Kapitel 5.6) können Proband:innen keine Hilfsmittel zur Aufgabenbearbeitung nutzen. *CTA* kann so zur Strukturierung der Gedanken während des Lösungsprozesses (Bojko, 2013, S. 109) als „verbaler Stift“ genutzt werden. Da die Proband:innen nicht unter Zeitdruck standen, selbst entscheiden konnten, wann sie die Aufgabenbearbeitung beenden wollen und „Denkpausen“ einlegen konnten, fallen die in Tabelle 5 aufgeführten Nachteile des *CTAs* vermutlich weniger stark ins Gewicht.

5.2. Operationalisierung

Aus den Forschungsfragen (Kapitel 4) ergeben sich die **unabhängigen Variablen Erfahrungsgrad und Modelleinsatz** sowie die **abhängige Variable Tree-Reading-Prozesse**, die wie folgt operationalisiert werden:

Unabhängige Variable 1: Erfahrungsgrad

Wie in Kapitel 2.4 erläutert wird der *a priori* angenommene Erfahrungsgrad mit DpV als Einteilung in 2 Vergleichsgruppen genutzt. Wissenschaftler:innen aus dem Feld der Evolutionsbiologie werden hiernach als „Personen mit hohem Erfahrungsgrad“ (HoEr) definiert. Lehramtsstudierende der Biologie, die Tree-Reading im Rahmen ihres Studiums bis zu einem gewissen Grad erlernen, als „Personen mit niedrigem Erfahrungsgrad“ (NiEr).

Hoher Erfahrungsgrad (HoEr): Wissenschaftler:innen mit abgeschlossenem Studium und Forschungsschwerpunkt Taxonomie, Systematik, Paläontologie o. ä..

Niedriger Erfahrungsgrad (NiEr): Studierende aus dem *Master of Education* Biologie, die ein Modul zur Morphologie und Evolution der Tiere erfolgreich abgeschlossen haben und somit die Grundlagen des Tree-Reading im Studium behandelt haben.

Unabhängige Variable 2: Modelleinsatz

Wie in Kapitel 2.3.2 erläutert können Modelle medial oder methodisch eingesetzt werden. Die Operationalisierung der Forschungsfragen erfolgte auf Grundlage der 3 Niveaus des Kompetenzmodells der Modellkompetenz (Upmeier zu Belzen & Krüger, 2010), da das um Niveau IIIa erweiterte Kompetenzmodell der Modellierkompetenz (Krüger & Upmeier zu Belzen, 2021) zu diesem Zeitpunkt noch nicht veröffentlicht war. Die Abgrenzung von Niveau I und II gegenüber Niveau III bzw. Niveau IIIa und IIIb ist jedoch beiden Modellen gemein.

Medialer Modelleinsatz: In der Aufgabe werden DpV als Modelle im Sinne von Niveau I und II des Kompetenzmodells der Modellkompetenz (Upmeier zu Belzen & Krüger, 2010) eingesetzt.

Methodischer Modelleinsatz: In der Aufgabe werden DpV als Modelle im Sinne von Niveau III des Kompetenzmodells der Modellkompetenz (Upmeier zu Belzen & Krüger, 2010) eingesetzt.

Abhängige Variable: Tree-Reading-Prozesse

Wie in Kapitel 2.1.2 definiert ist Tree-Reading das Interpretieren und Vergleichen mehrerer DpV (Halverson, 2011). Der Tree-Reading-Prozess, also der Verlauf des Interpretierens und Vergleichens von DpV, wird mithilfe der in Kapitel 2.2 erläuterten Stufen der Tree-Reading-Fähigkeiten charakterisiert:

TR_0: Oberflächliches Ablesen von Informationen aus einem oder mehreren DpV ohne Berücksichtigung der wissenschaftlichen Konventionen bzw. des Inhalts.

TR_1: Ablesen von Informationen aus einem oder mehreren DpV.

TR_1a wissenschaftlich adäquat: anhand wissenschaftlicher Konventionen.

TR_1b wissenschaftlich inadäquat: ohne Berücksichtigung wissenschaftlicher Konventionen.

TR_2: Erklären, Ableiten und Vorhersagen von Informationen, die nicht direkt im DpV dargestellt sind.

Kontroll-Variable 1: Tree-Reading-Fähigkeiten

In Abgrenzung zum Tree-Reading-Prozess sind die Tree-Reading-Fähigkeiten eine *a priori* vorliegende Größe, die die Tree-Reading-Prozesse ggf. beeinflusst. Sie wird mithilfe des *Basic Evolutionary Tree-Thinking Skills Instrument* erhoben, das als Multiple-Choice-Diagnoseinstrument für den Biologieunterricht entwickelt wurde (BETTSI; Bucklin et al., 2019; Jenkins et al. 2022). In der vorliegenden Studie wurde eine vorläufige Version übersetzt und angepasst, die im Jahr 2016 von Kristy Daniel, einer der Autor:innen von BETTSI, freundlicherweise zur Verfügung gestellt wurde (Anhang 1).

Kontroll-Variable 2: Modellkompetenz

Die Modellkompetenz ist die *a priori* vorliegende Größe, definiert von Upmeier zu Belzen und Krüger (2010; Kapitel 2.3.2). Jede Teilkompetenz wird mithilfe von offenen Fragen standardisiert getestet (Krell & Krüger, 2016).

Kontroll-Variable 3: Räumliches Denken

Das räumliche Denken hat ggf. einen Einfluss auf das mentale Rotieren von DpV (z. B. Gregory, 2008). Es wird mittels des 3DW-Tests (Gittler, 1990) ermittelt.

Kontroll-Variable 4: Kognitive Verarbeitungsgeschwindigkeit/ *Mental Speed*

Da die Erfahrungsgruppen klein sind und nicht randomisiert zusammengestellt wurden, wird die kognitive Verarbeitungsgeschwindigkeit als basales Maß für Intelligenz als potentieller Einflussfaktor auf die Lösungsprozesse erhoben (Bortz & Döring, 2006, S. 54). Sie wird anhand des Zahlenverbindungstests (ZVT; Oswald & Roth, 1987) ermittelt.

5.3. Pilotierung *BETTSI* und Analyse der anderen Kontrollvariablen

Die vorläufige *BETTSI*-Version wurde von der Autorin der vorliegenden Studie aus dem Englischen ins Deutsche übersetzt. Die sprachliche und sachliche Richtigkeit wurde von einem Biologieprofessor mit Englisch als Muttersprache und exzellenten Deutschkenntnissen überprüft, der nicht als Proband an der vorliegenden Studie teilnahm. Neben sprachlichen Anpassungen wurde auch das Layout für eine bessere Übersichtlichkeit leicht angepasst. Der übersetzte *BETTSI* wurde von Studierenden des *Master of Education* ($N = 29$) absolviert, die analog zu den Proband:innen der vorliegenden Studie ein Modul über Systematik der Tiere bzw. Pflanzen oder vergleichbares erfolgreich abgeschlossen haben, jedoch später nicht an der Studie teilnahmen. Die interne Konsistenz des übersetzten Fragebogens wurde in SPSS (IBM) anhand des Reliabilitätskoeffizienten Cronbachs α berechnet (Cronbach, 1951).

Für die anderen Kontrollvariablen und die Selbsteinschätzungen wurden standardisierte Tests genutzt. Für die Berechnung der Unterschiede zwischen HoEr und NiEr wurde ein ungepaarter t-Test durchgeführt und anschließend die Effektstärke mit Cohens d ermittelt (Cohen, 1988; Cohens $d > 0,2$ = schwacher Effekt, $d > 0,5$ = mittlerer Effekt; $d > 0,8$ = starker Effekt).

5.4. Aufgabenkonstruktion

5.4.1. Auswahl der Teilkompetenzen

Das in Kapitel 2.3.2 beschriebene Kompetenzmodell der Modellkompetenz (Upmeyer zu Belzen & Krüger, 2010) bildet die Grundlage für die Konstruktion von Aufgaben, in denen DpV medial bzw. methodisch eingesetzt werden. Zunächst wurden dazu die 5 Teilkompetenzen geprüft, inwiefern sie auf Tree-Reading anwendbar sind. Es ergaben sich Passungen für die Teilkompetenz *Alternative Modelle*, *Zweck von Modellen* und *Testen von Modellen*:

DpV mit derselben zugrundeliegenden Hypothese können in zahlreichen unterschiedlichen Repräsentationsformen, also entsprechend der Teilkompetenz *Alternative Modelle*, dargestellt werden. Dabei unterscheidet sich die äußere Form des Modellobjekts, jedoch nicht der Inhalt (medialer Modelleinsatz, *Alternative Modelle* Niveau I). Eine Hypothese über evolutive Verwandtschaftsverhältnisse als Ausgangsobjekt kann außerdem zu unterschiedlichen Modellen mit verschiedenen Fokussen führen (medialer Modelleinsatz, *Alternative Modelle* Niveau II). Eine Hypothese über evolutive Verwandtschaft kann z. B. mit dem Fokus auf gemeinsame Vorfahren und Merkmale oder mit der genetischen Distanz zwischen verschiedenen Taxa dargestellt werden. Anhand von DpV können außerdem verschiedene Hypothesen modelliert und verglichen werden, um neue Erkenntnisse über evolutive Verwandtschaftsverhältnisse derselben Taxa zu gewinnen (methodischer Modelleinsatz, *Alternative Modelle* Niveau III). Auf Grundlage von Daten aus DNA-Analysen, paläontologischen Funden und dem Vergleich der Knöchelstruktur wurden bspw. 3 Hypothesen über die verwandtschaftliche Einordnung der Wale in das Monophylum der Paarhufer aufgestellt, die von Wissenschaftler:innen miteinander verglichen und diskutiert werden (z. B. Thewissen, Madar, & Hussain, 1998).

In der Teilkompetenz *Zweck von Modellen* kann ein DpV als Medium zur Darstellung von evolutiven Verwandtschaftsverhältnissen eingesetzt werden (medialer Modelleinsatz, *Zweck von Modellen* Niveau I). Bekannte Zusammenhänge wie z. B. letzte gemeinsame Vorfahren oder Synapomorphien können mithilfe von DpV erklärt werden (medialer Modelleinsatz, *Zweck von Modellen* Niveau II). DpV können außerdem dazu genutzt werden, Voraussagen zu treffen (methodischer Modelleinsatz, *Zweck von Modellen*, Niveau III). Werden bspw. Fossilien entdeckt, können diese in DpV eingeordnet werden bzw. es können auch ohne Fossilienfunde Voraussagen über bestimmte Merkmale gemeinsamer Vorfahren getroffen werden.

Bei der Teilkompetenz *Testen von Modellen* kann das DpV als Modellobjekt bspw. darauf überprüft werden, ob es den wissenschaftlichen Standards entspricht (medialer Modelleinsatz, *Testen von Modellen*, Niveau I). Die Passung des DpV als Hypothese mit der Realität, also z. B. dem Vorkommen von Merkmalen (medialer Modelleinsatz, *Testen von Modellen*, Niveau II) ist ebenfalls möglich. Hypothesen über Merkmale o. ä. können mithilfe von DpV überprüft werden (methodischer Modelleinsatz, *Testen von Modellen*, Niveau III).

Im Gegensatz dazu sind die restlichen 2 Teilkompetenzen nicht auf Tree-Reading anwendbar. Die Teilkompetenz *Eigenschaften von Modellen* betrifft die Beziehung zwischen Modell und Original. Dies ist eher als übergreifender Aspekt des Tree-Thinking zu betrachten und somit dem Tree-Reading übergeordnet. Die Teilkompetenz *Ändern von Modellen* entspricht eher Tree-Building, da DpV bei Nicht-Passung angepasst werden.

Im Sinne der Zeitökonomie beim Versuchsablauf wurden 2 der 3 passenden Teilkompetenzen ausgewählt: *Zweck von Modellen* und *Alternative Modelle*. Hier wurde jeweils für den medialen Modelleinsatz Aufgabe Niveau II ausgewählt, da Niveau I sehr basal ist und ein zu großer Abstand zu Niveau III vorliegt.

5.4.2. Konstruktion der Aufgaben anhand der Teilkompetenzen

Für beide Teilkompetenzen wurden je 2 Aufgaben generiert, die aus je 3 Aufgaben bestehen. Der Aufgabenstamm und das bzw. die gezeigten DpV sind zwischen den Aufgaben identisch, die Aufgabenstellung unterscheiden sich. Der Aufgabenstamm enthält die Informationen, die zur Lösung der methodischen Aufgabe nötig ist. Die erste Aufgabenstellung dient dazu, dass sich die Proband:innen in DpV und Aufgabenstamm einarbeiten. Danach folgt je eine Aufgabenstellung mit medialem und eine mit methodischem Modelleinsatz. Hierzu wurden den Modelleinsätzen in der jeweiligen Teilkompetenz Tree-Reading-Fähigkeiten zugeordnet und Aufgabenstellungen formuliert (Tabelle 6).

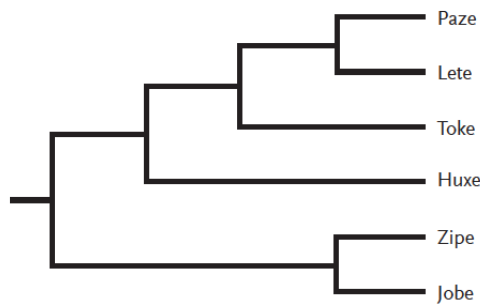
Tabelle 6: Aufgabenstellungen gemäß Teilkompetenz, Modelleinsatz und Tree-Reading-Fähigkeit für die Teilkompetenzen Zweck von Modellen und Alternative Modelle. Pro Teilkompetenz gibt es 2 Aufgaben, jede Aufgabe besteht aus 3 Teilen: Einarbeitung, medialer bzw. methodischer Modelleinsatz. Zweck1: Aufgabe 1 von 2 der Teilkompetenz Zweck von Modellen, analog dazu Zweck2, Alt1 und Alt2.

| Modelleinsatz | Tree-Reading-Fähigkeit | Aufgabenstellung |
|--|---|---|
| Teilkompetenz Zweck von Modellen | | |
| Einarbeitung | Ohne Vorgabe (Zweck1 und 2) | <i>Was sehen Sie auf der Abbildung? Sagen Sie alles, was Sie denken.</i> |
| Medial | <i>MRCA identifizieren + beschreiben (Zweck1 und 2)</i> | <i>Welche Stelle markiert den letzten gemeinsamen Vorfahren von [Taxon a] und [Taxon b].</i> |
| | <i>Merkmal eines MRCA ableiten anhand des Sparsamkeitsprinzips (Zweck1)</i> | <i>Welche Hypothese kann über das [Merkmal] des letzten gemeinsamen Vorfahren aller gezeigten Taxa unter Anwendung des Sparsamkeitsprinzips aufgestellt werden?</i> |
| Methodisch | <i>Konvergenz ableiten (Zweck2)</i> | <i>Welche Hypothese kann über die Entstehung des Merkmals [Merkmal] bei [Taxon a], [Taxon b] und [Taxon c] aufgestellt werden? Begründen Sie Ihre Antwort.</i> |
| Teilkompetenz Alternative Modelle | | |
| Einarbeitung | Ohne Vorgabe (Alt1 und 2) | <i>Was sehen Sie auf den Abbildungen? Sagen Sie alles, was Sie denken.</i> |
| Medial | 2 DpV vergleichen (Alt1 und 2) | <i>Welche Unterschiede und Gemeinsamkeiten weisen die Stammbäume auf? Vergleichen Sie sie miteinander.</i> |
| | <i>Passendes DpV identifizieren (Alt1)</i> | <i>Welcher der beiden Stammbäume unterstützt die genannte Hypothese unter Anwendung des Sparsamkeitsprinzips? Begründen Sie Ihre Antwort.</i> |
| Methodisch | <i>Passendes DpV identifizieren und Hypothese generieren (Alt2)</i> | <i>Welche Hypothese über das Monophylum, das [Taxon a] und [Taxon b] enthält, kann mithilfe der genetischen Daten aufgestellt werden? Begründen Sie Ihre Antwort.</i> |

5.4.3. Konstruktion der Repräsentationen und optische Gestaltung der Aufgaben

Aus der methodischen Entscheidung, visuelle und verbale Daten zu kombinieren, resultieren verschiedene Anforderungen sowohl für die Repräsentationsform der DpV als auch die Konstruktion des Aufgabenstamms. So ist für die Vergleichbarkeit der visuellen Daten zwischen den Aufgaben eine einheitliche Repräsentationsform für die DpV notwendig, für die das Baumformat mit rechtwinkligen Ecken mit einem Zeitverlauf von links nach rechts gewählt wurde, sodass die Taxa auf der rechten Seite jedes DpV waagrecht ausgerichtet sind (Abbildung 4). Dies unterstützt auch die Lesbarkeit der DpV und der Taxanamen, außerdem betrachten Proband:innen Stimuli eher horizontal als vertikal (Holmqvist et al., 2011, S. 80). Informationen über Merkmale sind schriftlich in den Aufgabenstamm eingebettet, um die DpV optisch einheitlich zu gestalten. Dabei bleibt zu beachten, dass die Angabe von Merkmalen innerhalb eines DpV das Tree-Reading erleichtern kann (Novick et al., 2010), die Analyse von visuellen Daten jedoch erschwert wird, da das DpV optisch verdichtet wird. Da die Aufmerksamkeit von Proband:innen im Zentrum des Bildschirms ist (Holmqvist et al., 2011, S. 80), sind die DpV zentral angeordnet.

Für die Bearbeitung der Einarbeitungs-Aufgabe und der Aufgabe mit medialem Modelleinsatz wird der Aufgabenstamm nicht benötigt, da er nur Informationen für die Aufgabe mit methodischem Modelleinsatz enthält. Zur Erhaltung einer optischen Kontinuität wird er jedoch bei allen Aufgaben angezeigt. Für die Taxa wurden fiktive Namen generiert, die immer aus 4 Buchstaben bestehen und mit dem Vokal „e“ enden. Dies vermeidet einerseits eine Beeinflussung der Aufgabenbearbeitung durch Vorwissen (Halverson & Friedrichsen, 2013; Novick & Catley, 2013) und generiert andererseits möglichst einheitliche visuelle Stimuli. Abbildung 4 veranschaulicht den Aufbau der Aufgaben anhand der Aufgabe *Zweck2* im medialen Modelleinsatz. Alle Aufgaben sind Anhang 2 bis 5 zu entnehmen.



Die Analyse verschiedener genetischer Marker ergab diesen Stammbaum. Eine zusätzliche morphologische Analyse zeigt, dass nur Paze, Lete und Zipe Beinpanzer haben. Weder die anderen Taxa noch zahlreiche gefundene Fossilien weisen Beinpanzer auf.

Welche Stelle markiert den letzten gemeinsamen Vorfahren von Jobe und Huxe?
Beschreiben Sie die Stelle.

Abbildung 4: Beispiel für die visuelle Darstellung einer Aufgabe (Zweck2 medialer Modelleinsatz) mit liegendem DpV, Aufgabenstamm und Aufgabenstellung.

5.5. Proband:innen

Proband:innen mit niedrigem Erfahrungsgrad (NiEr) wurden aus einem Studiengang *Master of Education* Biologie rekrutiert. Die Versuchsleiterin (Autorin der vorliegenden Studie) informierte in Seminaren zu ihrem Versuchsvorhaben und schrieb Seminarteilnehmer:innen per Mail mit einem Informationsbrief an. Für die Auswahl der Proband:innen mit hohem Erfahrungsgrad (HoEr) recherchierte die Versuchsleiterin auf den Homepages von Forschungseinrichtungen mit Schwerpunkt Biodiversität und Genetik und schrieb Personen direkt per Mail mit einem Informationsbrief an. Die Teilnahme am Versuch erfolgte freiwillig und wurde nicht finanziell oder anderweitig entlohnt.

5.6. Versuchsaufbau und -ablauf

Zur Randomisierung der Aufgabenreihenfolge wurden 2 verschiedene Versionen erstellt (Version 1: Aufgabe *Zweck1*, *Alt1*, *Zweck2*, *Alt2*; Version 2: Aufgabe *Alt2*, *Zweck2*, *Alt1*, *Zweck1*; jeweils alle 3 Aufgaben direkt nacheinander). Version 1 und Version 2 wurden immer abwechselnd für die Proband:innen innerhalb einer Erfahrungsgruppe eingesetzt, beginnend mit Version 1.

Die visuellen Daten wurden mit einem Remote Eyetracker RED 5 von SMI (SensoMotoric Instruments) erhoben, der mit einem Laptop (DELL Latitude E6530) zur Versuchssteuerung und Datenerfassung verbunden war. Die Stimuli wurden auf einem Monitor (22 Zoll,

Auflösung 1680 x 1050 Pixel) gezeigt, der Remote Eyetracker war direkt unter dem Bildschirm befestigt. Auf dem Monitor war eine Webcam (Logitech) mit integriertem Mikrofon angebracht (Abbildung 5a). Dieses wurde zur Aufnahme der verbalen Daten genutzt. Außerdem wurden die Proband:innen während des Versuchs gefilmt, um im Falle von Datenanomalien Gründe für diese zu finden (z. B. starke Bewegungen der Proband:innen). Der Monitor mit Eyetracker und Webcam befand sich in einem Schrank, der in der Höhe verstellbar und somit an die Größe der Proband:innen anpassbar war. Der Monitor stand im rechten Winkel zu einem abgedunkelten Fenster. Die Schranktüren wurden als zusätzlicher Lichtschutz eingesetzt. Auf dem Schreibtisch links vom Schrank stand der Steuerungslaptop, außerdem wurden dort die Paper-Pencil-Tests durchgeführt (Abbildung 5b). Die Länge des Verbindungskabels zwischen Eyetracker und Laptop limitierte die Entfernung der Versuchsleiterin zu den Proband:innen auf ca. 1,5 m, sodass die Schranktüren ebenfalls als räumliche Abtrennung dienten. Eine permanente Anwesenheit der Versuchsleiterin zur Überwachung der korrekten Datenerfassung war notwendig. Die Proband:innen saßen vor dem Monitor und konnten mit der Leertaste der Computertastatur selbstständig den nächsten Stimulus auslösen.

Die Kalibrierung erfolgte mit der Software iViewX (SensoMotoric Instruments) mit einer Toleranz von $0,5^\circ$. Die Stimuli wurden mithilfe der Software Experiment Center (SensoMotoric Instruments) präsentiert, die auch die visuellen und verbalen Daten erfasste. Die 4 Kontrollvariablen (Kapitel 5.2) und der allgemeine Fragebogen wurden mittels Paper-Pencil-Tests untersucht.

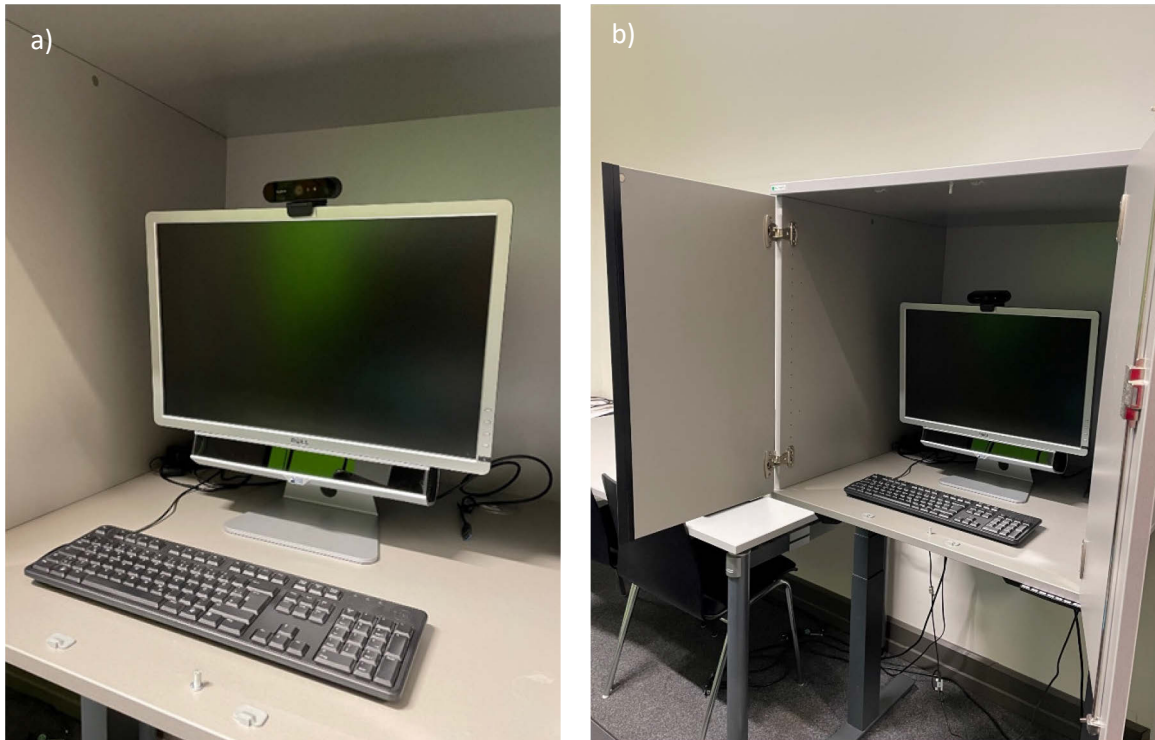


Abbildung 5: a) Remote Eyetracker von Sensomotoric Instruments unter Monitor montiert, Webcam auf Monitor und Tastatur davor. b) Aufbau in höhenverstellbarem Schrank mit Türen als Sicht- und Blendschutz. Hinter linker Tür Schreibtisch mit Steuerungslaptop für Versuchsleiterin.

Zunächst wurde die Proband:in begrüßt und gefragt, ob sie bereit für die Versuchsteilnahme sei. Die verschiedenen Teile des Versuches fanden in der folgenden Reihenfolge statt:

1. Erhebung der Kontrollvariablen (1) und Einverständnis

- a) Ausfüllen allgemeiner Fragebogen
- b) Ausfüllen Einverständniserklärung
- c) Durchführung 3DW-Test
- d) Durchführung ZVT

2. Erhebung der visuellen und verbalen Daten

- a) Instruktion Übungsaufgabe
- b) Übungsaufgabe
- c) Erläuterung Fachbegriffe
- d) Instruktion Aufgaben
- e) Aufgabe Zweck1 (Version 1)/ Aufgabe Alt2 (Version 2)
- f) Aufgabe Alt1 (Version 1)/ Aufgabe Zweck2 (Version 2)
- g) Aufgabe Alt2 (Version 1)/ Aufgabe Zweck1 (Version 2)
- h) Aufgabe Zweck2 (Version 1)/ Aufgabe Alt1 (Version 2)

3. Erhebung der Kontrollvariablen (2) und Abschluss

- a) Durchführung Modellkompetenztest
- b) Durchführung *BETTSI*
- c) Verabschiedung

Vor der Bearbeitung der Übungsaufgabe bzw. jeder Aufgabe wurde neu kalibriert. Außerdem wurde erst ein zentriertes schwarzes Kreuz auf weißem Hintergrund gezeigt, um die Aufmerksamkeit der Proband:in initial standardisiert auf die Mitte des Bildschirms zu lenken. Nach der Übungsaufgabe bzw. nach jeder Aufgabe konnten die Proband:innen eine Pause machen. Der gesamte Versuchsablauf dauerte ca. 1 bis 1,5 Stunden.

5.7. Datenaufbereitung

Mithilfe der Transkription werden verbale Daten verschriftlicht (Dresing & Pehl, 2018; Kuckartz, 2010) und somit für die spätere Kodierung aufbereitet. Die Transkription erfolgte mit der Software MAXQDA Version 18 (VERBI Software GmbH). Um eine einheitliche Transkription zu gewährleisten, wurden Transkriptionsregeln festgelegt (Anhang 6) und ein Beispieltranskript angefertigt. Die Transkription erfolgte durch 2 studentische Mitarbeitende und die Versuchsleiterin. Dafür wurde ein einfaches Transkriptionssystem gewählt, da der Fokus auf der Semantik, also dem Inhalt des *CTA* liegt und nicht auf nonverbalen Informationen oder sprachlichen Besonderheiten (Dresing & Pehl, 2018). Interaktionen mit Gegenständen oder anderen Personen waren durch das Versuchssetting nahezu ausgeschlossen - ausgenommen von vereinzelt Fragen der Proband:innen an die Versuchsleiterin und deren Antwort während des *CTA*. Beim Transkribieren erfolgte eine „Glättung“ des Inhaltes durch Anpassung an die deutsche Standardsprache unter Beibehaltung der ursprünglichen Grammatik (vgl. Dresing & Pehl, 2018; Kuckartz, 2010). Nach jedem Satz wurde ein Absatz und damit verbunden eine Zeitmarke eingefügt. Ein Satz ist die kleinste Analyseeinheit (Mayring, 2015). Hingen mehrere Sätze inhaltlich zusammen, wurden diese Sätze durch ein Memo verbunden. Eine Analyseeinheit kann somit aus einem bis mehreren Sätzen bestehen. Längere Pausen (> 5 sec) wurden als eigener Absatz transkribiert und als „keine Verbalisierung“ kodiert.

Die visuellen Daten wurden mithilfe der Software BeGaze (SensoMotoric Instruments) verarbeitet. Für die Aufgabe *Alt1* im medialen Modelleinsatz wurden hierzu *Areas of Interest* (AOI) erstellt, eine für jedes DpV sowie eine für Aufgabenstamm und Aufgabenstellung (Abbildung 6). Mithilfe einer *AOI Sequence Chart* (SensoMotoric Instruments, 2014) wurden die Fixationen innerhalb der definierten AOIs über die Zeit dargestellt.

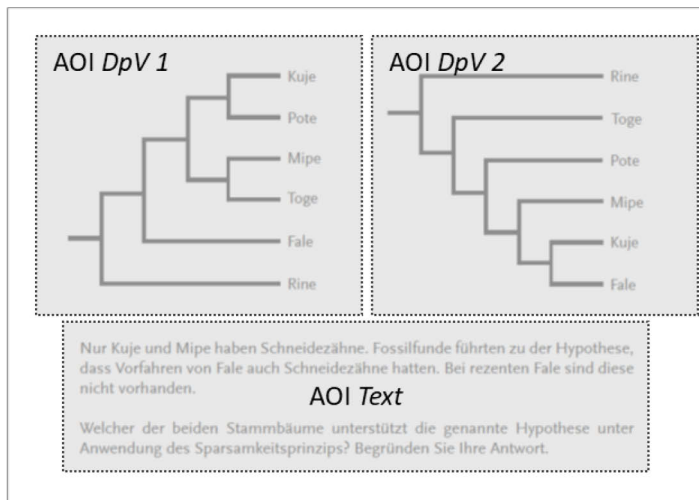


Abbildung 6: Areas of Interest (AOI) für *Alt1* medialer Modelleinsatz zur Auswertung der Fixationen über die Zeit. Graue Kästen stellen AOI dar; DpV1 = linkes DpV, DpV2 = rechtes DpV, Text = Aufgabenstamm und Aufgabenstellung.

5.8. Qualitative Auswertung der verbalen Daten durch Kodierung

Die qualitative Inhaltsanalyse kann auch als „kategoriegeleitete Textanalyse“ definiert werden, die systematisch anhand theoriebasierter Kategorien und bestimmten Regeln folgend durchgeführt wird (vgl. Mayring, 2015, S. 13). Sie stellt eine Zwischenform von quanti- und qualitativer Analyse dar. Das Material wird zwar interpretiert, die Codehäufigkeiten werden jedoch häufig quantitativ weiteranalysiert (vgl. Mayring 2015, S. 8). Die meistgenutzte Variante der Inhaltsanalyse ist die deduktive Kategorienanwendung bzw. Strukturierung. Hierzu werden zunächst theoriegeleitet Dimensionen gemäß der Fragestellungen erstellt und in Ausprägungen differenziert. Die ausdifferenzierten Dimensionen ergeben ein Kategoriensystem (vgl. Mayring, 2015, S. 97), das für das zu analysierende Material erstellt bzw. daran angepasst und am Material selbst auf seine Eignung geprüft wird. Dies erfolgt immer theoriebasiert, um eine hohe Validität zu sichern (vgl. Mayring, 2015, S. 52 f.).

5.8.1. Entwicklung der Kategorien

Die zugrundeliegenden Dimensionen des Kategoriensystems und deren Ausprägungen wurden deduktiv, also theoriebasiert festgelegt: In der Dimension „DpV-Bestandteile“ dienten die von verschiedenen Autor:innen definierten spezifischen Tree-Reading-Fähigkeiten (Tabelle 1; Baum & Smith, 2013; Blacquiére & Hoese, 2016; Halverson, 2011; Halverson & Friedrichsen, 2013; Kummer et al., 2019; Meir et al., 2007; Morabito et al., 2010; Novick & Catley, 2013, 2016; Novick, Stull, & Catley, 2012; Schramm et al., 2019) und *wissenschaftlich inadäquaten* Leseweisen (Auflistung in Kapitel 2.1.4) als Kategorien. Die Dimension „Tree-Reading-Fähigkeiten“ basiert auf der theoretisch hergeleiteten Einteilung von Tree-Reading-Fähigkeiten in *oberflächliches Ablesen* (TR_0), *wissenschaftlich adäquates* (TR_1a) bzw. *inadäquates Ablesen* (TR_1b) sowie *Ableiten* (TR_2; Kapitel 2.2). Jede Kategorie von „DpV-Bestandteile“ kann demnach auf verschiedenen Ebenen der „Tree-Reading-Fähigkeiten“ kodiert werden. Basierend auf den Charakteristika der Aufgabenstellungen wurde für Transkripte der Teilkompetenz *Alternative Modelle* die Dimension „Bezug“ aufgenommen. Diese gibt an, ob sich eine Aussage auf ein einzelnes DpV oder vergleichend auf beide DpV bezieht. Die Dimension „Aufgabenbearbeitung“ wurde ebenfalls hinzugefügt, um die Aufgabenlösungen hinsichtlich ihrer *Passung mit den wissenschaftlichen Konventionen* zu beurteilen. Die deduktiven Dimensionen „DpV-Bestandteile“ und „Aufgabenbearbeitung“ wurden induktiv anhand des Materials ergänzt (Mayring, 2015). Das Kategoriensystem wurde um den Kodierleitfaden zur Modellkompetenz von Krell und Krüger (2016) erweitert. Aus dem Kategoriensystem wurden 2 Kodierleitfäden mit Codes und Subcodes erstellt, die überwiegend identisch sind, jedoch an die 2 Teilkompetenzen *Zweck von Modellen* bzw. *Alternative Modelle* angepasst wurden (Anhang 7). Die Kodierung erfolgte aufgrund des Inhalts einer Aussage und nicht anhand verwendeter Fachbegriffe, sodass auch alltagssprachliche Aussagen als *wissenschaftlich adäquat* bzw. *den wissenschaftlichen Konventionen entsprechend* kodiert werden konnten. So sollte eine Verzerrung der Daten zum Nachteil der NiEr vermieden werden, die formell weniger erfahren im Umgang mit spezifischen Fachtermini sind als HoEr.

5.8.2. Gütekriterien

Göhner und Krell (2020) konnten Validität, Reliabilität und Objektivität als die am häufigsten genannten Gütekriterien in 50 ausgewählten qualitativen Studien identifizieren. Jedoch herrscht Uneinigkeit darüber, ob diese Gütekriterien ohne weiteres von der quantitativen auf die qualitative Forschung übertragen werden können (Hussy, Schreier, & Echterhoff, 2010) und mit welchen Maßnahmen sie erfüllt werden können (vgl. Göhner & Krell, 2020). Objektivität und Reliabilität sind hinsichtlich der Datenerhebung kaum erreichbar, da qualitative Studien meist den direkten Kontakt zu Individuen erforderlich machen und so jede Proband:in die Situation anders wahrnimmt und sich innerhalb eines Interviews oder der Bearbeitung offener Fragen weiterentwickeln kann. Bei der Datenauswertung kann Objektivität durch das Kriterium der Intersubjektivität untermauert werden, also der Übereinstimmung der Bewertung von Daten durch verschiedene Bewertende (vgl. Hussy et al., 2010, S. 266 f.). Laut Mayring (2003, 2015) ist dementsprechend ein Gütekriterium der qualitativen Inhaltsanalyse, dass verschiedene Personen die Kodierung unabhängig voneinander reproduzieren können. Die Reproduzierbarkeit kann durch die Überprüfung der Inter-coderreliabilität bewertet werden. Diese dient nicht nur zur Validierung des Kategoriensystems, sondern sollte vor allem dazu genutzt werden, Gründe für Unstimmigkeiten zu verstehen und diese formativ in die Anpassung des Kategoriensystems einzubinden (Mayring, 2015). Für die Ermittlung der Inter-coderreliabilität schulte die Versuchsleiterin eine Wissenschaftlerin aus ihrer Arbeitsgruppe im Umgang mit den Kodiermanualen. Beide Personen kodierten gleichzeitig an 2 verschiedenen Computern unabhängig voneinander dieselben zufällig ausgewählten Transkripte und verglichen ihre Kodierungen anschließend durch die Berechnung von Cohens Kappa (κ) als Maß für die Inter-coderreliabilität (Brennan und Prediger, 1981) in MAXQDA (Codeüberlappung an beiden Segmenten von mind. 90 %).

Die Grenzwerte wurden wie folgt angenommen (vgl. Kambach, 2018, S. 80):

| | |
|-----------------------|------------|
| $\kappa > 0.75$ | Sehr gut |
| $\kappa = 0.6 - 0.75$ | Gut |
| $\kappa = 0.4 - 0.6$ | Akzeptabel |
| $\kappa = 0.4$ | schlecht |

Lag κ unterhalb des akzeptablen Bereichs, wurden die betreffenden Stellen von den beiden Kodiererinnen diskutiert und das Kodiermanual angepasst. Anschließend wurde die Kodierung von beiden Kodiererinnen mit neuen Transkripten wiederholt. Es ergaben sich „gute“ bis „sehr gute“ Abschlusswerte der Intercoderreliabilität für alle Kodiermanuale (Tabelle 7).

Tabelle 7: Intercoderreliabilität für den medialen und methodischen Modelleinsatz der Teilkompetenzen Zweck von Modellen und Alternative Modelle anhand von Cohens Kappa nach Brennan und Prediger (1981).

| Teilkompetenz | κ Medial | κ Methodisch |
|---------------------|-----------------|---------------------|
| Zweck von Modellen | 0.74 | 0.62 |
| Alternative Modelle | 0.78 | 0.62 |

Im Gegensatz zu Objektivität und Reliabilität wird Validität als angemessenes Gütekriterium für die qualitative Forschung angesehen, jedoch unterschiedlich definiert (vgl. Hussy et al., 2010, S. 267). Die *Standards for Educational and Psychological Testing* (American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 2014) postulieren verschiedene Quellen für Validität, betonen jedoch, dass nicht jede Studie jedes Gütekriterium heranziehen muss. Zunächst kann in der vorliegenden Studie der *Test Content* als Quelle für Validität herangezogen werden. Dafür wurde durch ein:e HoEr überprüft, ob die Erwartungshorizonte der einzelnen Aufgaben den wissenschaftlichen Konventionen entsprechen. Diese HoEr nahm nicht als Proband:in an der vorliegenden Studie teil. Zusätzlich wurden verschiedene Kontrollvariablen erhoben, um bspw. zu überprüfen, ob HoEr und NiEr sich hinsichtlich ihrer Tree-Reading-Fähigkeiten unterscheiden und diese als weitere Quelle für Validität mit der Tree-Reading-Performanz im Versuch abzugleichen (*Relations to other variables*; AERA, APA, & NCME, 2014).

Die Validität visueller Daten kann geprüft werden, indem zu vergleichende Stimuli systematisch verändert werden und eine ausreichend große Anzahl an verschiedenen Stimuli verglichen wird, sodass der Einfluss irrelevanter Stimuli-Eigenschaften sich „herausmittelt“. Gleiche Stimuli können außerdem anhand verschiedener Personengruppen untersucht werden (Orquin & Holmqvist, 2018). Das letztgenannte Vorgehen wurde für die vorliegende Studie gewählt, da die Stichprobe klein und die verschiedenen Aufgaben inhaltlich zu unterschiedlich für einen Vergleich sind.

Wird der gleiche Forschungsgegenstand anhand verschiedener Datenquellen untersucht, kann dies ebenfalls als Quelle für Validität dienen. Hierbei ist zu beachten, dass verbale und

visuelle Daten ggf. unterschiedliche Aspekte des gleichen Forschungsgegenstandes beleuchten und so eher von einer komplementären als von einer validierenden Funktion gesprochen werden kann (vgl. Hussy et al., 2010, S. 269).

5.8.3. Codehäufigkeiten

Nach der Kodierung aller Transkripte mit MAXQDA Version 18 (VERBI Software GmbH) wurden die kodierten Segmente für jede:n Proband:in getrennt nach Teilkompetenz und Aufgabe in Excel-Tabellen exportiert. Diese Tabellen waren die Grundlage für alle durchgeführten Analysen. Für die allgemeinen Codehäufigkeiten wurden alle Sinneinheiten nach Erfahrungsgrad getrennt zusammengefasst, die prozentualen Anteile der verschiedenen Tree-Reading-Fähigkeiten an der Gesamtzahl der Codes berechnet und in Säulendiagrammen dargestellt. Die Häufigkeit der folgenden Kombinationen von Kodierungen aus 2 bzw. 3 Dimensionen wurden ermittelt und deren prozentuale Verteilung in Säulendiagrammen dargestellt:

DpV-Bestandteile x Tree-Reading-Fähigkeit (Teilkompetenz *Zweck von Modellen*) bzw. DpV-Bestandteile x Tree-Reading-Fähigkeit x Bezug (Teilkompetenz *Alternative Modelle*). Für die methodischen Aufgaben wurde noch die Dimension Aufgabenbearbeitung miteinbezogen.

5.9. Charakterisierung und Beschreibung von individuellen Lösungsprozessen

Für jedes Individuum wurde ermittelt, welche Tree-Reading-Fähigkeit in einer Aufgabe maximal codiert wurde. Anders als bei den prozentualen Anteilen spielte die Häufigkeit der maximal erreichten Tree-Reading-Fähigkeit oder anderer Tree-Reading-Fähigkeiten keine Rolle. Außerdem wurde klassifiziert, inwiefern die Lösungen im methodischen Modelleinsatz den wissenschaftlichen Konventionen entsprechen: *den wissenschaftlichen Konventionen entsprechend, von den wissenschaftlichen Konventionen abweichend, ohne Lösung* oder *nicht kodierbar*. Anhand der maximal kodierten Tree-Reading-Fähigkeit und der kodierten Entsprechung wissenschaftlicher Konventionen wurde eine Tabelle für den interindividuellen Vergleich von Tree-Reading-Prozessen erstellt und nach Anzahl der von den *wissenschaftlichen Konventionen abweichenden* Lösungen sortiert, getrennt nach Aufgabe und Modelleinsatz. Daraus ergaben sich 4 Gruppen hinsichtlich der Abweichungen: *ohne* = keine Abweichung von wissenschaftlichen Konventionen, *medial* = in mindestens einer

medialen Aufgabe, *methodisch* = in mindestens einer methodischen Aufgabe, *medial + methodisch* = in mindestens einer medialen und einer methodischen Aufgabe.

Um individuelle Lösungsprozesse zu erläutern und Abweichungen von den wissenschaftlichen Konventionen zu veranschaulichen, wurden aus jeder der 4 Gruppen individuelle Lösungsprozesse ausgewählt, die möglichst repräsentativ für die jeweilige Gruppe sind. Während in den vorausgegangenen Schritten Segmente nur hinsichtlich ihrer Codierung betrachtet wurden, wurden nun alle codierten Segmente eines Individuums miteinander in Bezug gesetzt und so Tree-Reading-Prozesse in ihrer Gesamtheit beschrieben. Für jedes codierte Segment wurde die Kombination von DpV-Bestandteil x Tree-Reading-Fähigkeit (Teilkompetenz *Zweck von Modellen*) bzw. DpV-Bestandteil x Tree-Reading-Fähigkeit x Bezug (Teilkompetenz *Alternative Modelle*) angegeben, um das Zusammenspiel der einzelnen Codierungen zu illustrieren.

5.10. Qualitative Auswertung der visuellen Daten und Triangulation

Obwohl die DpV für die Aufgaben anhand der in Kapitel 5.4.3 genannten Kriterien entworfen wurden, stellte sich bei der Analyse der Blickdaten heraus, dass die Äste und Knotenpunkte der DpV zu dicht aneinander liegen, um valide AOs zur quantitativen Auswertung anzulegen. Daher wurde eine qualitative Auswertung der Blickdaten bevorzugt (Viiri, Kekule, Isoniemi, & Hautala, 2017) und auf die Teilkompetenz *Alternative Modelle* angewendet, da diese als Stimuli 2 örtlich deutlich voneinander abgegrenzte DpV beinhaltet. Es wurden exemplarisch *Sequence Charts* für 1 HoEr und 2 NiEr erstellt und anhand des zeitlichen Verlaufs mit den verbalen Kodierungen der Dimension *Bezug* trianguliert, um zu untersuchen, inwiefern verbale Äußerungen mit visuellen Fixationen übereinstimmen. Diese qualitative Auswertung ist neben den in Kapitel 5.8.2 genannten Gütekriterien eine weitere Validitätsquelle, da die visuellen hier die verbalen Daten stützen können.

6. Ergebnisse

6.1. Induktive Erweiterung des Kategoriensystems

Das Kategoriensystem wurde um folgende Codes induktiv erweitert, um *wissenschaftlich inadäquates Ablesen* (TR_1b) zu charakterisieren, das über die in Kapitel 2.1.4 genannten *wissenschaftlich inadäquaten* Leseweisen hinausgeht:

1. *Verlust bei Nichtgebrauch und Entstehung bei Gebrauch*: Merkmale gehen verloren, wenn sie nicht mehr gebraucht werden bzw. entstehen, wenn sie gebraucht werden.
2. *Vererbung auf Individualebene*: Vererbung von Merkmalen auf Individualebene statt auf Populationsebene, analog zu Familienstammbaum.
3. *Genetische Prädisposition*: Ein Merkmal ist bereits im Genotyp vorhanden und äußert sich dann in Phänotyp.
4. *Abzweigungsreihenfolge*: DpV, bei denen von älteren Knotenpunkten Entwicklungslinien mit nur einem rezenten Taxon ausgehen, werden als sparsamer betrachtet als Entwicklungslinien mit Schwestertaxa.
5. *Nur MRCA für ein Taxon*: Ein Knotenpunkt ist nur MRCA für ein abgehendes Taxon, nicht von der durch ihn begründeten Klade.
6. *Kein Knotenpunkt = keine weiteren Vorfahren*: Wenn entlang einer Entwicklungslinie kein weiterer Knotenpunkt dargestellt wird, gab es keine weiteren MRCA.
7. *Knotenpunkte haben unterschiedliche Bedeutungen*: Knotenpunkte werden nicht immer als MRCA betrachtet, sondern als „Zwischenvorfahr“, der als weniger oder nicht relevant beschrieben wird.
8. *Gleiche Position = gleicher MRCA*: Befinden sich Knotenpunkte bei 2 DpV oberflächlich auf gleicher Höhe bzw. an der gleichen Position von der Wurzel aus gesehen, werden sie als identischer MRCA beschrieben.
9. *Abspaltungen/ Vorfahren zählen*: Als Vergleichskriterium von 2 DpV wird von der Wurzel an abgezählt, an welchem Knotenpunkt sich die Entwicklungslinie eines bestimmten Taxons abspaltet.

6.2. Pilotierung *BETTSI*

Zunächst wurden den 11 Items des vorläufigen *BETTSI* die entsprechenden Tree-Reading-Fähigkeiten (Kapitel 2.2) bzw. *wissenschaftlich inadäquaten* Leseweisen zugeordnet (Kapitel 2.1.4; Tabelle 8). Von 29 zur Pilotierung erhobenen *BETTSI*-Fragebögen konnten 26 ausgewertet werden. Die Reliabilitätsanalyse ergab ein ungenügendes Cronbachs α von 0,553 (gut: $> 0,8$; Bortz & Döring, 2006). Die Trennschärfe (Korrigierte Item-Skala-Korrelation) variierte stark zwischen den Items und war teilweise negativ. Nur Item 4, 9 und 10 wiesen eine zufriedenstellende Trennschärfe $> 0,3$ auf (Field, 2009, S. 679; Tabelle 8).

Die niedrige Reliabilität ist auf die oben genannten sehr unterschiedlichen Item-Inhalte zurückzuführen, die nicht durchgehend die gleichen Tree-Reading-Fähigkeiten testen. Daher wird *BETTSI* in der vorliegenden Arbeit nur als Ergänzung zu den qualitativen Daten genutzt.

Tabelle 8: Zuordnung der 11 BETTSI-Items zu Tree-Reading-Fähigkeiten sowie Item-Skala-Korrelation und Cronbachs α , wenn Item weggelassen wird (w. l. w.). Skalenmittelwert und Skalenvarianz w. l. w. sind ebenfalls angegeben.

| Item | Zugeordnete Tree-Reading-Fähigkeit | Skalenmittelwert w. l. w. | Skalenvarianz w. l. w. | Korrigierte Item-Skala-Korrelation | Cronbachs Alpha, w. l. w. |
|------|---|------------------------------|---------------------------|------------------------------------|------------------------------|
| 1 | Verwandtschaftsgrad | 5.88 | 3.066 | 0.287 | 0.514 |
| 2 | Höher und niedriger entwickelte Taxa (<i>wissenschaftlich inadäquat</i>) | 6.12 | 3.226 | 0.275 | 0.518 |
| 3 | Merkmale aus DpV ablesen | 5.50 | 4.020 | - 0.217 | 0.622 |
| 4 | Verwandtschaftsgrad | 5.88 | 2.826 | 0.439 | 0.463 |
| 5 | Bestimmung <i>MRCA</i> | 5.54 | 3.298 | 0.257 | 0.523 |
| 6 | Entwicklungslinie einzelner Taxa | 5.58 | 3.374 | 0.175 | 0.544 |
| 7 | Verwandtschaftsgrad | 6.27 | 3.565 | 0.192 | 0.540 |
| 8 | Taxa stammen voneinander ab (<i>wissenschaftlich inadäquat</i>) | 5.50 | 3.540 | 0.116 | 0.555 |
| 9 | Rotation an Knotenpunkten | 5.77 | 2,745 | 0.501 | 0.442 |
| 10 | Unterschiedliche Repräsentationsform | 6.04 | 2.838 | 0.489 | 0.452 |
| 11 | Unterschiedliche Repräsentationsform | 5.38 | 3.926 | - 0.166 | 0.583 |

6.3. Stichprobenbeschreibung

Insgesamt wurden Daten von 10 HoEr (Alter 35 ± 5 Jahre) und 15 NiEr (Alter 28 ± 3 Jahre) erhoben. Alle HoEr hatten einen fachbiologischen Hintergrund, teilweise in Kombination mit Geografie, Geologie und Paläoontologie. Als höchsten akademischen Grad gaben 7 HoEr eine abgeschlossene Promotion an, jeweils 1 HoEr Diplom bzw. Master. 1 HoEr machte keine Angaben zu Fächern und Abschlüssen. Die 3 HoEr ohne abgeschlossene Promotion bzw. ohne Angabe zum akademischen Grad gaben an, Promotionsstudierende zu sein.

Alle NiEr waren im *Master of Education* mit Biologie im Kern- oder Zweitfach eingeschrieben. 9 NiEr gaben Biologie und Chemie als Fächerkombination an, 6 Biologie in Kombination mit einem Fach außerhalb des MINT-Bereichs. 1 NiEr gab an, bereits beruflich tätig zu sein und ein Biologiestudium mit Diplom abgeschlossen zu haben, jedoch ohne Schwerpunkt auf DpV. 14 NiEr gaben an, eine Vorlesung mit Phylogenie-Anteil besucht zu haben, 1 NiEr machte keine Angaben.

Den ersten Kontakt mit DpV hatten nach eigener Aussage 3 HoEr in der Schule, 5 im Studium und 2 während der Promotion. 3 HoEr gaben an, Tree-Reading in der Schule gelernt zu haben, 7 im Studium. 11 NiEr befassten sich das erste Mal in der Schule mit DpV, 4 im Studium. 8 NiEr gaben an, Tree-Reading in der Schule erlernt zu haben, 7 im Studium.²⁰

HoEr schätzten die Aussage „*Ich habe schon oft mit evolutiven Stammbäumen gearbeitet*“ und „*Ich kann einen evolutiven Stammbaum wissenschaftlich korrekt lesen.*“ überwiegend mit „*Trifft zu*“ ein. NiEr schätzten sich bei der ersten Aussage niedriger ein („*Trifft eher nicht zu*“ und „*Trifft eher zu*“) als bei der zweiten („*Trifft eher zu*“) und insgesamt signifikant niedriger als HoEr (Tabelle 9). Im *BETTSI*-Test zu Tree-Reading-Fähigkeiten schnitten HoEr signifikant besser ab als NiEr (Tabelle 9). HoEr schnitten im 3D-Würfelttest zum räumlichen Denken signifikant besser ab als NiEr, der Mittelwert beider Gruppen befand sich dabei im „*durchschnittlichen*“ Bereich (26 - 74 %, Gittler, 1990, S. 45). Jedoch lagen bei beiden Erfahrungsgruppen starke Standardabweichungen von je über 20 % vor, sodass einzelne HoEr auch „*überdurchschnittlich*“ abschnitten (75 - 94 %; Gittler, 1990, S. 45). Die mittlere Bearbeitungszeit des ZVT unterschied sich nicht zwischen HoEr und NiEr (Tabelle 9).

²⁰ Es wurde erhoben, in welchen Kontexten es zu Kontakt mit DpV kam und das Lesen von DpV erlernt wurde. Hierbei war eine Mehrfachnennung möglich. Zur besseren Verständlichkeit wird hier die erste Bildungsumgebung genannt, in der laut der Proband:innen Kontakt zu DpV bestand.

Tabelle 9: Kontrollvariablen „Häufigkeit Arbeit mit DpV“ und „wissenschaftlich adäquates Lesen von DpV“ (Selbsteinschätzung), Tree-Reading-Fähigkeiten (BETTSI), räumliches Denken (3DW) und kognitive Bearbeitungsgeschwindigkeit (ZVT) für HoEr und NiEr. Mittelwert und Standardabweichung; T-Test ungepaart, zweiseitig, n. s. = nicht signifikant; Effektstärke Cohens $d > 0,2$ = schwacher Effekt, $d > 0,5$ = mittlerer Effekt; $d > 0,8$ = starker Effekt.

| Kontrollvariable | HoEr (n = 10) | NiEr (n = 15) | Signifikanz-niveau | Effektstärke Cohens d |
|---|---------------|---------------|--------------------|-----------------------|
| Selbsteinschätzung Häufigkeit Arbeit mit DpV (4 = trifft zu) | 3,8 ± 0,4 | 2,5 ± 0,6 | $p < 0,001$ | 2,56 (stark) |
| Selbsteinschätzung wissenschaftlich adäquates Lesen von DpV (4 = trifft zu) | 3,8 ± 0,4 | 3,0 ± 0,5 | $p < 0,001$ | 1,62 (stark) |
| BETTSI Anteil korrekte Antworten (%) | 91,8 ± 10,3 | 69,7 ± 10,3 | $p < 0,001$ | 2,05 (stark) |
| 3DW Anteil korrekte Antworten (%) | 67,5 ± 24,3 | 42,2 ± 24,8 | $p < 0,05$ | 0,99 (stark) |
| ZVT Mittlere Bearbeitungszeit (sec) | 56,4 ± 10 | 57,3 ± 8,9 | n. s. | - |

Hinsichtlich der Modellkompetenz unterscheidet sich die Niveauzuordnung von HoEr und NiEr jeweils zwischen den Teilkompetenzen. Die Antworten von HoEr wurden bei den Teilkompetenzen *Eigenschaften von Modellen* und *Ändern von Modellen* überwiegend Niveau II zugeordnet (je 90,0%), bei der Teilkompetenz *Alternative Modelle* überwiegend Niveau III (80,0%). Bei der Teilkompetenz *Zweck von Modellen* wurden die Antworten zu ähnlichen Teilen in Niveau I (40,0%) bzw. Niveau III (50,0%) eingeordnet, bei *Testen von Modellen* sowohl Niveau III (50%) als auch der didaktischen Perspektive (40,0%; Tabelle 10). Bei NiEr wurden die Antworten bei den Teilkompetenzen *Eigenschaften von Modellen* (71,4%), *Alternative Modelle* (73,3%) und *Ändern von Modellen* (80,0%) überwiegend Niveau II zugeordnet. Bei der Teilkompetenz *Zweck von Modellen* wurden die meisten Antworten in Niveau I eingeordnet (46,7%), jeweils gleich große Anteile jedoch auch Niveau II und III (je 26,7%). Bei der Teilkompetenz *Testen von Modellen* wurden die meisten NiEr der didaktischen Perspektive zugeordnet (53,3%), ein weiterer großer Anteil Niveau II (33,3%; Tabelle 10). Gemeinsam ist beiden Gruppen der große Anteil von Proband:innen, deren Antworten bei der Teilkompetenz *Testen von Modellen* der didaktischen Perspektive zugeordnet wurden.

Signifikante Unterschiede zwischen HoEr und NiEr in der Verteilung auf die verschiedenen Niveaus bestehen in den Teilkompetenzen *Eigenschaften von Modellen* und *Alternative Modelle* (Tabelle 10).

Tabelle 10: Kontrollvariable Modellkompetenz je Erfahrungsgrad. Prozentuale Anteile der Niveaus I | II | III | ggf. [didaktische Perspektive] pro Teilkompetenz. Bei der didaktischen Perspektive werden Modelle ausschließlich hinsichtlich ihres Einsatzes im Biologieunterricht bewertet. Die größten Anteile sind fett hervorgehoben. T-Test ungepaart, zweiseitig, n. s. = nicht signifikant; Effektstärke Cohens $d > 0,2$ = schwacher Effekt, $d > 0,5$ = mittlerer Effekt; $d > 0,8$ = starker Effekt.

| Teilkompetenz | HoEr (n = 10) | NiEr (n = 15) | Signifikanzniveau | Cohens d | Effektstärke r |
|----------------------------|---|--------------------------------------|-------------------|-----------------|----------------|
| <i>Eigenschaften v. M.</i> | 0 90,0 10,0 | 28,6 71,4 0 (n = 14) | p < 0,05 | 0,97 (stark) | 0,436 |
| <i>Alternative Modelle</i> | 0 20,0 80,0 | 6,7 73,3 13,3 [6,7] | p < 0,05 | 1,55 (stark) | 0,612 |
| <i>Zweck v. M.</i> | 40,0 10,0 50,0 | 46,7 26,7 26,7 | n. s. | - | - |
| <i>Testen v. M.</i> | 0 10,0 50,0 [40,0] | 0 33,3 13,3 [53,3] | n. s. | - | - |
| <i>Ändern v. M.</i> | 0 90,0 10,0 | 20,0 80,0 0 | n. s. | - | - |

Pro Proband:in wurden 8 Transkripte kodiert, pro Teilkompetenz je 2 pro Modelleinsatz. Von den insgesamt 200 Transkripten wurden einzelne aus der Analyse ausgeschlossen, wenn sie aus sprachlichen oder inhaltlichen Gründen nicht kodierbar oder auswertbar waren (Tabelle 11).

Tabelle 11: Von der Analyse ausgeschlossene Transkripte je Aufgabe und Erfahrungsgrad. „Zweck1“ bezeichnet Aufgabe 1 der Teilkompetenz Zweck von Modellen, analog ergeben sich die anderen Bezeichnungen der Aufgaben.

| Teilkompetenz | <i>Zweck von Modellen</i> | | | | <i>Alternative Modelle</i> | | | |
|---------------|---------------------------|--------|-------------------------------|----------|----------------------------|------|-------------|------|
| | Medial | | Methodisch | | Medial | | Methodisch | |
| Aufgabe | Zweck1 | Zweck2 | Zweck1 | Zweck2 | Alt1 | Alt2 | Alt1 | Alt2 |
| HoEr | - | - | - | - | - | - | - | - |
| NiEr | P09 | P09 | P11, P12, P15, P20, P22 | P08, P21 | - | - | P08, P11 | - |

Im medialen Modelleinsatz wurden für die Teilkompetenz *Zweck von Modellen* bei beiden Erfahrungsgruppen insgesamt deutlich weniger Sinneinheiten kodiert als für die Teilkompetenz *Alternative Modelle*: Bei HoEr für *Zweck von Modellen* 29 bzw. 30 Sinneinheiten, bei *Alternative Modelle* 107 bzw. 116 Sinneinheiten. Bei NiEr bei *Zweck von Modellen* 54 bzw. 39 Sinneinheiten, bei *Alternative Modelle* 173 bzw. 148 Sinneinheiten. Bis auf je 1 NiEr bei der Teilkompetenz *Zweck von Modellen* konnten alle Proband:innen in die Kodierung eingeschlossen werden. Beim methodischen Modelleinsatz variiert die Anzahl der Sinneinheiten je nach Aufgabe. Bei HoEr wurden bei Aufgabe *Alt2* 56 Sinneinheiten kodiert, bei den anderen Aufgaben je mehr als 100 (*Zweck*: 120 bzw. 134; *Alt1*: 106). Bei NiEr wurden bei der Teilkompetenz *Zweck von Modellen* 64 bzw. 94 und bei den Aufgaben der Teilkompetenz *Alternative Modelle* 94 bzw. 71 Sinneinheiten kodiert. Bei NiEr wurden bei Aufgabe *Zweck1* 5 Proband:innen von der Analyse ausgeschlossen, bei Aufgabe *Zweck2* und Aufgabe *Alt1* je 2 (Tabelle 12).

Tabelle 12: Anzahl Sinneinheiten und Proband:innen (n) je Aufgabe und Erfahrungsgrad. „Zweck1“ bezeichnet die Aufgabe 1 der Teilkompetenz *Zweck von Modellen*, analog ergeben sich die anderen Bezeichnungen der Aufgaben.

| Teilkompetenz | Zweck von Modellen | | | | Alternative Modelle | | | |
|---------------|--------------------|---------------|---------------|---------------|---------------------|-------------|-------------|-------------|
| | Medial | | Methodisch | | Medial | | Methodisch | |
| Aufgabe | <i>Zweck1</i> | <i>Zweck2</i> | <i>Zweck1</i> | <i>Zweck2</i> | <i>Alt1</i> | <i>Alt2</i> | <i>Alt1</i> | <i>Alt2</i> |
| HoEr | 29 | 30 | 120 | 134 | 107 | 116 | 106 | 56 |
| | (n = 10) | (n = 10) | (n = 10) | (n = 10) | (n = 10) | (n = 10) | (n = 10) | (n = 10) |
| NiEr | 54 | 39 | 64 | 94 | 173 | 148 | 94 | 71 |
| | (n = 14) | (n = 14) | (n = 10) | (n = 13) | (n = 15) | (n = 15) | (n = 13) | (n = 15) |

6.4. Intercoderreliabilität

Als Maß für die Intercoderreliabilität wurde Cohens Kappa (Brennan & Prediger, 1981) ermittelt. Hierfür wurde pro Aufgabe je ein Transkript von 3 Proband:innen herangezogen, sodass pro Modelleinsatz und Teilkompetenz 6 Transkripte von 3 Proband:innen einfließen. Ausgehend von einer Codeüberlappung von 90 % an beiden Segmenten kann die Intercoderübereinstimmung für den medialen Modelleinsatz der Teilkompetenz *Zweck von Modellen* als gut ($\kappa = 0.75$) und *Alternative Modelle* als sehr gut ($\kappa = 0.78$) eingeschätzt werden. Für den methodischen Modelleinsatz konnte eine gute Übereinstimmung für beide Teilkompetenzen erreicht werden (jeweils $\kappa = 0.62$; Tabelle 13).

Tabelle 13: Intercoderübereinstimmung je Modelleinsatz. Cohens Kappa (nach Brennan & Prediger, 1981) $\kappa > 0.75$: sehr gut, $\kappa = 0.6 - 0.75$: gut, $\kappa = 0.4 - 0.6$: akzeptabel, $\kappa < 0.4$: schlecht. Codeüberlappung an beiden Segmenten von mindestens 90 %. „Zweck1“ bezeichnet Aufgabe 1 der Teilkompetenz *Zweck von Modellen*, analog ergeben sich die anderen Bezeichnungen der anderen Aufgaben.

| Teilkompetenz | <i>Zweck von Modellen</i> | | | | <i>Alternative Modelle</i> | | | |
|------------------------------------|------------------------------|--------|------------------------------|--------|------------------------------|------|------------------------------|------|
| | Medial | | Methodisch | | Medial | | Methodisch | |
| Aufgabe | Zweck1 | Zweck2 | Zweck1 | Zweck2 | Alt1 | Alt2 | Alt1 | Alt2 |
| Cohens Kappa (Entsprechung) | 0.75 (gut) | | 0.62 (gut) | | 0.78 (sehr gut) | | 0.62 (gut) | |
| Transkripte | HoEr P14, NiEr P13, NiEr P27 | | HoEr P14, NiEr P16, NiEr P18 | | HoEr P05, NiEr P18, NiEr P22 | | HoEr P05, NiEr P18, NiEr P22 | |

6.5. Allgemeine Codehäufigkeiten im medialen Modelleinsatz

Für die Analyse der allgemeinen Tree-Reading-Fähigkeiten im medialen Modelleinsatz wurden pro Teilkompetenz und Aufgabe alle Sinneinheiten einer Erfahrungsgruppe zusammengefasst und die Anteile der Tree-Reading-Fähigkeiten an der Gesamtzahl der Sinneinheiten berechnet. Im Folgenden wird die Abkürzung *Zweck* genutzt, wenn prozentuale Anteile beider Aufgaben der Teilkompetenz *Zweck von Modellen* angegeben werden, die Abkürzung *Alt* für beide Aufgaben der Teilkompetenz *Alternative Modelle*. Die erste Prozentzahl entspricht der Aufgabe 1 der jeweiligen Teilkompetenz, die zweite Prozentzahl der zweiten. Werden prozentuale Anteile einzelner Aufgaben angegeben, wird dies explizit kenntlich gemacht: *Zweck1* entspricht bspw. der Aufgabe 1 der Teilkompetenz *Zweck von Modellen* im medialen oder methodischen Modelleinsatz. Dies kommt vor allem in den Säulendiagrammen vor.

Bei HoEr überwog *wissenschaftlich adäquates Ablesen* (TR_1a) sowohl bei der Teilkompetenz *Zweck von Modellen* (*Zweck*: 79,3 % bzw. 93,3 %; Abbildung 7a) als auch bei der Teilkompetenz *Alternative Modelle* (*Alt*: 75,7 % bzw. 74,1 %; Abbildung 7b). Die restlichen Anteile entfielen überwiegend auf *oberflächliches Ablesen* (TR_0; *Zweck*: 20,7 % bzw. 3,3 %; *Alt*: 22,4 % bzw. 25,0 %) und bei 3 Aufgaben zu sehr geringem Anteil auf *wissenschaftlich inadäquates Ablesen* (TR_1b; *Zweck*: 0 bzw. 3,3 %; *Alt*: 1,9 % bzw. 0,9 %).

NiEr *lasen* unabhängig von der Teilkompetenz je nach Aufgabe teilweise weniger als in der Hälfte der Aussagen *inhaltlich wissenschaftlich adäquat ab* (TR_1a; *Zweck*: 37,0 % bzw. 48,7 %; Abbildung 7a; *Alt*: 36,4 % bzw. 59,5 %; Abbildung 7b). Auf *oberflächliches Ablesen* (TR_0) entfielen je nach Aufgabe ca. ein Fünftel bis zur Hälfte der Aussagen mit größerem Anteil bei der Teilkompetenz *Zweck von Modellen* (*Zweck*: 48,2 % bzw. 33,3 %; *Alt*: 21,4 % bzw. 29,1 %). Bei allen Aufgaben kam es zu *wissenschaftlich inadäquatem Ablesen* (TR_1b): Bei der Teilkompetenz *Zweck von Modellen* zu ähnlichen Anteilen in den Aufgaben (*Zweck*: 14,8 % bzw. 17,95 %), bei der Teilkompetenz *Alternativen Modellen* unterschied sich der Anteil um das Vierfache zwischen den Aufgaben (*Alt*: 41 % bzw. 11,5 %). Bei Aufgabe *Alt1* kam bei NiEr zu 1,2 % *Ableiten* (TR_2) vor.

Insgesamt unterscheiden sich die Anteile der verschiedenen Tree-Reading-Fähigkeiten deutlich zwischen den beiden Erfahrungsgruppen. Außerdem unterscheiden sie sich innerhalb der Gruppe NiEr ebenfalls zwischen den Teilkompetenzen und bei der Teilkompetenz *Alternative Modelle* wiederum zwischen den Aufgaben.

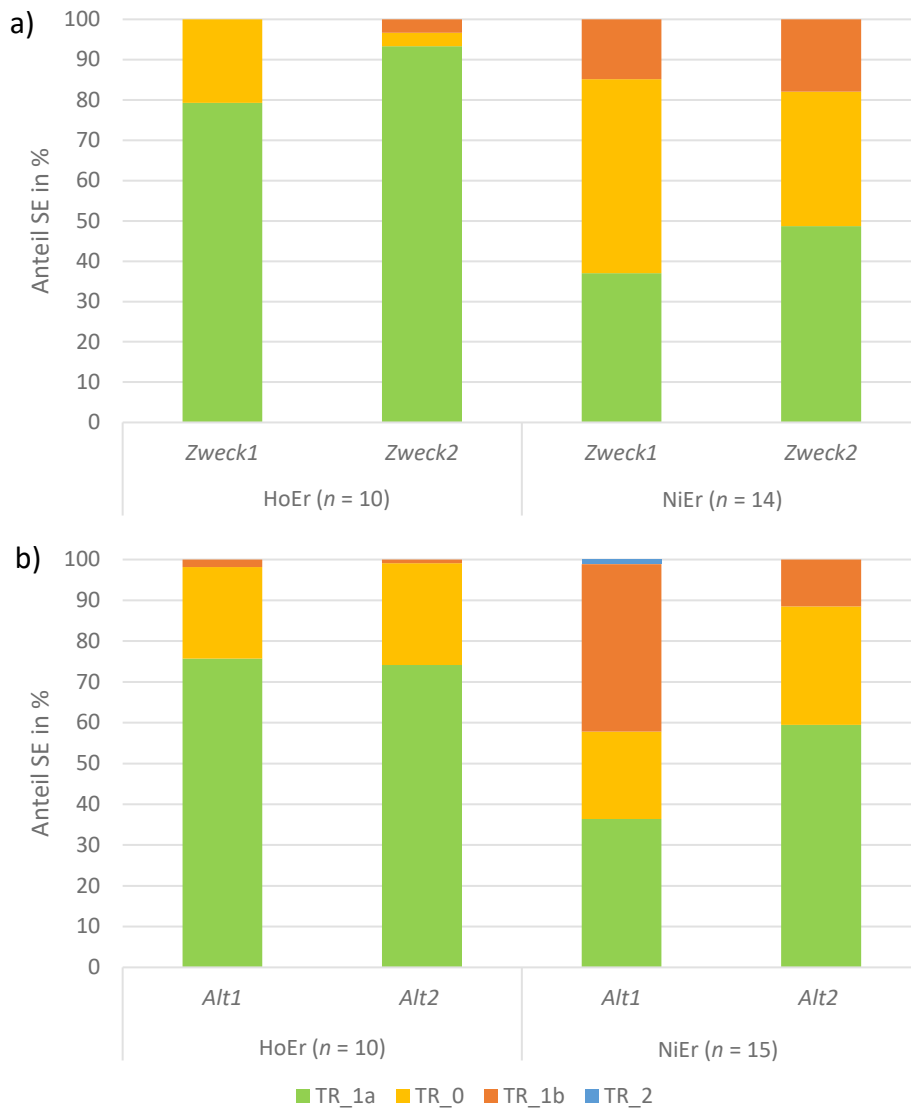


Abbildung 7: Allgemeine Codehäufigkeiten der Tree-Reading-Fähigkeiten in Prozent im medialen Modelleinsatz der Teilkompetenz Zweck von Modellen (a) und Alternative Modelle (b) für HoEr und NiEr. Anzahl kodierter Sinneinheiten bei Zweck von Modellen: HoEr Zweck1: 29, Zweck2: 30, Alt1: 107, Alt2: 116; NiEr Zweck1: 54; Zweck2: 39, Alt1: 173, Alt2: 148. Die Anzahl der analysierten Proband:innen (n) ist in Klammern angegeben. TR_1a = wissenschaftlich adäquates Ablesen, TR_0 = oberflächliches Ablesen, TR_1b = wissenschaftlich inadäquates Ablesen, TR_2 = Ableiten.

Betrachtet man die kodierten DpV-Bestandteile, bezogen sich bei der **Teilkompetenz Zweck von Modellen** (Abbildung 8) beide Erfahrungsgruppen zum Großteil auf MRCA (HoEr Zweck: 75,9 % bzw. 73,3 %; NiEr Zweck: 85,1 % bzw. 84,6 %). HoEr (Abbildung 8a) zeigten dabei überwiegend *wissenschaftlich adäquates Ablesen* (TR_1a; Zweck: 55,2 % bzw. 66,7 %), zu einem geringeren und zwischen den Aufgaben unterschiedlichen Anteil *oberflächliches Ablesen* (TR_0; Zweck: 20,7 % bzw. 3,3 %) und bei Aufgabe Zweck2 einen kleinen Anteil *wissenschaftlich inadäquates Ablesen* (TR_1b; Zweck2: 3,3 %). Bei NiEr (Abbildung 8b) lag der Anteil von *wissenschaftlich adäquatem Ablesen* (TR_1a) bei MRCA je nach Aufgabe bei 33,3 % bzw. 43,6 %. *Oberflächliches Ablesen* (TR_0) nutzten sie bei MRCA zu 37,0 % bzw. 23,1 % und

wissenschaftlich inadäquates Ablesen (TR_1b) kam mit ähnlichen Anteilen vor (Zweck: 14,8 % bzw. 17,9 %). Außerdem wurden von beiden Gruppen Klade genannt, jedoch zu einem deutlich geringeren Anteil (HoEr Zweck: 24,1 % bzw. 26,7 %; NiEr Zweck: 14,8 % bzw. 15,4 %). HoEr lesen hier ausschließlich *wissenschaftlich adäquat ab* (TR_1a), NiEr lasen Klade überwiegend *oberflächlich ab* (TR_0; Zweck: 11,1 % bzw. 10,3 %).

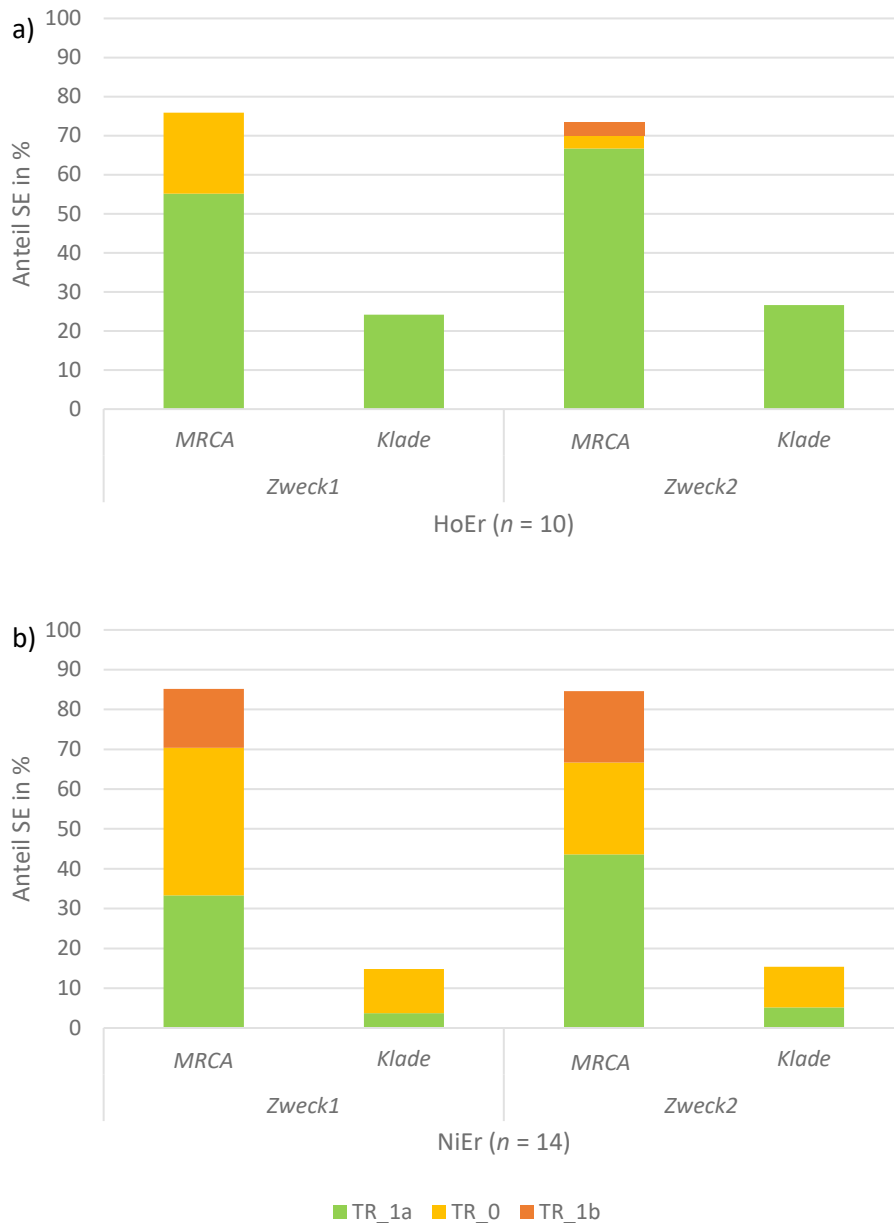


Abbildung 8: Zuordnung der DpV-Bestandteile zu Tree-Reading-Fähigkeiten in Prozent der Teilkompetenz Zweck von Modellen im medialen Modelleinsatz für HoEr und NiEr. Nur DpV-Bestandteile > 5 %. Anzahl kodierter Sinneinheiten HoEr Zweck1: 29, Zweck2: 30; NiEr Zweck1: 54; Zweck2: 39. Die Anzahl der analysierten Proband:innen (n) ist in Klammern angegeben. TR_1a = wissenschaftlich adäquates Ablesen, TR_0 = oberflächliches Ablesen, TR_1b = wissenschaftlich inadäquates Ablesen.

Bei der **Teilkompetenz *Alternative Modelle*** (Abbildung 9) wurden insgesamt mehr DpV-Bestandteile genannt als bei der Teilkompetenz *Zweck von Modellen*. HoEr (Abbildung 9a) beschrieben bei beiden Aufgaben mit je 70 % bzw. 73,3 % am häufigsten Kladen, überwiegend durch *wissenschaftlich adäquates Ablesen* (TR_1a; Alt: 58 % bzw. 54,3 %) und zu einem geringeren Teil durch *oberflächliches Ablesen* (TR_0; Alt: 12,1 bzw. 19 %). Bei *Alt1* nannten sie zusätzlich mit 21,5 % die Topologie zu etwa gleichen Teilen durch *wissenschaftlich adäquates Ablesen* (TR_1a; Alt1: 11,4 %) und *oberflächliches Ablesen* (TR_0; Alt1: 10,3 %). Bei Aufgabe *Alt2* lasen sie *wissenschaftlich adäquat* Verwandtschaftsbeziehungen ab (TR_1a; Alt2: 12,9 %) bzw. beschrieben Topologien zu etwa gleichen Teilen durch *wissenschaftlich adäquates Ablesen* (TR_1a; Alt1: 6,9 %) und *oberflächliches Ablesen* (TR_0; Alt2: 6 %).

NiEr (Abbildung 9b) beschrieben ebenfalls hauptsächlich Kladen, der Anteil variierte jedoch zwischen den Aufgaben stark mit insgesamt 26,6 % bei *Alt1* und 64,9 % bei *Alt2*. Hierbei nutzten NiEr jeweils überwiegend *wissenschaftlich adäquates Ablesen* (TR_1a; Alt: 17,9 % bzw. 44,6 %) und weniger *oberflächliches Ablesen* (TR_0; Alt: 8,7 % bzw. 20,3 %). NiEr nannten zusätzlich Verwandtschaftsbeziehungen (Alt: 8,1 % bzw. 6,8 %) ausschließlich mittels *wissenschaftlich adäquatem Ablesen* (TR_1a). Die Topologie wurde zwischen den Aufgaben unterschiedlich häufig beschrieben (Alt: 15,7 % bzw. 6,1 %), ebenso wie *MRCA* (Alt: 19,1% bzw. 8,8 %). Bei *Alt1* wurden außerdem zu 15 % *Taxa* genannt. Im Gegensatz zu den anderen DpV-Bestandteilen wurden *Taxa* (9,8 %) und *MRCA* (15,6 % bzw. 7,4 %) von NiEr zum Großteil *wissenschaftlich inadäquat abgelesen* (TR_1b).

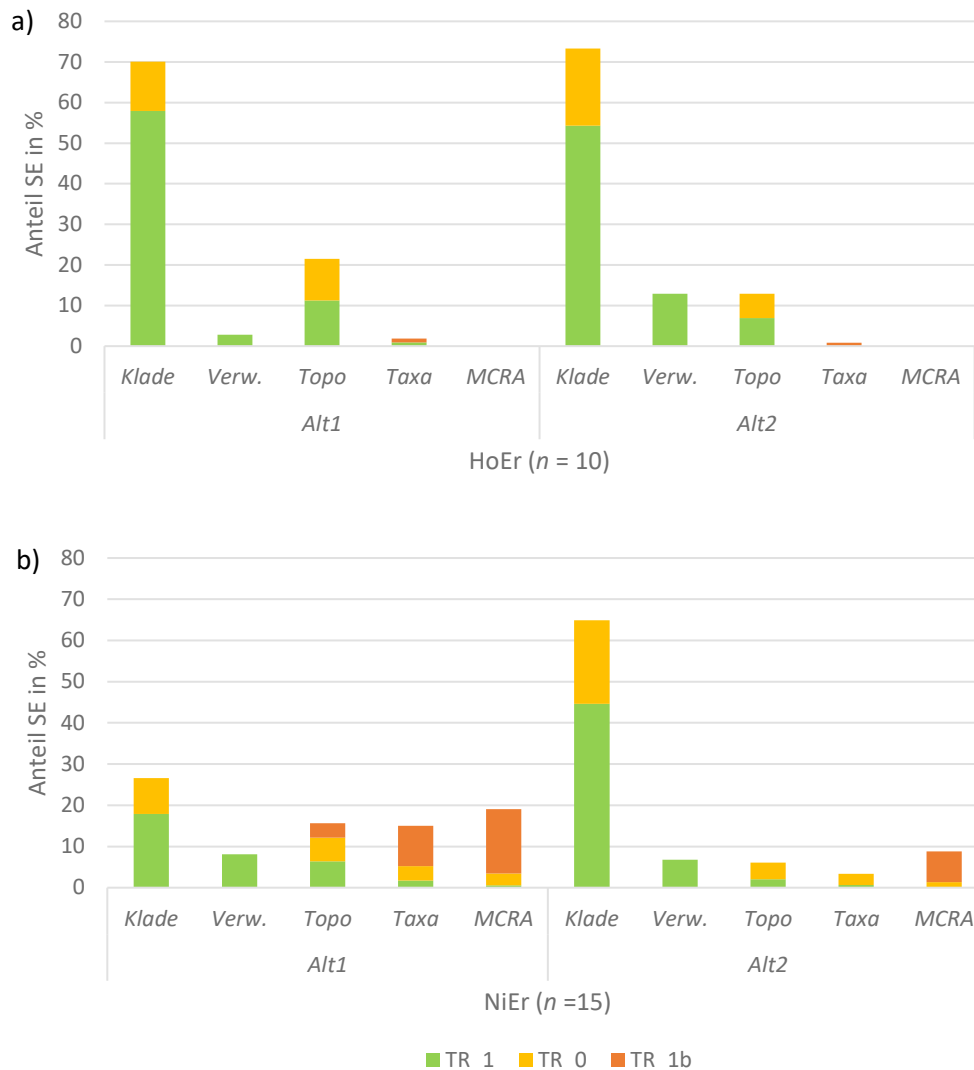


Abbildung 9: Zuordnung der DpV-Bestandteile zu den Tree-Reading-Fähigkeiten in Prozent der Teilkompetenz Alternative Modelle im medialen Modelleinsatz. Nur DpV-Bestandteile > 5%. HoEr Alt1: 107, Alt2: 116; NiEr Alt1: 173, Alt2: 148. Die Anzahl der analysierten Proband:innen (n) ist in Klammern angegeben. TR_2 = Ableiten, TR_1a = wissenschaftlich adäquates Ablesen, TR_0 = oberflächliches Ablesen, TR_1b = wissenschaftlich inadäquates Ablesen. DpV-Bestandteile: Verw = Verwandtschaft, Topo = Topografie.

Zusätzlich zum konkreten Bezug wurde kodiert, ob sich eine Aussage nur auf ein einzelnes DpV bezieht (einzeln) oder ein direkter Vergleich zum anderen DpV besteht (Vergleich). Unter Vergleich fällt ebenfalls, wenn gleichzeitig über beide DpV gesprochen wurde.

Bei der Beschreibung von DpV bezogen sich HoEr in Aufgabe *Alt1* zu 79,4 % vergleichend auf das jeweils andere DpV bzw. auf beide. DpV wurden entsprechend nur in 20,6 % einzeln, also ohne Bezug zum anderen DpV erwähnt. Bei Aufgabe *Alt2* verglichen HoEr zu 97,4 % DpV miteinander bzw. bezogen sich auf beide.

Auch NiEr beschrieben überwiegend vergleichend und bei beiden Aufgaben mit 74,6 % bzw. 81 % zu ähnlichen Anteilen. Entsprechend beschrieben sie DpV einzeln in 21,4 % bzw. 17,7 %.

Der Bezug der restlichen Aussagen ist unklar, da sie nicht eindeutig zugeordnet werden konnten (Abbildung 10).

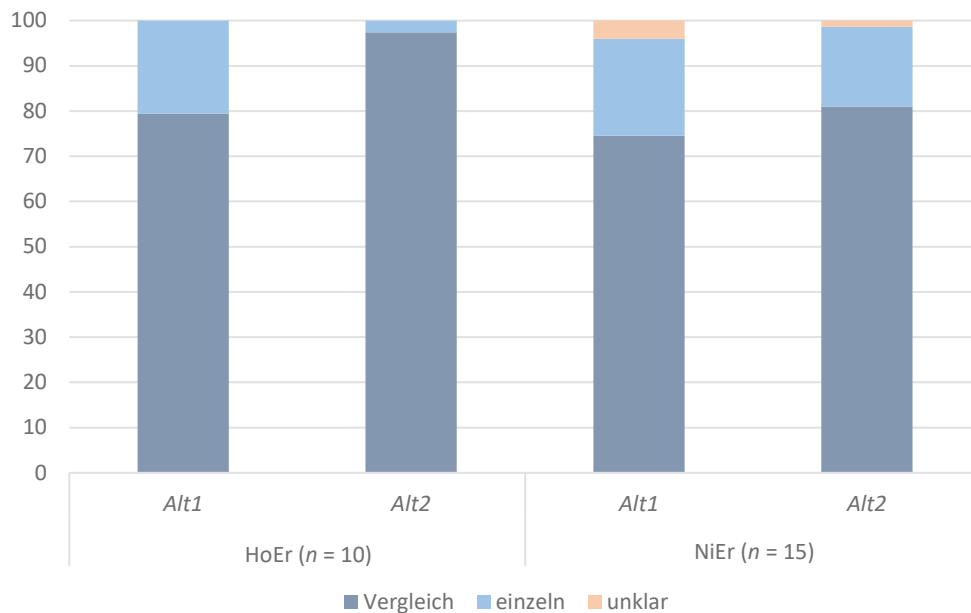


Abbildung 10: Bezug der Aussagen bei der Teilkompetenz Alternative Modelle im medialen Modelleinsatz. DpV können vergleichend oder einzeln beschrieben werden oder der Bezug ist unklar. Bei TR_1b wurde der Bezug nicht erfasst. Die Anzahl der analysierten Proband:innen (n) ist in Klammern angegeben.

Je nach Aufgabe und Teilkompetenz lasen vor allem NiEr bis zu 41 % *wissenschaftlich nicht adäquat ab* (TR_1b). Wie in Abbildung 9 ersichtlich, bezogen sie sich dabei am häufigsten auf MRCA, bei Aufgabe Alt1 auch auf Taxa (genaue Aufschlüsselung Kapitel 6.11). *Wissenschaftlich inadäquates Ablesen* (TR_1b) kam bei der Teilkompetenz *Zweck von Modellen* bei 5 bzw. 3 NiEr vor. Bei der Teilkompetenz *Alternative Modelle lasen* 14 bzw. 10 NiEr und 2 bzw. 1 HoEr *wissenschaftlich inadäquat ab* (Abbildung 11).

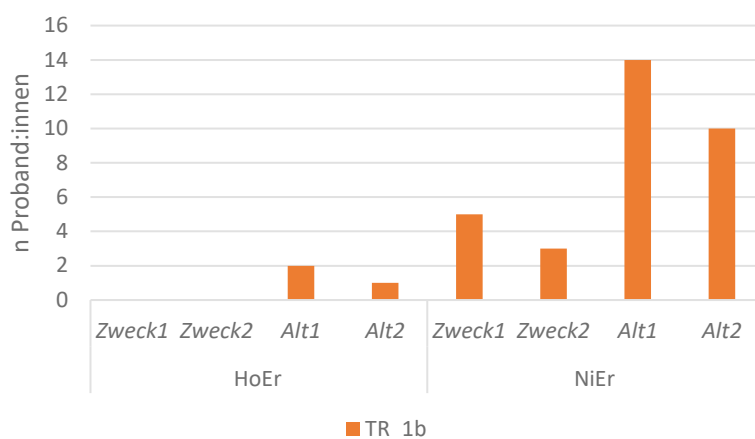


Abbildung 11: Anzahl n der Proband:innen, bei denen im medialen Modelleinsatz wissenschaftlich nicht adäquates Ablesen (TR_1b) kodiert wurde.

6.6. Allgemeine Codehäufigkeiten im methodischen Modelleinsatz

HoEr *leiteten* bei 3 Aufgaben überwiegend *ab* (TR_2; Zweck: 80 % bzw. 76,1 %; Alt1: 89,6 %). *Wissenschaftlich adäquates Ablesen* war zu geringeren Anteilen vertreten (TR_1a; Zweck: 17,5 % bzw. 23,9 %; Alt1: 9,4 %). Bei der Aufgabe *Alt2 lasen* sie überwiegend inhaltlich *ab* (TR_1a; 80,4 %). *Oberflächliches Ablesen* kam bei allen Aufgaben zu sehr geringen Anteilen bzw. gar nicht vor (TR_0; Zweck: 2,5 % bzw. 0 %; Alt: 1 % bzw. 7,1 %). Bei Aufgabe *Alt2* wurde zu 7,2 % *wissenschaftlich nicht adäquat abgelesen* (TR_1b; Abbildung 12a & b).

NiEr *leiteten* bei 3 Aufgaben ebenfalls *ab* (TR_2; Zweck: 62,5 % bzw. 34 %; Alt1: 41,5 %), die Anteile sind insgesamt jedoch niedriger als bei HoEr. Bei Aufgabe *Alt2* überwog *wissenschaftlich adäquates Ablesen* (TR_1a; 67,6 %) ähnlich wie bei HoEr. Bei den anderen Aufgaben war *wissenschaftlich adäquates Ablesen* unterschiedlich häufig vertreten (TR_1a Zweck: 23,4 % bzw. 45,7 %; Alt1: 29,8 %), insgesamt jeweils häufiger als bei HoEr. *Oberflächliches Ablesen* kam bei NiEr bei allen Aufgaben mit geringen Anteilen vor (TR_0; Zweck: 4,7 % bzw. 9,6 %; Alt: 7,4 % bzw. 7 %). Der Anteil an *wissenschaftlich inadäquatem Ablesen* (TR_1b) war insgesamt höher als bei HoEr und variierte zwischen den Teilkompetenzen mit 9,4 % bzw. 10,6 % bei der Teilkompetenz *Zweck von Modellen* und 21,3 % bzw. 25,4 % bei der Teilkompetenz *Alternative Modelle* (Abbildung 12a & b).

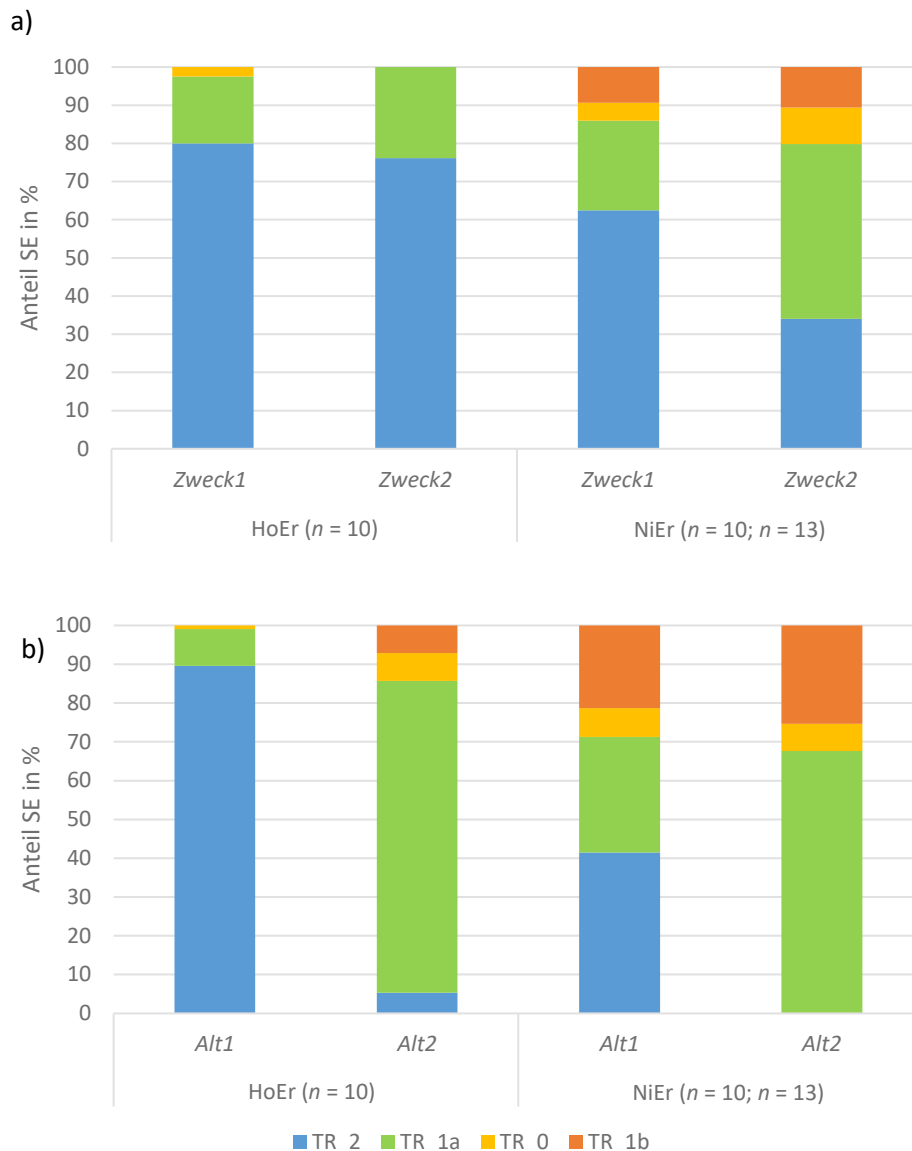


Abbildung 12: Allgemeine Codehäufigkeiten der Tree-Reading-Fähigkeiten in Prozent im methodischen Modelleinsatz der Teilkompetenz Zweck von Modellen (a) und Alternative Modelle (b) für HoEr und NiEr. Kodierte Sinneinheiten bei Zweck von Modellen: HoEr Zweck1: 120, Zweck2: 134, Alt1: 106, Alt2: 56; NiEr Zweck1: 64; Zweck2: 94, Alt1: 94, Alt2: 71. Die Anzahl der analysierten Proband:innen (n) ist in Klammern angegeben. TR_2 = Ableiten, TR_1a = wissenschaftlich adäquates Ablesen, TR_0 = oberflächliches Ablesen, TR_1b = wissenschaftlich inadäquates Ablesen.

Bezogen auf die DpV-Bestandteile in der Teilkompetenz *Zweck von Modellen leiteten* HoEr in beiden Aufgaben am häufigsten Entstehungs- und Verlustereignisse *ab* (E/V; TR_2; 46,6 % bzw. 35 %). Merkmale des *MRCA* (TR_2; 21,7 % bzw. 12 %) wurden ebenfalls *abgeleitet*. Konvergenz wurde bei Aufgabe *Zweck2* häufiger *abgeleitet* (TR_2; 23,1 %) als bei Aufgabe *Zweck1* (TR_2; 5,8 %). Das Sparsamkeitsprinzip wurde von HoEr in beiden Aufgaben mit ähnlichen Anteilen genannt (TR_2; 5,8 % bzw. 6 %).

In beiden Aufgaben wurden Merkmale von *Taxa wissenschaftlich adäquat abgelesen* (TR_1a; 15 % bzw. 9 %). Weitere DpV-Bestandteile, die *wissenschaftlich adäquat abgelesen* wurden, kamen vor allem in Aufgabe *Zweck2* vor (TR_1a; 13,4 %).

NiEr *leiteten* bei beiden Aufgaben am häufigsten Merkmale von *MRCA ab* (TR_2; 32,8 % bzw. 16 %). Bei Aufgabe *Zweck1* folgten in der Häufigkeit Entstehungs- und Verlustereignisse (TR_2; 25 %), bei Aufgabe *Zweck2* Konvergenz (TR_2; 11,7) und deutlich seltener Entstehungs- und Verlustereignisse (TR_27,45 %). Bei beiden Aufgaben *lasen* NiEr häufig Merkmale von *Taxa wissenschaftlich adäquat ab* (TR_1a; 21,9 % bzw. 26,6 %). Bei Aufgabe *Zweck2* kamen noch weitere DpV-Bestandteile dazu, die *wissenschaftlich adäquat abgelesen* wurden (TR_1a; 18,1 %). Bei beiden Aufgaben konnten insgesamt 14 % bzw. 20,2 % der Aussagen *oberflächlichem Ablesen* (TR_0) bzw. *wissenschaftlich inadäquatem Ablesen* (TR_1b) zugeordnet werden (Abbildung 13).

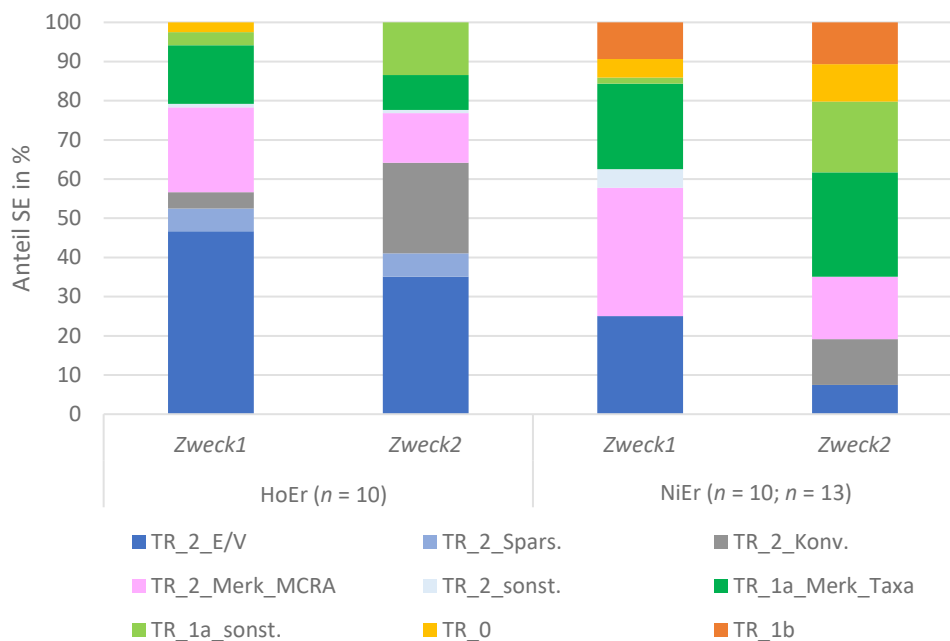


Abbildung 13: Zuordnung der DpV-Bestandteile zu Tree-Reading-Fähigkeiten in Prozent der Teilkompetenz Zweck von Modellen im methodischen Modelleinsatz für HoEr und NiEr. Zweck1: 120, Zweck2: 134; NiEr Zweck1: 64; Zweck2: 94. Die Anzahl der analysierten Proband:innen (n) ist in Klammern angegeben. TR_2 = Ableiten, TR_1a = wissenschaftlich adäquates Ablesen, TR_0 = oberflächliches Ablesen, TR_1b = wissenschaftlich inadäquates Ablesen. DpV-Bestandteile: E/V = Entstehungs-/ Verlustereignis, Spars. = Sparsamkeitsprinzip, Konv. = Konvergenz, Merk_MCRA = Merkmal eines MRCA, Merk_Taxa = Merkmal eines Taxons/ von Taxa, sonst = sonstiges.

Beim methodischen Modelleinsatz der Teilkompetenz *Alternative Modelle* erwähnten sowohl HoEr als auch NiEr in beiden Aufgaben unterschiedliche DpV-Bestandteile. Daher werden die beiden Aufgaben in Abbildung 14 getrennt voneinander dargestellt.

HoEr *leiten* in Aufgabe *Alt1* überwiegend Entstehungs- und Verlustereignisse *ab* (E/V; TR_2; 65,1 %) und zu geringeren Anteilen das Sparsamkeitsprinzip (TR_2; 9,4 %) sowie Merkmale von Taxa *ab* (TR_2; 9,4 %). *Wissenschaftlich adäquat* wurden zu einem kleinen Teil Kladen *abgelesen* (TR_1a; 8,5 %; Abbildung 14a).

Auch NiEr *leiten* in Aufgabe *Alt1* am häufigsten Entstehungs- und Verlustereignisse *ab* (TR_2; 18,1 %), jedoch deutlich weniger als HoEr. Sie *leiten* außerdem auch Merkmale von MRCA (TR_2; 11,7 %) und Merkmale von Taxa (TR_2; 7,45 %) *ab*. NiEr *lasen* zu einem recht hohen Anteil *wissenschaftlich adäquat* vor allem Kladen (TR_1a; 22,34 %) sowie sonstige DpV-Bestandteile (TR_1a; 7,45 %) *ab*. Weiterhin wurden mit großen Anteilen DpV-Bestandteile *oberflächlich* (TR_0; 7,45 %) bzw. *wissenschaftlich nicht adäquat* (TR_1b; 21,28 %) *abgelesen* (Abbildung 14a).

Bei Aufgabe *Alt2* *lasen* HoEr vor allem Kladen (TR_1a; 59 %) sowie Verwandtschaftsbeziehungen (TR_1a; 21,43 %) *wissenschaftlich adäquat ab*. Die restlichen Sinneinheiten entfielen mit 5 % bis 7 % auf die 3 anderen Tree-Reading-Fähigkeiten (Abbildung 14b).

NiEr *lasen* ebenfalls vor allem Kladen (TR_1a; 45,07 %) und Verwandtschaftsbeziehungen (TR_1a; 16,9 %) *wissenschaftlich adäquat ab*. Ein großer Anteil der Sinneinheiten entfiel auf *wissenschaftlich nicht adäquates Ablesen* (TR_1b; 25,35 %; Abbildung 14b).

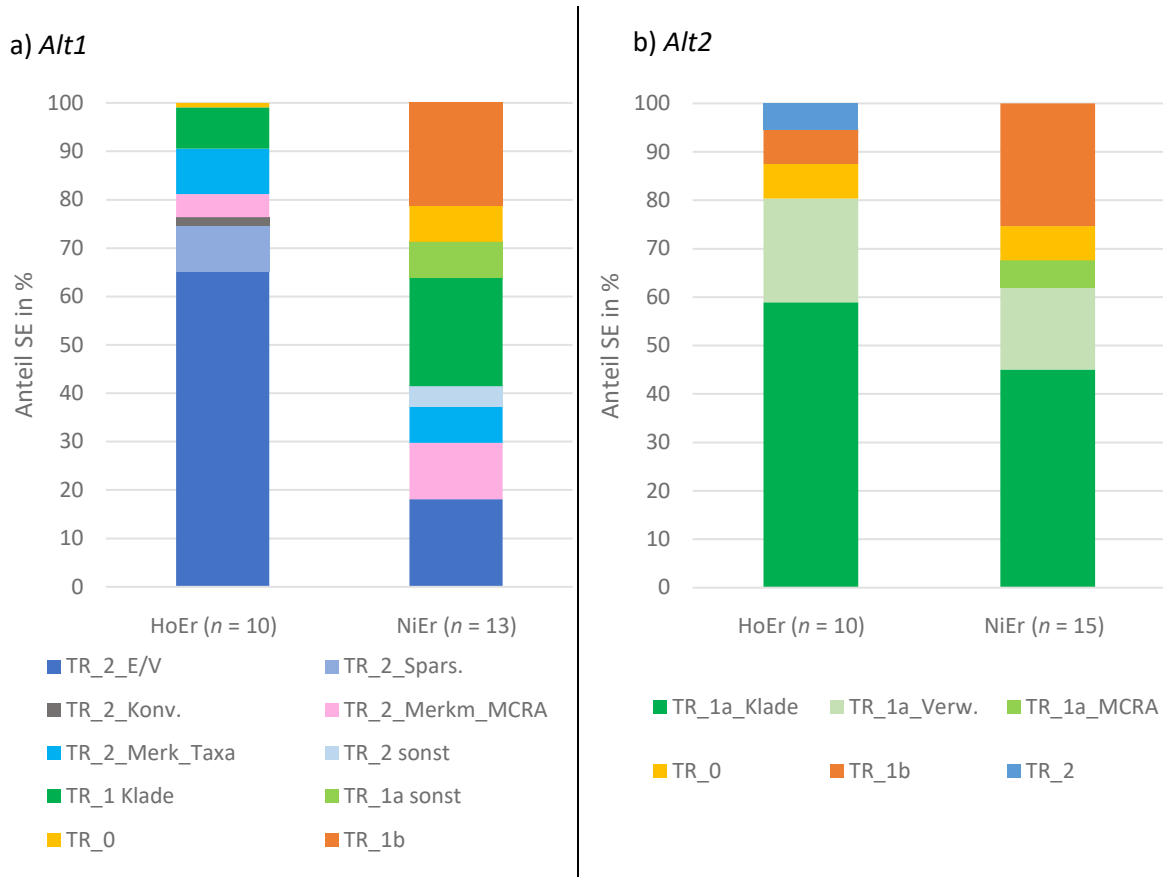


Abbildung 14: Zuordnung der DpV-Bestandteile zu Tree-Reading-Fähigkeiten in Prozent der Teilkompetenz Alternative Modelle im methodischen Modelleinsatz für HoEr und NiEr. Kodierte Sinneinheiten HoEr Alt1: 106, Alt2: 56; NiEr Alt1: 94, Alt2: 71. Die Anzahl der analysierten Proband:innen (n) ist in Klammern angegeben. TR_2 = Ableiten, TR_1a = wissenschaftlich adäquates Ablesen, TR_0 = oberflächliches Ablesen, TR_1b = wissenschaftlich inadäquates Ablesen. DpV-Bestandteile: E/V = Entstehungs-/ Verlustereignis, Konv. = Konvergenz, Merk_Taxa = Merkmal eines Taxons/ von Taxa, Spars. = Sparsamkeitsprinzip, Merkm_MCRA = Merkmal eines MRCA, sonst = sonstiges, Verw. = Verwandtschaft.

Der Bezug auf beide oder ein einzelnes DpV unterscheidet sich bei den Erfahrungsgruppen je nach Aufgabe. HoEr verwendeten bei Aufgabe *Alt1* die DpV überwiegend vergleichend (68 %), bei Aufgabe *Alt2* eher einzeln (57,1 %). Bei NiEr waren die Anteile entgegengesetzt verteilt, mit überwiegendem Erwähnen eines einzelnen DpV (59,6 %) bei Aufgabe *Alt1* und überwiegendem Vergleichen bei Aufgabe *Alt2* (67,6 %; Abbildung 15).

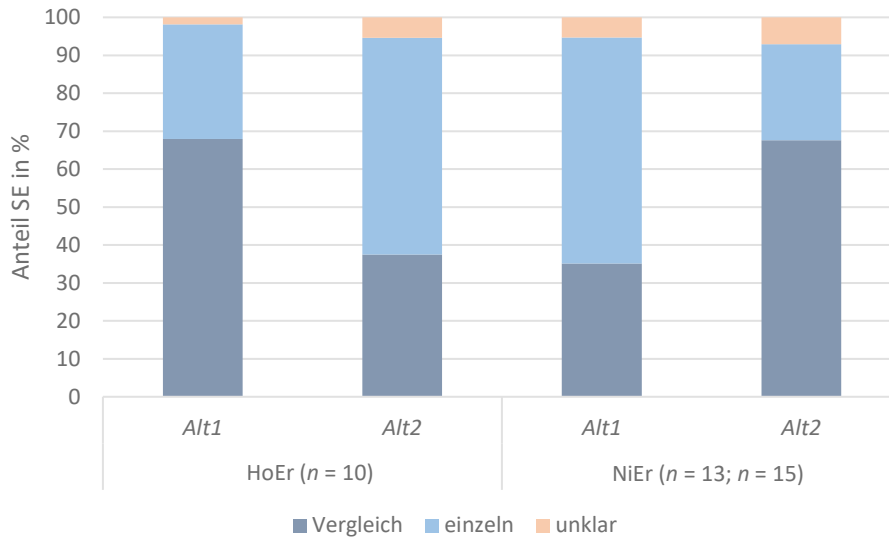


Abbildung 15: Bezug der Aussagen bei der Teilkompetenz Alternative Modelle im methodischen Modelleinsatz. DpV können vergleichend oder einzeln beschrieben werden oder der Bezug ist unklar. Kodierte Sinneinheiten HoEr Alt1: 106, Alt2: 56; NiEr Alt1: 94, Alt2: 71. Die Anzahl der analysierten Proband:innen (n) ist in Klammern angegeben.

Betrachtet man die absolute Anzahl von Proband:innen, die *wissenschaftlich inadäquat ablesen*, finden sich bei der Teilkompetenz *Zweck von Modellen* 3 bzw. 5 NiEr, bei der Teilkompetenz *Alternative Modelle* 9 bzw. 8 NiEr, bei Aufgabe *Alt2* auch 3 HoEr (Abbildung 16). In Kapitel 6.11 wird näher erläutert, inwiefern *wissenschaftlich inadäquat abgelesen* wird.

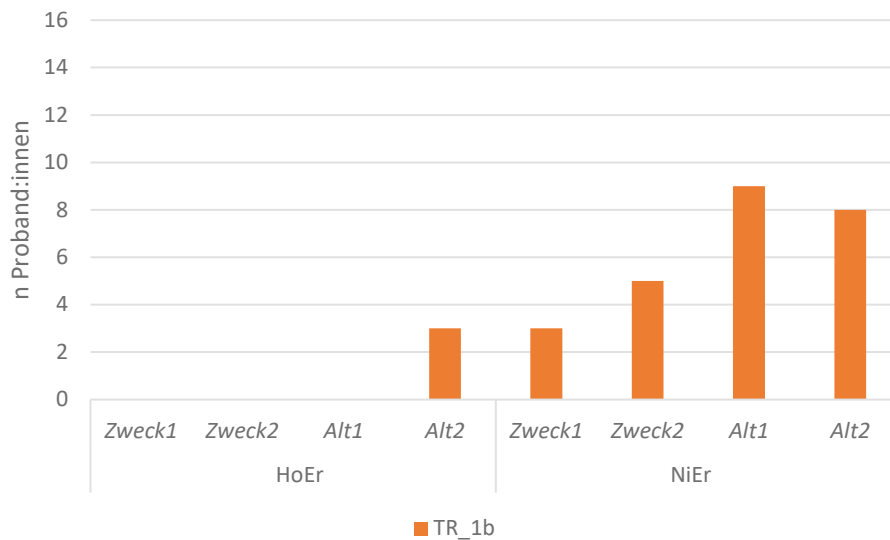


Abbildung 16: Anzahl n der Proband:innen, bei denen im methodischen Modelleinsatz wissenschaftlich nicht adäquates Ablesen (TR_1b) kodiert wurde.

6.7. Maximale Tree-Reading-Fähigkeit und Lösung der Aufgaben

Im Gegensatz zur allgemeinen Analyse der Tree-Reading-Fähigkeiten, DpV-Bestandteile und Bezüge im vorangehenden Kapitel wird nun die individuelle Ebene betrachtet. Hierfür wurden zunächst die maximal gezeigte Tree-Reading-Fähigkeit der einzelnen Proband:innen ermittelt, basierend auf der Annahme, dass inhaltliches, *wissenschaftlich adäquates Ablesen* (TR_1a) anspruchsvoller ist als *wissenschaftlich inadäquates Ablesen* (TR_1b) bzw. *oberflächliches Ablesen* (TR_0). Am anspruchsvollsten wird das erkenntnisgenerierende Erklären und *Ableiten von Informationen* (TR_2) definiert. So wurde innerhalb jedes Transkripts die anspruchsvollste Tree-Reading-Fähigkeit ausgewählt, auch wenn sie anteilig seltener vorkam als andere Tree-Reading-Fähigkeiten.

Alle 10 HoEr zeigten bei allen medialen Aufgaben als maximale Tree-Reading-Fähigkeit *wissenschaftlich adäquates Ablesen* (TR_1a). Bei NiEr variierten die maximalen Tree-Reading-Fähigkeiten je nach Teilkompetenz und Aufgabe. Bei der Teilkompetenz *Zweck von Modellen* zeigten 7 bzw. 10 von 14 NiEr *wissenschaftlich adäquates Ablesen* (TR_1a) als maximale Tree-Reading-Fähigkeit, 6 bzw. 4 *oberflächliches Ablesen* (TR_0) und 1 NiEr bei Aufgabe *Zweck1* *wissenschaftlich inadäquates Ablesen* (TR_1b). Je 1 NiEr konnte nicht kodiert werden. Bei der Teilkompetenz *Alternative Modelle* zeigten je 13 von 15 NiEr *wissenschaftlich adäquates Ablesen* (TR_1a) als maximale Tree-Reading-Fähigkeit, 2 bzw. 1 *oberflächliches Ablesen* (TR_0) und 1 NiEr bei Aufgabe *Alt1* *wissenschaftlich inadäquates Ablesen* (TR_1b; Abbildung 17).

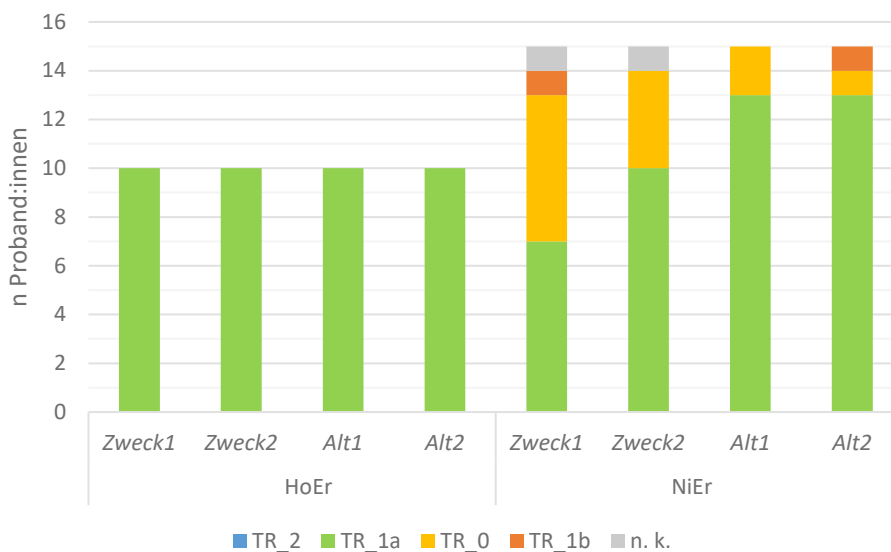


Abbildung 17: Maximal erreichte Tree-Reading-Fähigkeit im medialen Modelleinsatz der einzelnen Proband:innen. TR_2 = Ableiten, TR_1a = wissenschaftlich adäquates Ablesen, TR_0 = oberflächliches Ablesen, TR_1b = wissenschaftlich inadäquates Ablesen, n. k. = nicht kodierbar.

Beim methodischen Modelleinsatz zeigten bei den Aufgaben *Zweck1*, *Zweck2* und *Alt1* alle 10 HoEr *Ableiten* (TR_2) als maximale Tree-Reading-Fähigkeit. Bei Aufgabe *Alt2* kam es bei einer HoEr zu *Ableiten* (TR_2) und 9 HoEr zu *wissenschaftlich adäquatem Ablesen* (TR_1a) als maximale Tree-Reading-Fähigkeit. Wie beim medialen Modelleinsatz unterschieden sich die maximalen Tree-Reading-Fähigkeiten bei NiEr zwischen den Teilkompetenzen. Bei der Teilkompetenz *Zweck von Modellen* wiesen 10 bzw. 11 NiEr *Ableiten* (TR_2) als maximale Tree-Reading-Fähigkeit auf, bei Aufgabe *Zweck2* zwei NiEr *wissenschaftlich adäquates Ablesen* (TR_1a). Die restlichen 5 bzw. 2 NiEr konnten nicht kodiert werden. Bei der Teilkompetenz *Alternative Modelle* zeigten 9 NiEr bei Aufgabe *Alt1* *Ableiten* (TR_2), 3 *wissenschaftlich adäquates Ablesen* (TR_1a) und 1 NiEr *wissenschaftlich inadäquates Ablesen* (TR_1b) als maximale Tree-Reading-Fähigkeit. 2 NiEr konnten nicht kodiert werden. Bei Aufgabe *Alt2* kam es bei 14 von 15 NiEr zu *wissenschaftlich adäquatem Ablesen* (TR_1a) und bei 1 NiEr zu *oberflächlichem Ablesen* (TR_0) als maximale Tree-Reading-Fähigkeit (Abbildung 18).

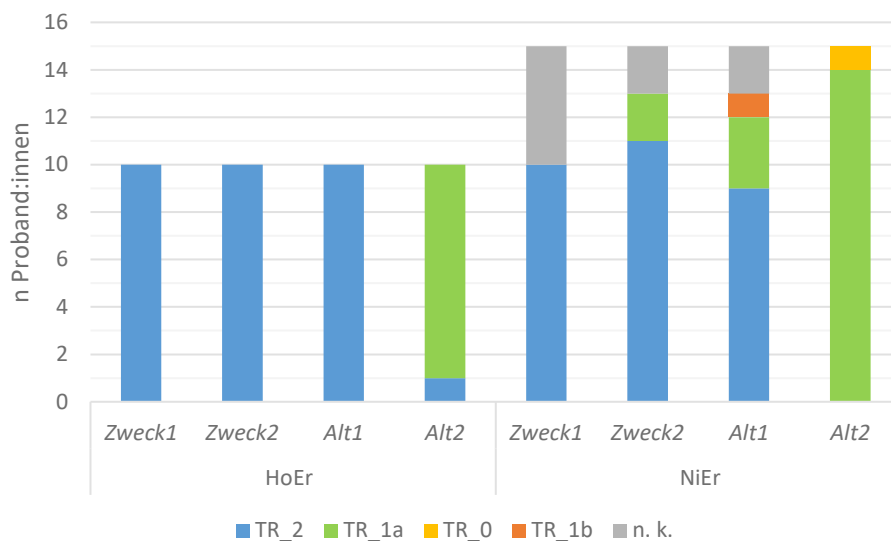


Abbildung 18: Maximal erreichte Tree-Reading-Fähigkeit im methodischen Modelleinsatz der einzelnen Proband:innen. TR_2 = Ableiten, TR_1a = wissenschaftlich adäquates Ablesen, TR_0 = oberflächliches Ablesen, TR_1b = wissenschaftlich inadäquates Ablesen, n. k. = nicht kodierbar.

Hinsichtlich der wissenschaftlichen Qualität der Lösungsprozesse durch die Proband:innen werden 2 unterschiedliche Maße herangezogen. Für den medialen Modelleinsatz die oben beschriebenen maximalen Tree-Reading-Fähigkeiten, die bei der Teilkompetenz *Zweck von Modellen* im Lösungsprozess zum Beschreiben von *MRCA* gezeigt und bei der Teilkompetenz *Alternative Modelle* zum Beschreiben bzw. Vergleichen der *DpV* genutzt wurden (Abbildung 17).

Beim methodischen Modelleinsatz wird als Maß die individuelle Lösung einer Aufgabe hinsichtlich ihrer Entsprechung der wissenschaftlichen Konventionen beurteilt: *den wissenschaftlichen Konventionen entsprechend, von den wissenschaftlichen Konventionen abweichend, ohne Lösung und nicht kodierbar.*²¹ Im Gegensatz zu den vorhergehenden Maßen wird hier nicht der allgemeine Lösungsprozess betrachtet, sondern die abschließende Lösung der Aufgabe, auf die sich ein:e Proband:in verbal festlegt.

Bei der Teilkompetenz *Zweck von Modellen* lösten alle 10 HoEr die beiden Aufgaben *den wissenschaftlichen Konventionen entsprechend*. Bei der Teilkompetenz *Alternative Modelle* lösten 9 von 10 HoEr die Aufgabe *den wissenschaftlichen Konventionen entsprechend* und je 1 HoEr *abweichend*.

Bei der Teilkompetenz *Zweck von Modellen* bearbeiteten 7 bzw. 6 NiEr die Aufgaben *den wissenschaftlichen Konventionen entsprechend*, 3 bzw. 5 *abweichend* und bei Aufgabe *Zweck2* nannten 2 NiEr keine identifizierbare Lösung (*ohne*). Die Bearbeitungen von je 5 bzw. 2 NiEr waren *nicht kodierbar*. Bei der Teilkompetenz *Alternative Modelle* lösten 7 bzw. 12 NiEr die Aufgaben *den wissenschaftlichen Konventionen entsprechend*, bei Aufgabe *Alt1* wichen 5 NiEr davon *ab* und die Bearbeitungen von 2 NiEr waren *nicht kodierbar*. Bei beiden Aufgaben nannten 1 bzw. 3 NiEr keine identifizierbare Lösung (*ohne*; Abbildung 19).

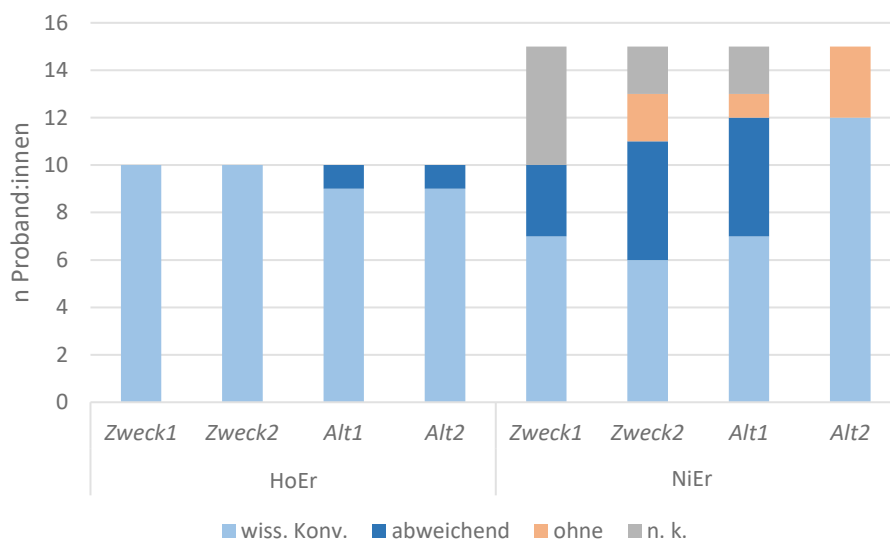


Abbildung 19: Klassifizierung der individuellen Lösungen von HoEr und NiEr im methodischen Modelleinsatz hinsichtlich ihrer Entsprechung der wissenschaftlichen Konventionen. *Wiss. Konv.* = den wissenschaftlichen Konventionen entsprechend, *abweichend* = von den wissenschaftlichen Konventionen abweichend, *ohne* = ohne Lösung, *n. k.* = nicht kodierbar.

²¹ Im Kodierleitfaden (Anhang 7) wurden ursprünglich andere Formulierungen verwendet, die aber bedeutungsgleich sind.

6.8. Interindividueller Vergleich von Lösungsprozessen

Für den Vergleich der Lösungsprozesse verschiedener Proband:innen über alle Aufgaben hinweg sind in Tabelle 14 die in Kapitel 6.7 berichteten maximalen Tree-Reading-Fähigkeiten (medialer und methodischer Modelleinsatz) und individuellen Lösungen (methodischer Modelleinsatz) aufgetragen. Die Proband:innen sind innerhalb ihrer Erfahrungsgruppe danach geordnet, in welchem Modelleinsatz ihre maximalen Tree-Reading-Fähigkeiten bzw. ihre individuellen Lösungen von den wissenschaftlichen Konventionen abweichen. Es ergeben sich 4 Gruppen: *ohne* Abweichung, Abweichungen nur in Aufgaben des *medialen* Modelleinsatzes, Abweichungen nur in Aufgaben des *methodischen* Modelleinsatzes sowie Abweichungen in Aufgaben des *medialen und methodischen* Modelleinsatzes.

8 von 10 HoEr wichen weder in der maximalen Tree-Reading-Fähigkeit noch in ihren Lösungen von den wissenschaftlichen Konventionen ab. Lediglich HoEr P25 und HoEr P29 lösten je eine Aufgabe im *methodischen Modelleinsatz* abweichend.

Die Lösungen von 2 NiEr (P07, P13) entsprachen in allen Aufgaben den wissenschaftlichen Konventionen. 2 NiEr (P16, P17) wichen jeweils bei einer *medialen* Aufgabe der Teilkompetenz *Zweck von Modellen* davon ab. 4 NiEr (P17, P19, P20, P22) zeigten beim *methodischen* Modelleinsatz Abweichungen bzw. ihre Aussagen sind in einzelnen Aufgaben *nicht kodierbar*. Pro Proband:in kamen 2 bis 3 Abweichungen vor. Mit 7 NiEr (P08, P09, P11, P12, P15, P21, P27) weicht knapp die Hälfte der insgesamt 15 NiEr sowohl im *medialen als auch im methodischen* Modelleinsatz ein- bis viermal bzw. zwei- bis viermal von den wissenschaftlichen Konventionen ab oder die Aussagen sind *nicht kodierbar* (Tabelle 14).

Im Folgenden werden individuelle Lösungsprozesse von Proband:innen vorgestellt, die repräsentativ für die jeweiligen Aufgaben bzw. Formen der Abweichungen sind. Tabelle 15 gibt einen Überblick über die Auswahl.

Tabelle 14: Individuelle Lösungsprozesse bzw. Lösungen der Aufgaben. Zeilen entsprechen Proband:innen, Spalten entsprechen Aufgaben, geordnet nach Teilkompetenz und Modelleinsatz. Proband:innen sind in **4 Gruppen hinsichtlich ihrer Abweichungen** geordnet: ohne = keine Abweichung von wissenschaftlichen Konventionen, medial = in mind. einer medialen Aufgabe, methodisch = in mind. einer methodischen Aufgabe, medial + methodisch = in mind. einer medialen und methodischen Aufgabe. Die **maximal erreichte Tree-Reading-Fähigkeit ist durch die Farbe der Zelle in der Aufgabe angegeben**: blau = Ableiten (TR_2), grün = wissenschaftlich adäquates Ablesen (TR_1a), gelb = oberflächliches Ablesen (TR_0), orange = wissenschaftlich inadäquates Ablesen (TR_1b). **Zellenbeschriftung im Falle einer Abweichung** der Lösung von wissenschaftlichen Konventionen bzw. maximaler Tree-Reading-Fähigkeit: n. k. = nicht kodierbar, Tree-Reading = Abweichende Tree-Reading-Fähigkeit, and. = andere Bearbeitung, ohne = keine erkennbare Lösung, PS = Punkte und Streifen als Merkmal des MRCA, MRCA = MRCA hat Merkmal, späterer Verlust, Topo = Topologie.

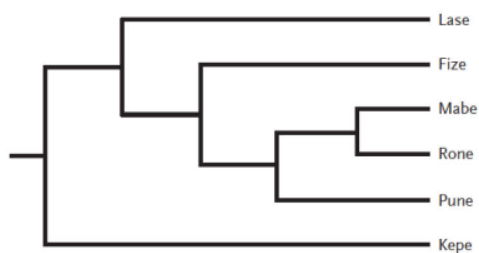
| Erfahrungsgrad | Abweichung | Proband:in | Medialer Modelleinsatz | | | | Methodischer Modelleinsatz | | | |
|----------------|---------------------|--|-----------------------------------|--------|-------------------|-------|----------------------------|---------------------|------------------------|----------------------|
| | | | Zweck1 | Zweck2 | Alt1 | Alt2 | Zweck1 | Zweck2 | Alt1 | Alt2 |
| | | | MRCA identifizieren + beschreiben | | 2 DpV vergleichen | | Merkmal MRCA ableit. | Konvergenz ableiten | Passendes DpV identif. | Hypothese aufstellen |
| HoEr | ohne | P10, P28, P05, P24, P26, P06, P14, P23 | | | | | | | | |
| | methodisch | P25 | | | | | | and. | | |
| | | P29 | | | | | | | and. | |
| NiEr | ohne | P07 | | | | | | | | |
| | | P13 | | | | | | | | |
| | medial | P18 | TR | | | | | | | |
| | | P16 | | TR | | | | | | |
| | methodisch | P17 | | | | | PS | MRCA | MRCA | |
| | | P20 | | | | | n. k. | | TR, ohne | |
| | | P22 | | | | | n. k. | | TR, Topo | ohne |
| | | P19 | | | | | | MRCA | MRCA | |
| | medial + methodisch | P21 | TR | | | | | n. k. | | ohne |
| | | P27 | TR | | | | PS | MRCA | | |
| | | P09 | n. k. | n. k. | | | PS | MRCA | TR, and. | |
| | | P11 | TR | TR | | | n. k. | TR, ohne | n. k. | 1 |
| | | P08 | TR | TR | | TR | | n. k. | n. k. | |
| P15 | | TR | TR | TR | TR | n. k. | MRCA | | | |
| P12 | | TR | | TR | | n. k. | TR, ohne | TR, teilw | TR, ohne | |

Tabelle 15: Repräsentative Beispiele für Lösungsprozesse sind farblich hervorgehoben. Zeilen entsprechen Proband:innen, Spalten entsprechen Aufgaben, geordnet nach Teilkompetenz und Modelleinsatz. Proband:innen sind in **4 Gruppen hinsichtlich ihrer Abweichungen** geordnet: ohne = keine Abweichung von wissenschaftlichen Konventionen, medial = in mind. einer medialen Aufgabe, methodisch = in mind. einer methodischen Aufgabe, medial + methodisch = in mind. einer medialen und methodischen Aufgabe. Die **maximal erreichte Tree-Reading-Fähigkeit** ist durch die Farbe der Zelle in der jeweiligen Aufgabe angegeben und markiert ausgewählte Aufgaben: blau = Ableiten (TR_2), grün = wissenschaftlich adäquates Ablesen (TR_1a), gelb = oberflächliches Ablesen (TR_0), orange = wissenschaftlich inadäquates Ablesen (TR_1b). **Zellenbeschriftung im Falle einer Abweichung** der Lösung von wissenschaftlichen Konventionen bzw. maximaler Tree-Reading-Fähigkeit: n. k. = nicht kodierbar, Tree-Reading = Abweichende Tree-Reading-Fähigkeit, PS = Punkte und Streifen als Merkmal des MRCA, MRCA = MRCA hat Merkmal, späterer Verlust.

| | | | Aufgaben | | | | | | | |
|----------------|---------------------|------------|-----------------------------------|--------|-------------------|------|----------------------------|---------------------|------------------------------|----------------------|
| | | | Medialer Modelleinsatz | | | | Methodischer Modelleinsatz | | | |
| | | | Zweck1 | Zweck2 | Alt1 | Alt2 | Zweck1 | Zweck2 | Alt1 | Alt2 |
| Erfahrungsgrad | Abweichung | Proband:in | MRCA identifizieren + beschreiben | | 2 DpV vergleichen | | Merkmal MRCA ableiten | Konvergenz ableiten | Passendes DpV identifizieren | Hypothese aufstellen |
| HoEr | ohne | P06 | | | | | | | | |
| | | P14 | | | | | | | | |
| NiEr | medial | P16 | | TR | | | | | | |
| | methodisch | P17 | | | | | PS | MRCA | MRCA | |
| | | P20 | | | | | n. k. | | TR, ohne | |
| | medial + methodisch | P15 | TR | TR | TR | TR | n. k. | MRCA | | |
| | | P27 | TR | | | | PS | MRCA | | |

6.8.1. MRCA identifizieren und beschreiben (Aufgabe Zweck1 medialer Modelleinsatz)

Beim medialen Modelleinsatz der Teilkompetenz *Zweck von Modellen* soll bei beiden Aufgaben der MRCA bestimmter Taxa identifiziert und beschrieben werden. Im Folgenden werden Beispiele der Aufgabe Zweck1 (Abbildung 20) beschrieben. Die Zitate sind für die bessere Lesbarkeit sprachlich leicht geglättet, indem Füllwörter, abgebrochene Wörter und kurze Pausen entfernt wurden. Vor den wörtlichen Zitaten sind in eckigen Klammern die jeweils kodierte Tree-Reading-Fähigkeit und der DpV-Bestandteil angegeben. Alle Proband:innen beider Erfahrungsgruppen haben den korrekten MRCA bestimmt, unterschieden sich jedoch in der Beschreibung.



Die Analyse verschiedener genetischer Marker ergab diesen Stammbaum. Eine zusätzliche morphologische Analyse zeigt, dass nur Lase, Pune und Kepe Streifen haben. Alle anderen gezeigten Taxa haben nur Punkte.

Welche Stelle markiert den letzten gemeinsamen Vorfahren von Rone und Fize? Beschreiben Sie die Stelle.

Abbildung 20: Aufgabe Zweck1 medialer Modelleinsatz: Identifizierung und Beschreibung des MRCA zweier Taxa.

HoEr beschrieben den MRCA überwiegend durch *wissenschaftlich adäquates Ablesen* (TR_1a) und nannten vor allem MRCA und Kladen. Beispielhaft hierfür ist **HoEr P06**, die den MRCA *wissenschaftlich adäquat* (TR_1a) als MRCA einer Klade verortete und zusätzlich die Schwesterklade nannte:

[TR_1a, MRCA] „[der] letzte gemeinsame Vorfahre von Rone und Fize ist der, zu der Gruppierung ab Fize geführt hat

[TR_1a, Klade] und damit die Schwestergruppe zu Lase.“

(HoEr P06, Zweck1 medialer Modelleinsatz)

Eine weitere, bei HoEr häufige Art des Lösungsprozesses ist die Kombination aus *wissenschaftlich adäquater* (TR_1a) und *oberflächlicher* (TR_0) Beschreibung des MRCA. **HoEr P14** beschrieb den MRCA zunächst *wissenschaftlich adäquat* (TR_1a), indem sie alle Taxa bzw. Kladen angab, die sich von diesem MRCA abspalten. Zusätzlich verortete HoEr P14 den MRCA durch *oberflächliches* (TR_0) Abzählen der Knotenpunkte von der Wurzel an:

[TR_1a, MRCA] *„Der letzte gemeinsame Vorfahre von Fize und Rone ist an dem Knotenpunkt gewesen, an dem sich der Baum zu Fize (.) ab/ also die eine Seite vom Baum zu Fize abspaltet und die andere Seite zu der Linie, die zu der Großgruppe aus Mabe, Rone und Pune führt (...)*

[TR_0, MRCA] *der eins, zwei, sozusagen dritte Knoten, wenn man von unten zählen würde (..)*“

(HoEr P14, Zweck1 medialer Modelleinsatz)

Bei NiEr kam es selten zu rein *wissenschaftlich adäquaten* (TR_1a) Beschreibungen des MRCA, sondern wie bei HoEr P14 zu Kombinationen mit *oberflächlichen* (TR_0) Beschreibungen. Am häufigsten kamen Kombinationen der Tree-Reading-Fähigkeiten *wissenschaftlich adäquates* (TR_1a), *oberflächliches* (TR_0) und *wissenschaftlich inadäquates Ablesen* (TR_1b).

NiEr P20 beschrieb den MRCA zunächst *wissenschaftlich inadäquat* (TR_1b), indem sie pro Knotenpunkt nur die Abspaltung eines Taxons nannte und nicht berücksichtigte, dass die anderen Taxa auch von diesem MRCA abstammen. Anschließend beschrieb sie den MRCA anhand der in der Instruktion genannten Taxa: Zunächst beschrieb sie *oberflächlich* (TR_0), indem sie von einem in der Instruktion genannten Taxon aus in Richtung Wurzel die Knotenpunkte zählte. Abschließend beschrieb sie den MRCA *wissenschaftlich adäquat* (TR_1a) in Bezug zum anderen in der Instruktion genannten Taxon:

[TR_1b, MRCA] *„der letzte gemeinsame Vorfahr zwischen den beiden ist der Vorfahr, aus dem eben Fize entstanden ist*

[TR_1b, MRCA] *und eben dann darauffolgend Pune*

[TR_1b, MRCA] *und die weitere Unterteilung aus diesem Vorfahr von Pune ist eben dann Mabe und Rone,*

[TR_0, MRCA] *das heißt, würde man zurückgehen, wäre es also von Rone aus gesehen der zweite, dritte Vorfahr,*

[TR_1a, MRCA] *und bei Fize der direkte“*

(NiEr P20, Zweck1 medialer Modelleinsatz)

Ebenfalls häufig war die rein *oberflächliche* (TR_0) Beschreibung des MRCA wie z. B. bei **NiEr P15** anhand des Linienverlaufs von der Wurzel bis zum MRCA bzw. zu den in der Instruktion genannten Taxa:

[TR_0, MRCA] *„wenn man davon ausgeht, dass links die Urart ist,*

[TR_0, MRCA] *geht man den Stamm oben horizontal hoch, geht dann nach rechts, geht dann nach unten und geht dann nach rechts. So, da haben wir einen Vorfahren. Das ist der letzte gemeinsame Vorfahre (..)*

[TR_0, Klade] *Geht es hoch, horizontal und rechts zu Fize (..)*

[TR_0, Klade] *runter, rechts, hoch, rechts, runter, rechts zur Rone (..)“*

(NiEr P15, Zweck 1 medialer Modelleinsatz)

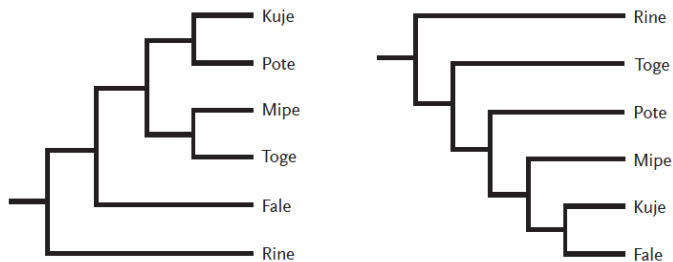
NiEr P27 beschrieb ebenfalls ausschließlich *oberflächlich* (TR_0) durch Abzählen der Knotenpunkte:

[TR_0, MRCA] *„also von links gelesen der eins, zwei, drei/ die dritte Gabelung von links gelesen war der letzte gemeinsame Vorfahre von Fize und Rone.“*

(NiEr P27, Zweck1 medialer Modelleinsatz)

6.8.2. DpV vergleichen (Aufgabe *Alt1* medialer Modelleinsatz)

Beim medialen Modelleinsatz der Teilkompetenz *Alternative Modelle* sollen in beiden Aufgaben 2 DpV hinsichtlich ihrer Gemeinsamkeiten und Unterschiede miteinander verglichen werden. Beide Erfahrungsgruppen verglichen überwiegend in beiden Aufgaben. Die repräsentativen Fallbeispiele stammen aus beiden Aufgaben *Alt1* (Abbildung 21) und *Alt2* (Abbildung 22). Zusätzlich zur Tree-Reading-Fähigkeit und zum DpV-Bestandteil wird angegeben, ob sich die Proband:innen auf ein *einzelnes* DpV (*_e*) oder *vergleichend* auf ein anderes (*_v*) beziehen. Das links abgebildete DpV wird mit DpV1 bezeichnet, das rechts abgebildete DpV als DpV2. Findet ein Wechsel von einem DpV zum anderen statt, ist dies durch eine fette Hervorhebung gekennzeichnet.



Nur Kuje und Mipe haben Schneidezähne. Fossilfunde führten zu der Hypothese, dass Vorfahren von Fale auch Schneidezähne hatten. Bei rezenten Fale sind diese nicht vorhanden.

Welche Unterschiede und Gemeinsamkeiten weisen die Stammbäume auf? Vergleichen Sie sie miteinander.

Abbildung 21: Aufgabe *Alt1* medialer Modelleinsatz: Vergleichen von 2 DpV.

HoEr beschrieben die DpV allgemein überwiegend *wissenschaftlich adäquat* (TR_1a), teilweise auch *oberflächlich* (TR_0) und ausschließlich mit *wissenschaftlich adäquatem Ablesen* (TR_1a) als maximale Tree-Reading-Fähigkeit. Bei beiden Aufgaben verglichen sie zum Großteil die DpV miteinander.

HoEr P14 beschrieb insgesamt ausschließlich *wissenschaftlich adäquat* (TR_1a) erst DpV2 *einzel*, indem sie die verschachtelten Kladen ausgehend von der kleinsten Klade in Richtung Wurzel nacheinander nannte. Anschließend nannte sie *vergleichend* in DpV1 das Schwestertaxon von Kuje als kleinste Klade und deren Beziehung zu einer weiteren Klade:

- [TR_1a, **DpV2_e**, Klade] *„der eine Stammbaum stellt eine monophyletische Gruppe aus Kuje und Fale (..)“*
- [TR_1a, DpV2_e, Klade] *deren Schwestergruppe Mipe ist (.)*
- [TR_1a, DpV2_e, Klade] *Die Schwestergruppe von dieser (.) Gruppe Mipe, Kuje, Fale (.) wäre Pote (.)*
- [TR_1a, DpV2_e, Klade] *Und Toge wiederum steht im Schwestergruppenverhältnis zu dieser Gruppe aus Pote, Mipe, Kuje, Fale (.)*
- [TR_1a, DpV2_e, Klade] *Rine stellt als frühesten abzweigende Linie wieder die Schwestergruppe zu der Großgruppe aus Toge, Pote, Mipe, Kuje und Fale (..)“*
- [TR_1a, **DpV1_v**, Klade] *Und der linke Baum stellt Kuje als Schwestergruppe zu Pote (..)“*
- [TR_1a, DpV1_v, Klade] *und (.) diese Gruppe als Schwestergruppe zu Mipe und Toge (4)“*

Nachfolgend beschrieb HoEr P14 in DpV1 *wissenschaftlich adäquat* (TR_1a) das Verhältnis zweier Taxa zueinander (Fale und Kuje) und verglich es ebenfalls *wissenschaftlich adäquat* (TR_1a) mit DpV2:

- [TR_1a, DpV1_v, Klade] *„Hier steht Fale nicht als Schwestergruppe zu Kuje, sondern als Schwestergruppe zu der Großgruppe aus Mipe, Toge, Pote und Kuje,*
- [TR_1a, DpV1_v, Klade] *das heißt als eine der recht früh abzweigenden Linien im Baum*
- [TR_1a, **DpV2_v**, Taxa] *auf der rechten Seite sind Fale und Kuje (4) sehr spät abzweigende oder naja, das würde man ja so nicht unbedingt sagen, aber ja sehr abgeleitete Gruppen (4)“*

Abschließend *verglich* HoEr P14 beide DpV direkt hinsichtlich der Außengruppe und der allgemeinen Verwandtschaftsbeziehungen. Insgesamt nannte sie alle Taxa beider DpV:

[TR_1a, **beide**, Klade] *„in beiden Stammbäumen steht Rine als Schwestergruppe zu allen anderen Taxa (..)“*

[TR_1a, beide, Klade] *die beiden Bäume sind sich einig darin, dass Rine die Schwestergruppe zum Rest ist (.)*

[TR_1a, beide, Topol.] *aber die restlichen Verwandtschaftsverhältnisse sind keine Übereinstimmungen (..)“*

(HoEr P14, Alt1 medialer Modelleinsatz)

NiEr beschrieben bzw. verglichen DpV insgesamt je nach Aufgabe zu großen Teilen durch *wissenschaftlich inadäquates Ablesen* (TR_1b), insbesondere trat dies bei Aufgabe Alt1 auf. Bei beiden Aufgaben *verglichen* NiEr überwiegend.

NiEr P15 *verglich* zunächst oberflächlich die Taxanamen zwischen beiden DpV und beschrieb DpV1 *wissenschaftlich inadäquat* (TR_1b) als weniger komplex als DpV2:

[TR_0, **beide**, Taxa] *„Zunächst einmal sind das die gleichen Arten die in den Stammbäumen dargestellt sind (..)“*

[TR_1b, **DpV1_v**, Topol.] *Der linke Stammbaum zeigt ein etwas vereinfachtes Bild, wie das zu den Arten gekommen ist (..)“*

Anschließend *verglich* sie *wissenschaftlich inadäquat* (TR_1b) die Anzahl von MRCA als „Zwischenstationen“ zwischen verschiedenen Taxa. Ausgehend von DpV1 wechselte sie dabei häufig zwischen DpV1 und DpV2:

[TR_0, DpV1_v, MRCA] *„Es gibt eine Stammart, die sich auf der linken Seite befindet, die Ausgangsart (.)“*

[TR_1b, DpV1_v, MRCA] *dann gibt es mehrere Zwischenstationen,*

[TR_1b, DpV1_v, MRCA] *die sind aber nicht so viele wie auf dem rechten Stammbaum, (..)“*

- [TR_1b, **beide**, MRCA] *Rine hat bloß eine Zwischenstation durchlaufen auf beiden Stammbäumen (..)*
- [TR_1b, **DpV1_v**, MRCA] *dann sehen wir, dass bei Fale drei Zwischenstationen sind, bis es zu der Art gekommen ist, zu der rezenten Art,*
- [TR_1b, **DpV2_v**, MRCA] *auf dem rechten Stammbaum sehen wir deutlich mehr Zwischenstationen (..)*
- [TR_1b, **DpV2_v**, MRCA] *bei Fale sind das sogar auf den ersten Blick die meisten Zwischenstationen, (.)*
- [TR_1b, **DpV1_v**, MRCA] *da haben wir ja auf dem linken Stammbaum mehrere Arten, die zu Toge geführt haben (.)*
- [TR_1b, **DpV2_v**, MRCA] *und auf der rechten Seite nur vier, äh, nur drei (.) nach der Stammart (...)*“

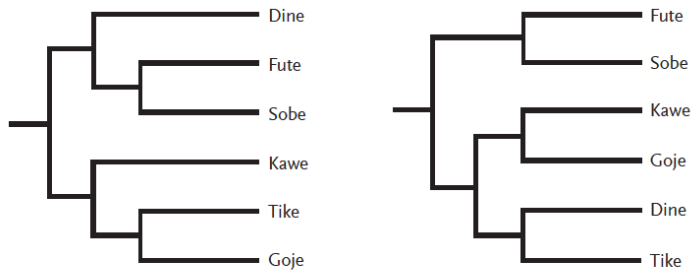
Wie bereits am Anfang des Transkripts beschrieb NiEr P15 wieder *wissenschaftlich inadäquat* (TR_1b) die unterschiedliche Komplexität der 2 DpV mit Bezug auf die Topologie, revidierte diese Aussage jedoch direkt, um abschließend ihr vorheriges Vorgehen des Zählens von MRCA als „Zwischenstationen“ als Vergleichskriterium zusammenfassend zu nennen:

- [TR_1b, **DpV2**, Topo] *„insgesamt verhält es sich so, als wäre auf den ersten Blick der rechte Stammbaum der komplexere, was sich aber im zweiten Blick relativiert, (..)*
- [TR_1b, **unklar**, MRCA] *um dann Rückschlüsse darauf zu ziehen, wie viele Zwischenstationen sie durchlaufen haben (.)*.“

(NiEr P15, *Alt1* medialer Modelleinsatz)

Insgesamt bezog sich der *Vergleich* von NiEr P15 auf 3 von 6 Taxa (Rine, Fale, Toge). Die 3 Taxa wurden nicht miteinander in Bezug gesetzt. Das vorherrschende Vergleichskriterium war die Anzahl der MRCA, die *wissenschaftlich inadäquat* (TR_1b) als „Zwischenstationen“ gezählt wurden. Dieses Vorgehen trat bei vielen NiEr auf.

Bei Aufgabe *Alt2* (Abbildung 22) wurden von NiEr vor allem Kladen beschrieben bzw. verglichen und es kam seltener zu *wissenschaftlich inadäquatem Ablesen* (TR_1b) als bei Aufgabe *Alt1*.



Einer der Stammbäume beruht auf morphologischen, der andere auf genetischen Daten. Die morphologischen Daten deuten darauf hin, dass nur Kawe und Goje ein Monophylum mit ihrem letzten gemeinsamen Vorfahren bilden. Die genetischen Daten führen zu einer anderen Hypothese über dieses Monophylum.

Welche Unterschiede und Gemeinsamkeiten weisen die Stammbäume auf? Vergleichen Sie sie miteinander.

Abbildung 22: Aufgabe Alt2 medialer Modelleinsatz: Vergleichen von 2 DpV.

NiEr **P20** verglich Kladen zwischen den beiden DpV, indem sie häufig zwischen den DpV wechselte. Dabei *las* sie durchgehend *wissenschaftlich adäquat ab* (TR_1a):

- [TR_1a, DpV1_v, Klade] *„auf dem linken Stammbaum sehe ich, dass Fute und Sobe vom selben Vorfahren abstammen*
- [TR_1a, DpV1_v, Klade] *und dazu noch einen weiteren Partner haben, Dine.*
- [DpV2_v] *das ist auf dem rechten Stammbaum nicht zu sehen,*
- [TR_1a, DpV2_v, Klade] *dort gehört Dine mit Tike in eine gemeinsame Gruppe,*
- [TR_1a, DpV2_v, Klade] *also beziehungsweise denselben Vorfahren,*
- [DpV1_v] *das ist links nicht zu sehen.*
- [TR_1a, DpV1_v, Klade] *wenn man den Stammbaum weiter nach unten verfolgt, dann gehören auf der linken Seite Tike und Goje zusammen,*
- [TR_1a, DpV2_v, Klade] *also rechts, wie schon erwähnt, gehören Dine und Tike zusammen,*

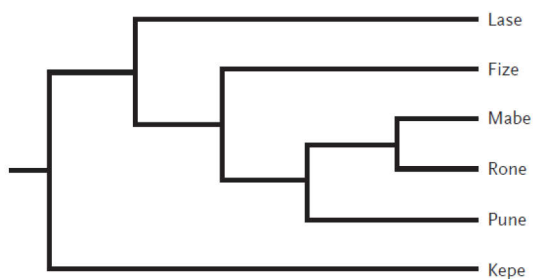
- [TR_1b, **beide**, MRCA] *sie haben nicht denselben Vorfahren.*
- [TR_0, **DpV1_v**, Klade] *Außerdem ist auf dem linken Stammbaum Kawe eine eigene Gruppe für sich und hat keine weiteren Partner,*
- [TR_1a, **DpV2_v**, Klade] *da gehört Goje noch zusammen, (..)“*

(NiEr P20, Alt2 medialer Modelleinsatz)

6.8.3. Merkmal des *MRCA* ableiten (Aufgabe *Zweck1* methodischer Modelleinsatz)

Beim methodischen Modelleinsatz der Teilkompetenz *Zweck von Modellen* soll in Aufgabe *Zweck1* anhand von angegebenen Taxamerkmale *abgeleitet* werden, dass der *MRCA* aller dargestellten Taxa Streifen als Merkmal hatte (Abbildung 23).

Entstehungs- bzw. Verlustereignisse werden mit „E/V“ abgekürzt, konvergente Merkmalsentstehung mit „Konv.“, das Sparsamkeitsprinzip mit „Spars.“ und Merkmal mit „M.“ (z. B. bei Merkmal *MRCA*).



Die Analyse verschiedener genetischer Marker ergab diesen Stammbaum. Eine zusätzliche morphologische Analyse zeigt, dass nur Lase, Pune und Kepe Streifen haben. Alle anderen gezeigten Taxa haben nur Punkte.

Welche Hypothese kann über das Muster des letzten gemeinsamen Vorfahren aller gezeigten Taxa unter Anwendung des Sparsamkeitsprinzips aufgestellt werden?

Abbildung 23: Aufgabe *Zweck1* methodischer Modelleinsatz: Merkmal des *MRCA* ableiten.

Alle HoEr *leiteten* das der Instruktion entsprechende Merkmal des *MRCA* den *wissenschaftlichen Konventionen entsprechend ab* (TR_2), je 5 HoEr nannten dabei 1 bzw. 2 Hypothesen.

HoEr P14 *leitete* zunächst das der Instruktion entsprechende Merkmal des letzten gemeinsamen *MRCA ab* (TR_2; Streifen), indem sie die konvergente Entstehung von Punkten im DpV verortete:

- | | |
|-------------------------|---|
| [TR_2, Spars.] | „ich sage der letzte gemeinsame Vorfahre hatte, wenn man Sparsamkeitsprinzip anwendet (...) |
| [TR_2, M. <i>MRCA</i>] | <i>Streifen, (4)</i> |
| [TR_2, E/V] | <i>weil in diesem Falle wurden die Streifen bei Rone,</i> |
| [TR_2, E/V] | <i>Mabe (.)</i> |
| [TR_2, E/V] | <i>und Fize in Punkte aufgelöst.</i> |

- [TR_2, E/ V] *Und zwar einmal auf der Linie, die zu Fize führt*
 [TR_2, E/ V] *und einmal auf der Linie, die zu Mabe und Rone führt (..)*
 [TR_2, Konv.] *das wären zweimal konvergent Streifen zu Punkte aufgelöst werden würden (..)“*

Anschließend nannte sie eine alternative Hypothese, indem sie von einem anderen Merkmal des MRCA (Punkte) ausgehend die Entstehungsereignisse für Streifen *ableitete* (TR_2):

- [TR_2, M. MRCA] *„Wenn man im Gegensatz dazu Punkte als plesiomorphes Merkmal annehmen würde (..)*
 [TR_2, Konv.] *dann müssten Streifen bei (.) Lase (.),*
 [TR_2, Konv.] *Pune*
 [TR_2, Konv.] *und Kepe (4) dreimal konvergent entstanden sein (..)“*

Zum Schluss verglich sie die beiden Hypothesen hinsichtlich des Sparsamkeitsprinzips und nannte das sparsamer erklärbare Merkmal als wahrscheinlicher:

- [TR_2, Konv.] *„und dreimal konvergentes Entstehen von Streifen auf Punkten (..)“*
 [TR_2, Spars.] *würde ich als weniger sparsam einschätzen*
 [TR_2, E/ V] *als zweimal das Auflösen von Streifen zu Punkten (..)“*

(HoEr P14, Zweck1 methodischer Modelleinsatz)

7 von 15 NiEr konnten das Merkmal des *MRCA* den *wissenschaftlichen Konventionen entsprechend ableiten* (TR_2), 3 *leiteten* eine *abweichende Lösung ab* und die Äußerungen von 5 NiEr waren nicht kodierbar. Insgesamt nannten 9 von 10 Proband:innen 1 Hypothese.

NiEr P16 *leitete* das Merkmal des *MRCA* aller Taxa den *wissenschaftlichen Konventionen entsprechend ab* (TR_2). Im Gegensatz zu HoEr P14 verortete NiEr P16 die Merkmalsentstehung von Punkten jedoch nicht als konvergente Entstehung, sondern *leitete* zunächst ein Entstehungsereignis beim *MRCA* von Fize, Mabe, Rone und eine spätere erneute Merkmalsänderung zu Streifen bei Pune *ab* (TR_2). Ebenso wie die von HoEr P14 genannten Entstehungs- und Verlustereignisse ist auch diese Variante fachlich korrekt und wurde von 5 HoEr bzw. 3 NiEr genannt:

- [TR_1a, M. Taxon] *„Dadurch, dass Pune jetzt auch noch Streifen besitzt,*
 [TR_2, E/ V] *ist es jetzt fraglich ob der gemeinsame Vorfahre von Fize,*
Mabe, Rone und Pune schon eine Mutation zu Punkten
hatte
 [TR_2, E/ V] *und dann Pune also nochmal eine weitere Mutation*
wieder hin zu Streifen durchlaufen hat.“

Anschließend *leitete* NiEr P16 das Merkmal des *MRCA* aller Taxa aus den bekannten Merkmalen der Taxa Kepe und Lase *ab* (TR_2) und wiederholte die entsprechenden Entstehungsereignisse von zunächst Punkten und dann Streifen:

- [TR_0, MRCA] *„[...] dass der gemeinsame Vorfahre von Lase und Kepe,*
der also ganz links die erste Abzweigung des Stammbaums
bezeichnet,
 [TR_2, M. MRCA] *dass der Streifen haben muss,*
 [TR_1a, M. Taxon] *lässt sich aus der Tatsache, dass Lase und Kepe beide*
Streifen haben, ziehen. [...]
 [TR_2, M. MRCA] *Also alle anderen Vorfahren von Fize, Mabe und Rone*
sozusagen/ ob die auch alle Streifen hatten
 [TR_2, E/ V] *und die die alle verloren haben.*
 [TR_2, E/ V] *das wären zwei Änderungen.*

- [TR_2, M. MRCA] *Weil der gemeinsame Vorfahre von Mabe und Rone hätte ja schon keine Streifen haben müssen, sondern Punkte*
- [TR_1a, M. Taxon] *und daraus folgend dann Mabe und Rone.*
- [nicht kodierbar] *ich denke, dass Pune irgendeine Ausnahme darstellt.*
- [TR_1b, Verwandtschaft] *dadurch dass Lase und Kepe am nächsten noch mit dem verwandt sind,*
- [TR_2, M. MRCA] *würde ich sagen, dass der letzte gemeinsame Vorfahre auch Streifen hatte.*
- [nicht kodierbar] *Und Pune irgendeine Ausnahme darstellt, [...]“*

(NiEr P16, Zweck1 methodischer Modelleinsatz)

Insgesamt konnte NiEr P16 das Merkmal des *MRCA* den *wissenschaftlichen Konventionen entsprechend ableiten* (TR_2) und benannte Pune als Ausnahme, spezifizierte dies jedoch nicht näher.

NiEr P17 leitete als Merkmal des *MRCA* sowohl Streifen als auch Punkte auf genetischer Ebene *ab* (TR_2) und begründete diese erste Hypothese durch rezessive Merkmalsausprägung, also auf Individualebene bzw. wie in einem Familienstammbaum. 2 weitere NiEr zeigten ebenfalls eine ähnliche *von den wissenschaftlichen Konventionen abweichende* Lösung mit Punkten und Streifen als ursprüngliches Merkmal:

- [TR_2, M. MRCA] *„dass der gemeinsame Vorfahre aller der gezeigten Taxa, beide Anlagen genetisch involviert haben muss, also sowohl eine Anlage für Streifen als auch eine Anlage für Punkte, da beides auftritt im Laufe der Evolution.*
- [TR_1b, Individualebene] *man kann davon ausgehen, dass es rezessiv ist, da es nicht immer bei jedem auftritt, sondern halt zu bestimmten Gelegenheiten.“*

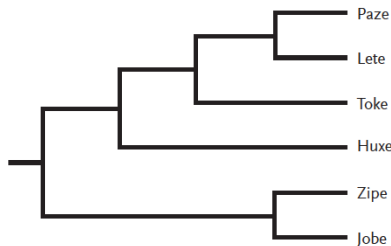
Anschließend stellte sie eine weitere Hypothese auf, legte sich aber nicht endgültig auf eine Hypothese als Lösung fest:

[TR_2, unspez. Mechanismus] *„Eine weitere Hypothese kann sein, dass sich die jeweilige Mustergebung auch durch eine Mutation entwickelt hat, [TR_1a, Verwandtschaft] was wiederum die weiter auseinanderliegende genetische Verwandtschaft von, zum Beispiel, Kepe und Pune bestätigen würde.“*

(NiEr P17, Zweck1 methodischer Modelleinsatz)

6.8.4. Konvergente Merkmalsentstehung ableiten (Aufgabe *Zweck2* methodischer Modelleinsatz)

In Aufgabe *Zweck2* im methodischen Modelleinsatz soll anhand der angegebenen Merkmale von Taxa und *MRCA* *abgeleitet* werden, dass das Merkmal Beinpanzer konvergent entstanden ist (Abbildung 24).



Die Analyse verschiedener genetischer Marker ergab diesen Stammbaum. Eine zusätzliche morphologische Analyse zeigt, dass nur Paze, Lete und Zipe Beinpanzer haben. Weder die anderen Taxa noch zahlreiche gefundene Fossilien weisen Beinpanzer auf.

Welche Hypothese kann über die Entstehung des Merkmals Beinpanzer bei Paze, Lete und Zipe aufgestellt werden? Begründen Sie Ihre Antwort.

Abbildung 24: Aufgabe Zweck2 methodischer Modelleinsatz: Konvergente Merkmalsentstehung ableiten.

Alle HoEr *leiteten* die konvergente Entstehung des Merkmals Beinpanzer als Lösung den *wissenschaftlichen Konventionen entsprechend ab* (TR_2). 5 von 10 HoEr nannten dabei 2 *verschiedene* Hypothesen, 4 jeweils 1 Hypothese und 1 HoEr nannte 3.

HoEr P14 (2 Hypothesen) *leitete* zuerst *ab* (TR_2), dass das Merkmal Beinpanzer konvergent entstanden ist und *leitete* die Entstehungsereignisse korrekt mit dem Sparsamkeitsprinzip *ab* (TR_2):

- [TR_2, Konv.] *„meine Schlussfolgerung ist, dass Beinpanzer konvergent entstanden sind (.)*
- [TR_2, Konv.] *und zwar einmal in der Linie, die zu Paze und Lete führt (.)*
- [TR_2, Konv.] *und einmal konvergent dazu in der Stammgruppe, wo heute Zipe vorhanden ist (...)*
- [TR_2, Spars.] *Und meine Begründung dafür wäre, dass es sich um das sparsamere, also parsimonischere Prinzip handelt (.)*“

Anschließend erläuterte sie eine weitere Hypothese der Merkmalsentstehung ausgehend von Beinpanzern als Merkmal des *MRCA* aller Taxa, indem sie die Anzahl und die Verortung der Entstehungsereignisse *ableitete* (TR_2):

- [TR_2, M. *MRCA*] *„Weil wenn in der Stammart von unserer ganzen Großgruppe Beinpanzer vorhanden gewesen wären,*
 [TR_2, Konv.] *dann wären die eins, zwei, dreimal unabhängig voneinander entstanden. [sic: verloren gegangen]*
 [TR_2, Konv.] *Und zwar einmal bei Toke, (.)*
 [TR_2, Konv.] *einmal bei Huxe (..)*
 [TR_2, Konv.] *und einmal bei Jobe.“*

(HoEr P14, Zweck2 methodischer Modelleinsatz)

6 von 15 NiEr *leiteten* wie HoEr ebenfalls den *wissenschaftlichen Konventionen entsprechend* Konvergenz als Begründung für die angegebenen Taxamerkmale *ab* (TR_2), 5 von ihnen nannten 1 Hypothese, 1 NiEr nannte 2 Hypothesen. Die Transkripte von je 2 NiEr konnten nicht kodiert werden bzw. schilderten unspezifische Lösungsprozesse durch *wissenschaftlich adäquates Ablesen* (TR_1a). Weitere 5 NiEr *leiteten ab* (TR_2), dass das Merkmal Beinpanzer bereits beim *MRCA* aller Taxa vorhanden war (*abweichend von den wissenschaftlichen Konventionen*). 1 NiEr nannte hierbei 2 Hypothesen, alle weiteren 1. Das Nicht-Vorhandensein von Beinpanzern bei mehreren Taxa wurde nur durch 2 der 5 NiEr begründet:

NiEr P15 *leitete ab* (TR_2), dass das Merkmal Beinpanzer bereits beim *MRCA* aller gezeigten Taxa vorhanden war und begründete dies mit der frühen Auftrennung der in der Instruktion genannten Taxa Paze, Lete und Zipe. Das Nicht-Vorhandensein von Beinpanzern in den restlichen Taxa begründete NiEr P15 *wissenschaftlich inadäquat* (TR_1b) damit, dass Merkmale nur auftreten, wenn diese durch Umwelteinflüsse für die Taxa nötig werden:

- [TR_2, M. *MRCA*] *„das Merkmal Beinpanzer bei Paze, Lete und Zipe scheint nun also ein Merkmal zu sein, was auf einen ganz früheren Vorfahren zurückzuführen ist,*

- [TR_1a, Zeitverlauf] *Da Paze, Lete und Zipe sich schon sehr früh in der Evolution getrennt haben, sehr früh, also quasi gleich zu Anfang an,*
- [TR_1b, E/ V] *Es kann sein, dass sie, äh, den gleichen Lebensraum bewohnen, gleiche Fressfeinde haben, wo auch immer Beinpanzer gebraucht werden, scheinen diese Arten diese im Zuge der Evolution nun gebraucht zu haben (.)*
- [TR_1a, Zeitverlauf] *und obwohl sie sich in der Evolution früh voneinander getrennt haben, (..)*
- [TR_1a, M. Taxon] *Wenn die drei Arten Paze, Lete und Zipe alle Beinpanzer ausbilden, (.)*
- [TR_1a, MRCA] *müssen sie einen gemeinsamen Vorfahren haben (4)*
- [abstrakt] *Diese Beinpanzer können Generationen übersprungen haben und können dann im Endeffekt erst wieder in der Art aufgetaucht sein, (.)*
- [abstrakt] *die Anlagen dafür waren aber auf jeden Fall in den Vorfahren immer enthalten, von Anfang an (..)*
- [TR_1a, Klade] *Zum Beispiel haben Toke, Lete und Paze auch einen gemeinsamen Vorfahren.*
- [TR_1a, M. Taxon] *Toke hat dieses Merkmal nicht ausgeprägt, (.)*
- [TR_1a, M. Taxon] *die anderen beiden schon, (..)*
- [TR_1b, E/ V] *was darauf schließen lässt, dass diese Anlage vorhanden war (.) und bei den verschiedenen Arten entweder gebraucht werden konnte oder halt nicht (..)*
- [TR_1a, Klade] *Genauso verhält sich das bei Zipe und Jobe, die haben auch einen gemeinsamen Vorfahren,*
- [TR_1a, M. Taxon] *Zipe hat das Merkmal ausgeprägt,*
- [TR_1a, M. Taxon] *Jobe nicht.“*

(NiEr P15, Zweck2 methodischer Modelleinsatz)

6.8.5. Vorgegebene Hypothese einem DpV zuordnen (Aufgabe *Alt1* methodischer Modelleinsatz)

Beim methodischen Modelleinsatz der Teilkompetenz *Alternative Modelle* wird im Aufgabenstamm der Aufgabe *Alt1* eine Hypothese formuliert, die einem der dargestellten DpV (DpV2) begründet zugeordnet werden soll (Abbildung 25).

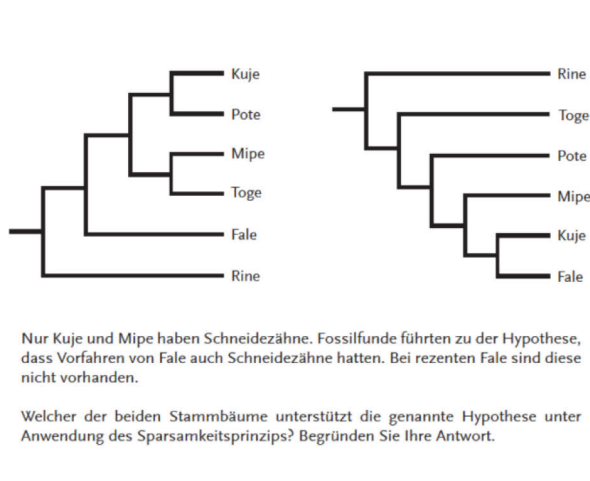


Abbildung 25: Aufgabe *Alt1* methodischer Modelleinsatz: Vorgegebene Hypothese einem DpV zuordnen.

9 von 10 HoEr wendeten die genannte Hypothese auf beide DpV an und ordneten sie abschließend *den wissenschaftlichen Konventionen entsprechend* dem passenden DpV (DpV2) zu.

HoEr P14 ordnete die Hypothese DpV2 zu, indem sie zunächst die Entstehung bzw. den Verlust des Merkmals Schneidezähne *ableitete* (TR_2).

- [TR_1a, DpV2_e, Klade] „*das ist der rechte Baum, nämlich, in dem Mipe, Kuje und Fale eine monophyletische Gruppe bilden (..)*
- [TR_2, DpV2_e, M. MRCA] *Dann hätte nämlich die Stammart, die zu Mipe, Kuje und Fale führt auch Schneidezähne gehabt (..)*
- [TR_2, DpV2_e, E/ V] *Und Schneidezähne wären bei Fale reduziert worden“*

Anschließend wendete sie die Hypothese zusätzlich auf DpV1 an und *leitete* eine konvergente Entstehung des Merkmals *ab* (TR_2). Abschließend wiederholte sie die Zuordnung zu DpV2:

- [TR_2, DpV1_v, Konv.] „*Der andere Baum würde für Kuje und Mipe quasi die Entwicklung von Schneidezähnen unabhängig voneinander sonst vorschlagen (4)*

[DpV2_e] *aber der rechte Baum unterstützt die Hypothese, dass die Schneidezähne in der Stammart vorhanden sind (..)“*

(HoEr P14, Alt1 methodischer Modelleinsatz)

4 Transkripte von NiEr wiesen keine eindeutige Lösung auf, 2 waren nicht kodierbar und 2 zeigten eine abweichende Lösung. 7 von 15 NiEr wendeten wie HoEr die Hypothese ebenfalls auf beide DpV an und ordneten sie dann *den wissenschaftlichen Konventionen entsprechend* dem korrekten DpV zu:

NiEr P16 wendete zunächst die Hypothese auf DpV1 an und *leitete* Entstehungs- bzw. Verlustereignisse *ab* (TR_2):

[TR_1b, DpV1_e, MRCA] *„[...] Dann sehe ich, dass da relativ viele Schritte zu Toge, Mipe, Pote und Kuje sind, also vom Vorfahren von Fale und den anderen vier eben genannten*

[TR_2, DpV1_e, M. MRCA] *Das heißt, der Vorfahre von Fale hätte Schneidezähne gehabt.*

[TR_2, DpV1_e, E/ V] *Fale selber hat sie im Verlauf der Evolution verloren.*

[TR_2, DpV1_e, M. MRCA] *Die gemeinsamen Vorfahren der vier hätten auch wieder Schneidezähne gehabt.*

[TR_1a, DpV1_e, Verwandtschaft] *Kuje und Mipe, die also keine direkte Verwandtschaft im linken Stammbaum besitzen.*

[TR_2, DpV1_e, M. MRCA] *die gemeinsamen Vorfahren der vier Kuje, Pote, Mipe und Toge müssten auch noch Schneidezähne haben,*

[TR_2, DpV1_e, M. MRCA] *weil der Vorfahre von Fale ja auch Schneidezähne hatte.*

[TR_2, DpV1_e, M. MRCA] *Dann die gemeinsamen Vorfahren von Kuje, Pote, Mipe und Toge also die beiden, die aus denen wiederum folgten, hätten auch noch Schneidezähne gehabt.*

- [TR_2, DpV1_e, E/ V] *Und dann hätten Toge und Pote die Schneidezähne wieder verloren.*
- [TR_2, DpV1_e, E/ V] *wir hätten insgesamt drei Merkmalsänderungen.
[Äh. Im linken Stammbaum]*
- [TR_2, DpV1_e, E/ V] *weil Fale ja verloren hat*
- [TR_1b, DpV1_e, MRCA] *und Toge*
- [TR_2, DpV1_e, E/ V] *und Pote auch.“*

Anschließend wendete sie die Hypothese auf DpV2 an, indem sie Entstehungs- bzw. Verlustereignisse *ableitete* (TR_2):

- [TR_1a, **DpV2_v**, MRCA] *„[Jetzt suche ich das gleiche im rechten Stammbaum.]
Da ist der Vorfahre von Fale auch der Vorfahre von Kuje,*
- [TR_2, DpV2_v, M. MRCA] *schließt aber auch nicht aus, dass der Vorfahre von Fale und Kuje und Mipe auch Schneidezähne hatten.
Die könnten also da auch schon vorhanden sein.*
- [TR_1b, DpV2_v, MRCA] *der Vorfahre von Mipe hatte Schneidezähne, hat die an Mipe weitergegeben und auch an den Vorfahren von Kuje und Fale*
- [TR_2, DpV2_v, E/ V] *und dann hat Fale im Verlauf der Evolution die Schneidezähne wieder verloren.*
- [TR_2, DpV2_v, E/ V] *wir hätten also nur eine Änderung,“*

Abschließend verglich sie die Anzahl der Entstehungs- bzw. Verlustereignisse zwischen beiden DpV und wählte DpV2 als passend zur Hypothese aus:

- [TR_2, **DpV1_v**, E/ V] *„im Vergleich zum linken Stammbaum, der drei Änderungen drinne hätte*
- [**DpV2**] *Dann äh. würde ich zur Unterstützung der Hypothese den rechten Stammbaum vorziehen.“*

(NiEr P16, Alt1 methodischer Modelleinsatz)

Im Gegensatz dazu leitete NiEr P17 zunächst das im Aufgabenstamm genannte Merkmal Schneidezähne als Merkmal des MRCA aller gezeigten Taxa in DpV1 ab (TR_2) und begründete das spätere Nicht-Vorhandensein über *wissenschaftlich inadäquates Ablesen* (TR_1b):

- [TR_2, DpV_un, M. MRCA] *„(..) Unter der Annahme, dass Kuje und Mipe Schneidezähne haben, müsste deren gemeinsamer Vorfahre ebenfalls Schneidezähne gehabt haben.
[Pause]*
- [TR_1b, DpV1_e, E/ V] *dass die sich aus irgendwelchen evolutiven Vorteilsgründen zurückgebildet haben,*
- [TR_0, DpV1_e, MRCA] *[würde zeigen, dass dann äh. der gemeinsame Vorfahre von Kuje und Mipe und Fale,] also auf der zweiten Ebene der Gliederung,*
- [T_1b, DpV1_e, MRCA] *welcher dann sich zur gleichen Zeit abgespalten hat wie Rine*
- [TR_2, DpV1_e, M. MRCA] *und Schneidezähne gehabt hat, sprich die Ausgangsspezies hat Schneidezähne gehabt. [Das würde sich mit der Darstellung auf der linken Seite konform gehen.]“*

Anschließend fokussierte sich NiEr P17 erneut auf das Merkmal des MRCA aller gezeigten Taxa und leitete ab (TR_2), dass dieser keine Schneidezähne hatte:

- [TR_1a, DpV2_v, Klade] *„[Bei der rechten Seite,] Kuje und Mipe, die hätten den gemeinsamen Vorfahren*
- [TR_1a, DpV2_v, Klade] *und Fale hat ebenfalls den gemeinsamen Vorfahren von Mipe und Kuje.*
- [TR_2, DpV2_v, M. MRCA] *Ich würde aber nicht schlussfolgern, dass auch die Ausgangsspezies dann die Schneidezähne hat,*
- [TR_1a, DpV2_e, Klade] *weil sich dann ja noch weitere Taxa abspalten - Rine,*
- [TR_1a, DpV2_e, Klade] *Toge*
- [TR_1a, DpV2_e, Klade] *und Pote*

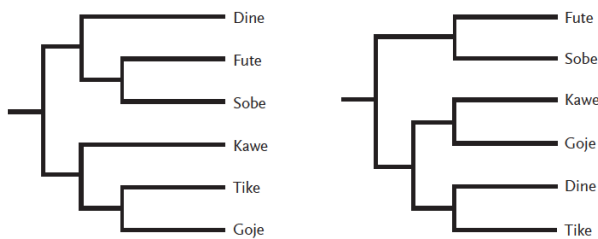
[nicht kodierbar]

da kann noch einiges evolutionär dazwischen passiert sein.“

(NiEr P17, *Alt1* methodischer Modelleinsatz)

6.8.6. Vorgegebene Hypothese einem DpV zuordnen und neue Hypothese generieren (Aufgabe *Alt2* methodischer Modelleinsatz)

In Aufgabe *Alt2* der Teilkompetenz *Alternative Modelle* im methodischen Modelleinsatz soll eine im Aufgabenstamm vorgegebene Hypothese zunächst einem der beiden DpV (DpV1) zugeordnet werden, um dann eine entsprechende Hypothese über das andere DpV aufzustellen (Klade enthält neben Kawe und Goje auch Tike; Abbildung 26).



Einer der Stammbäume beruht auf morphologischen, der andere auf genetischen Daten. Die morphologischen Daten deuten daraufhin, dass nur Kawe und Goje ein Monophylum mit ihrem letzten gemeinsamen Vorfahren bilden. Die genetischen Daten führen zu einer anderen Hypothese über dieses Monophylum.

Welche Hypothese über das Monophylum, das Kawe und Goje enthält, kann mithilfe der genetischen Daten aufgestellt werden? Begründen Sie Ihre Antwort.

Abbildung 26: Aufgabe *Alt2* methodischer Modelleinsatz: Vorgegebene Hypothese einem DpV zuordnen und neue Hypothese ableiten.

9 von 10 HoEr bearbeiteten diese Aufgabe *den wissenschaftlichen Konventionen entsprechend*, 6 nannten dabei 1 Hypothese, 3 HoEr 2 Hypothesen.

HoEr P14 ordnete die Hypothese DpV1 zu und las *wissenschaftlich adäquat* (TR_1a) die Klade *ab*, in der sich Kawe und Goje befinden. Sie erwähnte ausschließlich DpV1:

- [DpV1_e] „(..) Die genetischen Daten sind die auf der linken Seite (..)
- [TR_1a, DpV1_e, Klade] und über das gesamte Monophylum aus Kawe, Tike und Goje
- [TR_1a, DpV1_e, Klade] dass es als Schwestergruppe zu den drei anderen Taxa steht (..)
- [TR_1a, DpV1_e, Verw.] Tike und Goje näher miteinander verwandt sind (.)
- [TR_1a, DpV1_e, Klade] und Kawe die früh abzweigende Linie/“

(HoEr P14, *Alt2* methodischer Modelleinsatz)

12 von 15 NiEr lösten diese Aufgabe *entsprechend der wissenschaftlichen Konventionen*, die Transkripte von 3 NiEr enthielten keine eindeutige Lösung.

NiEr P16 löste gemäß der Instruktion und insgesamt sehr kurz durch *wissenschaftlich adäquates Ablesen* (TR_1a):

[TR_1a, DpV1_e, Klade] *„Aus dem Monophylum, das auch Kawe und Goje hervorbrachte kam auch das ja/ die Art Tike hervor,“*

(NiEr P16, Alt2 methodischer Modelleinsatz)

NiEr P15 löste die Aufgabe *gemäß der wissenschaftlichen Konventionen*, las dafür aber teilweise *wissenschaftlich inadäquat ab* (TR_1b):

[TR_1b, DpV_be, Merkmale] *„Kawe ist ziemlich ähnlich zu dem letzten gemeinsamen K/ Also in den beiden Stammbäumen zu den letzten gemeinsamen Vorfahren von Kawe und Goje (.)*

[TR_1b, DpV1_v, MRCA] *Gab es zwei Zwischenstationen, bis es zum gemeinsamen Vorfahren von Kawe und Goje kam, (.)*

[TR_1b, DpV2_v, MRCA] *und auf dem rechten Stammbaum sind das eins, zwei, drei, vier, also deutlich mehr (..) auf dem rechten sind das deutlich mehr evolutionäre Zwischenstufen*

[DpV1_e] *Ich gehe davon aus, dass links der genetische Stammbaum ist,*

[TR_1a, DpV1_e, Klade] *das ist gar kein Monophylum mehr, weil Tike auch noch dabei ist (.)*“

(NiEr P15, Alt2 methodischer Modelleinsatz)

6.9. Triangulation von verbalen und visuellen Daten

Wie in Kapitel 6.8.1 beschrieben, können sich Aussagen bei der Teilkompetenz *Alternative Modelle* im medialen Modelleinsatz beim Vergleichen auf ein DpV (DpV1 oder DpV2) oder auf beide DpV beziehen. Im Folgenden wird anhand von *Sequence Charts* für 3 Proband:innen (Abbildung 27 bis Abbildung 29) erläutert, inwiefern diese verbalen Daten mit den Fixationen auf die DpV in Verbindung stehen.

NiEr P13 verglich verbal bis Sekunde 32 beide DpV, indem sie die Aussagen explizit auf beide DpV bezog. Gleichzeitig fixierte NiEr P13 beide DpV abwechselnd (Abbildung 27, Kasten 1). In den folgenden Sekunden bezogen sich die Äußerungen von NiEr P13 erst auf DpV1 (bis Sekunde 37), dann auf DpV2 (bis Sekunde 45), jeweils bezugnehmend auf das jeweils andere DpV. In dieser Zeit wurden beide DpV abwechselnd fixiert, mit zunächst etwas mehr Fixationen bei DpV1 (Kasten 2) und dann DpV2 (Kasten 3). Abschließend sprach NiEr P13 vergleichend über beide DpV und fixierte ebenfalls beide (Kasten 4). Insgesamt verglich NiEr P13 verbal kontinuierlich die DpV, überwiegend bezogen sich die Aussagen dabei auf beide DpV vergleichend bzw. gleichzeitig. Dies spiegelt sich in den Fixationen wider, die permanent zwischen den DpV wechselten. Wenn NiEr P13 über ein DpV sprach, begannen die Fixationen auf diesem DpV bereits etwas früher als die verbale Sinneinheit (Abbildung 27).

HoEr P14 beschrieb die ersten 57 Sekunden ausschließlich DpV2, ohne sich auf DpV1 zu beziehen. Nach einer 5 Sekunden langen anfänglichen Phase, in der sie beide DpV abwechselnd fixierte (Abbildung 28, Kasten 1), fixierte sie anschließend ausschließlich DpV2 (Kasten 2). Anschließend bezog sich HoEr P14 verbal bis Sekunde 101 auf DpV1, dann bis Sekunde 124 auf DpV2, jeweils mit Bezug auf das jeweils andere DpV. Damit übereinstimmend fixierte sie zunächst fast ausschließlich DpV1 (Kasten 3), danach ausschließlich DpV2 (Kasten 4). Bis zum Ende des Transkripts bezog sich HoEr P14 auf beide DpV und fixierte sie immer abwechselnd (Kasten 5). Insgesamt stimmen verbaler Bezug und Fixationen bei HoEr P14 überein. Im Gegensatz zu NiEr P13 bezog sich HoEr P14 verbal nicht permanent auf beide DpV, sondern beschrieb die DpV separat. Die Fixationen blieben lange auf einem einzelnen DpV, es fanden wenige Fixationswechsel zwischen den DpV statt. Wie bei NiEr P13 begannen die Fixationen auf ein bestimmtes DpV etwas früher als die verbale Sinneinheit (Abbildung 28).

Verbal zeigte NiEr P17 die gleiche Bezugsreihenfolge wie NiEr P13 (beide DpV, DpV1 vergleichend, DpV2 vergleichend, beide DpV). Die Fixationen unterscheiden sich jedoch

jeweils in den Zeiträumen, in denen sich die Proband:innen auf ein DpV bezogen und mit einem anderen verglichen. NiEr P17 sprach bis Sekunde 20 über beide DpV und fixierte auch beide abwechselnd. Auch in den beiden Pausen vor und nach den Sinneinheiten fixierte sie beide abwechselnd (Abbildung 29, Kasten 1). Anschließend bezog sie sich bis Sekunde 43 auf DpV1, beziehend auf DpV2. Währenddessen fixierte sie ausschließlich DpV1 (Kasten 2). Bis Sekunde 59 bezogen sich die verbalen Sinneinheiten auf DpV2, vergleichen zu DpV1. Die Fixationen in diesem Zeitraum fanden fast ausschließlich beim rechten DpV statt (Kasten 3). Abschließend bezogen sich die Sinneinheiten auf beide DpV und auch die Fixationen wechselten zwischen beiden DpV hin und her (Kasten 4). Wie bei den anderen beiden Proband:innen begannen die Fixationen auf einem oder beiden DpV jeweils kurz bevor der verbale Bezug beginnt (Abbildung 29).

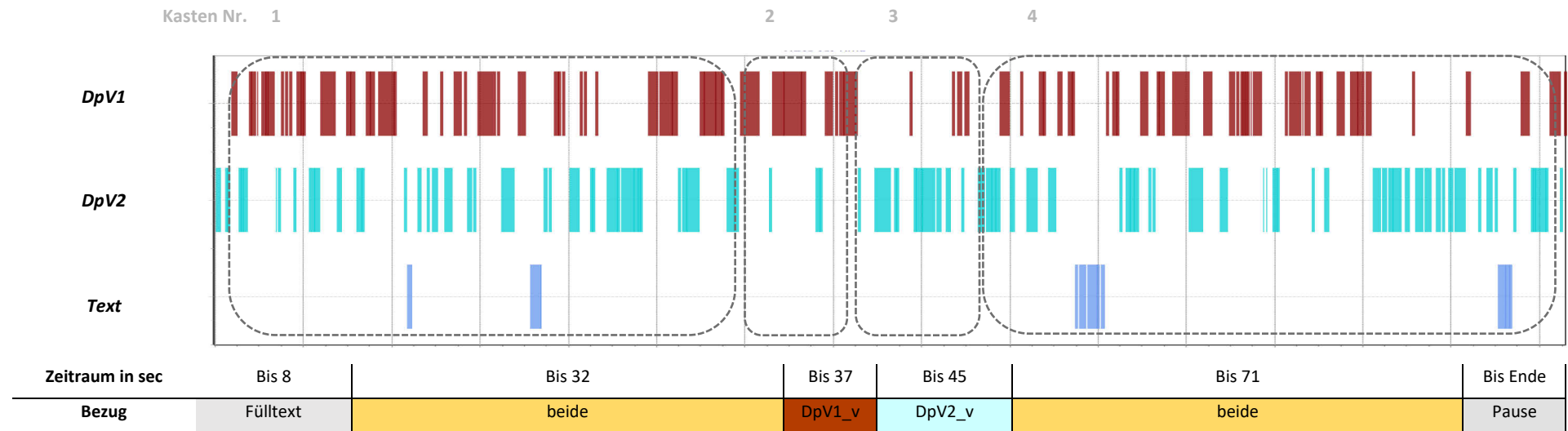


Abbildung 27: Triangulation von Fixationen und verbalen Bezügen von NiEr P13 in Aufgabe Alt1 im medialen Modelleinsatz. DpV1 = linkes DpV, DpV2 = rechtes DpV; ein Längsstrich entspricht 5 sec. Zeitraum in sec: Dauer der Sinneinheit in Sekunden, Bezug: Bezug der Sinneinheit. _v = vergleichend, _e = einzeln. Zahlen über den Kästen geben die jeweilige Kastennummer an.

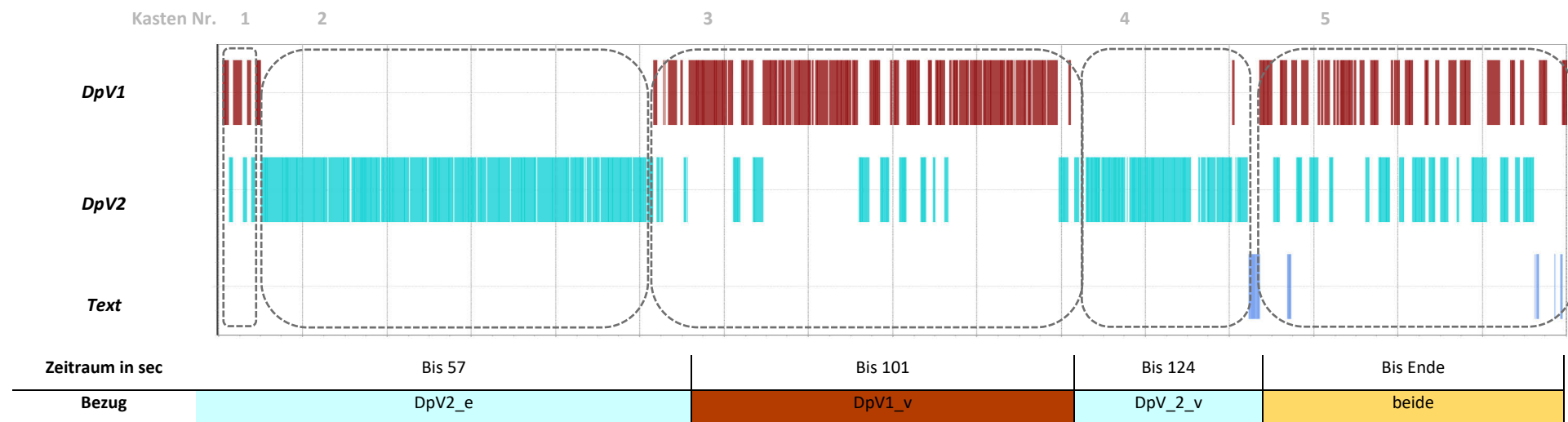
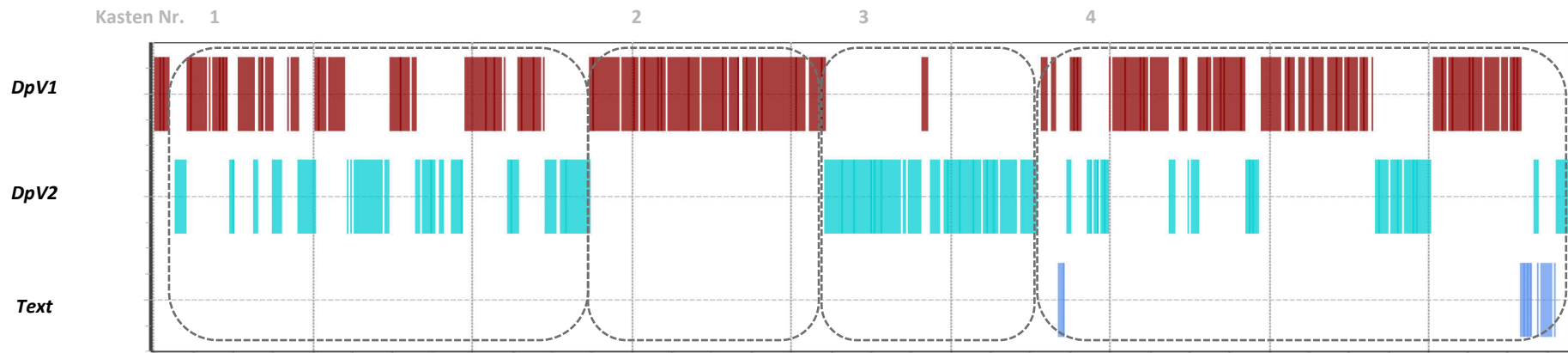


Abbildung 28: Triangulation von Fixationen und verbalen Bezügen von HoEr P14 in Aufgabe Alt1 im medialen Modelleinsatz. DpV1 = linkes DpV, DpV2 = rechtes DpV; ein Längsstrich entspricht 10 sec. Zeitraum in sec: Dauer der Sinneinheit in Sekunden, Bezug: Bezug der Sinneinheit. _v = vergleichend, _e = einzeln. Zahlen über den Kästen geben die jeweilige Kastennummer an.



| | | | | | | |
|------------------------|-------|--------|--------|--------|--------|----------|
| Zeitraum in sec | Bis 3 | Bis 20 | Bis 29 | Bis 43 | Bis 59 | Bis Ende |
| Bezug | Pause | beide | Pause | DpV1_v | DpV2_v | beide |

Abbildung 29: Triangulation von Fixationen und verbalen Bezügen von NiEr P17 in Aufgabe Alt1 im medialen Modelleinsatz. DpV1 = linkes DpV, DpV2 = rechtes DpV; ein Längsstrich entspricht 10 sec. Zeitraum in sec: Dauer der Sinneinheit in Sekunden, Bezug: Bezug der Sinneinheit. _v = vergleichend, _e = einzeln. Zahlen über den Kästen geben die jeweilige Kastennummer an.

6.10. Versionsunterschiede

Um eventuelle Effekte der Aufgabenreihenfolge detektieren zu können, wurden den Proband:innen innerhalb einer Erfahrungsgruppe abwechselnd Version 1 bzw. 2 zugeordnet (Kapitel 5.6). Die Versionen unterschieden sich in der Reihenfolge der Teilkompetenz und Aufgabe (Version 1: *Zweck1, Alt1, Zweck2, Alt2*; Version 2: *Alt2, Zweck2, Alt1, Zweck1*; jeweils Aufgabe zum Einarbeiten, Aufgabe medialer Modelleinsatz, Aufgabe methodischer Modelleinsatz direkt nacheinander). In Tabelle 16 sind die individuellen Lösungen nach Versuchsversion geordnet. Bei HoEr zeigten sich keine Unterschiede in den Lösungen zwischen den beiden Versionen. Bei NiEr kommen die abweichenden Lösungen im methodischen Modelleinsatz (PS, MRCA) ausschließlich bei Proband:innen mit Version 1 vor. Bei P17 und P19 kommt es zusätzlich sowohl bei der Aufgabe *Zweck2* als auch *Alt1* zu einem ähnlichen Lösungsansatz.

Tabelle 16: Individuelle Lösungen bzw. Lösungsprozesse geordnet nach Versuchsversion 1 (V1) und 2 (V2). Individuelle Lösungsprozesse bzw. Lösungen der Aufgaben. Zeilen entsprechen Proband:innen, Spalten entsprechen Aufgaben, geordnet nach Teilkompetenz und Modelleinsatz. Proband:innen sind anhand der von ihnen bearbeiteten Version 1 oder 2 geordnet. Zellenbeschriftung bei Abweichung der Lösung von wissenschaftlichen Konventionen bzw. maximaler Tree-Reading-Fähigkeit: n. k. = nicht kodierbar, Tree-Reading = Abweichende Tree-Reading-Fähigkeit, PS = Punkte und Streifen als Merkmal des MRCA, MRCA = MRCA hat Merkmal, späterer Verlust, Topo = Topologie.

| Aufgaben | | | | | | | | | | | |
|---------------------|---------------------|--|------------------------|--------|------|-------|----------------------------|----------|-----------|----------|--|
| Erfahr. | Abweichung | Proband:in | Medialer Modelleinsatz | | | | Methodischer Modelleinsatz | | | | |
| | | | Zweck1 | Zweck2 | Alt1 | Alt2 | Zweck1 | Zweck2 | Alt1 | Alt2 | |
| HoEr | ohne | P10, P28, P05, P24, P26, P06, P14, P23 | | | | | | | | | |
| | methodisch | P25 (Vers. 1) | | | | | | | and. | | |
| | | P29 (Vers. 2) | | | | | | | | and. | |
| NIer | Version 1 | | | | | | | | | | |
| | ohne | P07 | | | | | | | | | |
| | methodisch | P17 | | | | | PS | MRCA | MRCA | | |
| | | P19 | | | | | | MRCA | MRCA | | |
| | medial + methodisch | P12 | TR | | TR | | n. k. | TR, ohne | TR, teilw | TR, ohne | |
| | | P15 | TR | TR | TR | TR | n. k. | MRCA | | | |
| | | P09 | n. k. | n. k. | | | PS | MRCA | TR, and. | | |
| | | P21 | TR | | | | | n. k. | | ohne | |
| | | P27 | TR | | | | PS | MRCA | | | |
| | Version 2 | | | | | | | | | | |
| | ohne | P13 | | | | | | | | | |
| | medial | P18 | TR | | | | | | | | |
| | | P16 | | TR | | | | | | | |
| | methodisch | P20 | | | | | n. k. | | TR, ohne | | |
| P22 | | | | | | n. k. | | TR, Topo | ohne | | |
| medial + methodisch | P11 | TR | TR | | | n. k. | TR, ohne | n. k. | 1 | | |
| | P08 | TR | TR | | TR | | n. k. | n. k. | | | |

6.11. Prävalenz wissenschaftlich inadäquates Ablesen

Wie in Kapitel 6.5 beschrieben, kommt es je nach Aufgabe zu unterschiedlichen Anteilen von *wissenschaftlich inadäquatem Ablesen* (TR_1b). In Tabelle 17 sind die einzelnen Codehäufigkeiten von *wissenschaftlich inadäquatem Ablesen* (TR_1b) nach Kategorien geordnet. Bei Aufgabe *Alt1* im medialen Modelleinsatz kommt es besonders häufig zu den *wissenschaftlich inadäquaten Leseweisen* (TR_1b) *MRCA* zu zählen, *Taxa* in älter und jünger einzuteilen und Verwandtschaftsverhältnisse anhand alternativer Kriterien zu bestimmen.

Tabelle 17: Absolute Häufigkeiten von Sinneinheiten der Tree-Reading-Fähigkeit wissenschaftlich nicht adäquates Ablesen (TR_1b) pro Aufgabe für NiEr und ggf. HoEr. Anzahl NiEr / ggf. Anzahl HoEr (bei Einzelcodes nur angegeben wenn > 0). Unterstrichen, wenn Anteil der wiss. inadäquaten Sinneinheiten > 20 %. Keine Angabe wenn 0 in beiden Erfahrungsgruppen.

| Aufgabe | Medialer Modelleinsatz | | | | Methodischer Modelleinsatz | | | |
|--|------------------------|--------------|----------------------|----------------------|----------------------------|---------------|----------------------|----------------------|
| | Zw1 | Zw2 | Alt1 | Alt2 | Zw1 | Zw2 | Alt1 | Alt2 |
| Sinneinheiten gesamt | 54 / 39 | 39 / 30 | 173 / 107 | 148 / 116 | 64 / 120 | 94 / 134 | 92 / 106 | 71 / 56 |
| Sinneinheiten wiss. Inadäquat | 9 / 0 | 7 / 1 | <u>71 / 2</u> | <u>16 / 1</u> | 6 / 0 | 10 / 0 | <u>20 / 0</u> | <u>18 / 4</u> |
| MRCA gesamt | 9 / 0 | 6 / 1 | 33 / 0 | 12 / 0 | 1 / 0 | 2 / 0 | 10 / 0 | 6 / 0 |
| <i>MRCA zählen</i> | | 4 | 27 | 6 | | | 5 | 5 |
| <i>Nur MRCA für ein Taxon/ zwei Taxa</i> | 8 | 1 | | | 1 | 2 | 1 | |
| <i>gleiche Position = gleicher MRCA</i> | | | 3 | 1 | | | | 1 |
| <i>kein Knotenpunkt = keine weiteren Vorfahren</i> | 1 | 1 / 1 | 2 | | | | 3 | |
| <i>Outgroup = MRCA</i> | | | 1 | | | | | |
| <i>Knotenpunkte haben unterschiedliche Bedeutung</i> | | | 1 | 5 | | | 1 | |
| Merkmale gesamt | | | 2 / 0 | | | 4 / 0 | | |
| <i>Veränderung nur an Knotenpunkten</i> | | | 1 | | | | | |
| <i>Verlust bei Nichtgebrauch</i> | | | | | | 3 | 1 | |
| <i>Durchgehender Ast= keine Veränderung</i> | | | 1 | | | 1 | | |
| Verwandtschaft gesamt | | | 9 / 0 | 1 / 1 | | | 3 / 0 | 7 / 4 |
| <i>Räumliche Nähe</i> | | | 2 | | | | | |
| <i>Knotenzählen</i> | | | 3 | | | | 1 | |
| <i>Taxa stammen voneinander ab</i> | | | 1 | 1 / 1 | | | 1 | |
| <i>Similarity vs. Relatedness</i> | | | 3 | | | | 1 | 7 / 4 |
| Zeitverlauf gesamt | | | 2 / 1 | | | 1 / 0 | | |
| <i>Astlänge hat Bedeutung</i> | | | 1 | | | 1 | | |
| <i>Abzweigungsreihenfolge</i> | | | 1 / 1 | | | | | |
| Ältere und jüngere Taxa gesamt | | | 17 / 1 | 1 | 3 | 3 | | 3 |
| Sonstige gesamt | | 1 / 0 | 7 / 0 | 2 / 0 | 2 / 0 | | 6 / 0 | 2 / 0 |
| <i>Oberflächliche Topologie</i> | | | 6 | 1 | | | 6 | 2 |
| <i>Individualebene</i> | | | | | 1 | | | |
| <i>unklar</i> | | 1 | 1 | 1 | 1 | | | |

6.12. Kontrollvariablen

In Tabelle 18 wurden den individuellen Lösungsprozessen der Proband:innen die Kontrollvariablen-Werte zugeordnet. Beim *BETTSI*-Test, der Tree-Reading-Fähigkeiten misst, erreichten 5 HoEr einen Score von 100 %, 3 von 90,9 % und 2 von 72,7 %. Unabhängig vom *BETTSI*-Score lösten alle HoEr die Aufgaben mit *wissenschaftlich adäquatem Ablesen (TR_1a)* bzw. *den wissenschaftlichen Konventionen entsprechend*. Bei den NiEr erreichte 1 Proband:in 91 %, je 3 NiEr 82 bzw. 73 %, 6 je 64 % und 2 NiEr 55 % im *BETTSI*-Test. Im Vergleich zu HoEr erreichten also ca. die Hälfte der NiEr einen ähnlich hohen Score (> 70 %).

Betrachtet man die Einteilung in Gruppen anhand der Abweichungen der Lösungen, finden sich NiEr mit *BETTSI*-Scores > 70 % in allen 4 Gruppen, die größten Anteile jedoch in den Gruppen *ohne* Abweichung (2 von 2) oder Abweichungen in *methodischen* Aufgaben (3 von 4). In der Gruppe, deren Lösungen sowohl in den *medialen als auch den methodischen* Aufgaben abweicht, finden sich sowohl die 2 NiEr mit dem niedrigsten Score von 54,5 % als auch die NiEr mit dem höchsten *BETTSI*-Score von 90,9 %.

Der Großteil der HoEr löste beim Test zu dreidimensionalem Denken (3DW) mehr als 60 % der Aufgaben korrekt, 2 die Hälfte und 2 ein Drittel. Bei den NiEr lösten je 3 Proband:innen mehr als 60 % bzw. 50 % der Aufgaben, 9 NiEr ein Drittel. Diese 9 NiEr finden sich in allen Abweichungsgruppen.

Bei der Modellkompetenz variieren die Niveaus je nach Teilkompetenz: So weisen die meisten HoEr und NiEr bei der Teilkompetenz *Eigenschaften von Modellen* und *Ändern von Modellen* Niveau II auf, 4 bzw. 3 NiEr Niveau I. Bei der Teilkompetenz *Alternative Modelle* wurden 8 von 10 HoEr und 2 von 15 NiEr Niveau III zugeordnet, die restlichen Niveau II, 1 NiEr Niveau I. Bei der Teilkompetenz *Zweck von Modellen* wurden 5 von 10 HoEr und 4 von 15 NiEr Niveau III zugeordnet, 1 HoEr bzw. 4 NiEr Niveau II und 4 HoEr und 7 NiEr Niveau I. Bei der Teilkompetenz *Testen von Modellen* wurden 5 HoEr und 2 NiEr Niveau III zugeordnet, 1 HoEr und 5 NiEr Niveau II und 4 HoEr bzw. 8 NiEr der didaktischen Perspektive. Insgesamt lässt sich keine Häufung der Niveaus bei bestimmten Abweichungsgruppen finden.

Tabelle 18: Zuordnung der Kontrollvariablen zu individuellen Lösungsprozesse bzw. Lösungen. Zeilen entsprechen Proband:innen, Spalten entsprechen Aufgaben, geordnet nach Teilkompetenz und Modelleinsatz. Proband:innen sind in 4 Gruppen hinsichtlich ihrer Abweichungen geordnet: ohne = keine Abweichung von wissenschaftlichen Konventionen, medial = in mind. einer medialen Aufgabe, methodisch = in mind. einer methodischen Aufgabe, medial + methodisch = in mind. einer medialen und methodischen Aufgabe. **Kontrollvariablen:** BETTSI: korrekt gelöste Tree-Reading-Aufgaben in Prozent, 3DW: korrekt gelöste Aufgaben Dreidimensionales Denken in Prozent, Zellenfarbe: grün = > 70 %, gelb = < 70 %, ZVT: Zahlenverbindungstest Bearbeitungsdauer in Sekunden, grün = < 60 sec, gelb = > 60 sec., Modellkompetenz: grün = Niveau 1, gelb = Niveau 2, rot = Niveau 1, grau = didaktisch. **Aufgaben:** Die maximal erreichte Tree-Reading-Fähigkeit ist durch die Farbe der Zelle in der jeweiligen Aufgabe angegeben: blau = Ableiten (TR_2), grün = wissenschaftlich adäquates Ablesen (TR_1a), gelb = oberflächliches Ablesen (TR_0), orange = wissenschaftlich inadäquates Ablesen (TR_1b). Zellenbeschriftung im Falle einer Abweichung der Lösung von wissenschaftlichen Konventionen bzw. maximaler Tree-Reading-Fähigkeit: n. k. = nicht kodierbar, Tree-Reading = Abweichende Tree-Reading-Fähigkeit, and. = andere Bearbeitung, ohne = keine erkennbare Lösung, PS = Punkte und Streifen als Merkmal des MRCA, MRCA = MRCA hat Merkmal, späterer Verlust, Topo = Topologie.

| Erfahrungsgrad | Abweichung | Prob. | Kontrollvariablen | | | | | | | | Aufgaben | | | | | | | | |
|----------------|---------------------|-------|-------------------|---------------|----------|-----------------|---------|-------|--------|------|------------------------|---------|------|------|----------------------------|----------|-----------|----------|------|
| | | | BETTSI korrekt % | 3DW korrekt % | ZVT sec. | Modellkompetenz | | | | | Medialer Modelleinsatz | | | | Methodischer Modelleinsatz | | | | |
| | | | | | | Eigensch. | Alt. M. | Zweck | Testen | Änd. | Zweck 1 | Zweck 2 | Alt1 | Alt2 | Zweck1 | Zweck2 | Alt1 | Alt2 | |
| HoEr | ohne | P10 | 72,7 | 33,3 | 50,25 | 2 | 3 | 3 | 3 | 2 | | | | | | | | | |
| | | P28 | 72,7 | 33,3 | 68 | 2 | 2 | 3 | 3 | 2 | | | | | | | | | |
| | | P05 | 90,9 | 66,7 | 40,25 | 2 | 3 | 1 | did | 2 | | | | | | | | | |
| | | P24 | 90,9 | 100,0 | 58 | 2 | 3 | 1 | did | 2 | | | | | | | | | |
| | | P26 | 90,9 | 100,0 | 49,75 | 3 | 3 | 2 | did | 2 | | | | | | | | | |
| | | P06 | 100,0 | 91,7 | 40 | 2 | 3 | 3 | 2 | 3 | | | | | | | | | |
| | | P14 | 100,0 | 50,0 | 64,75 | 2 | 3 | 1 | did | 2 | | | | | | | | | |
| | | P23 | 100,0 | 66,7 | 65,5 | 2 | 3 | 3 | 3 | 2 | | | | | | | | | |
| | methodisch | P25 | 100,0 | 83,3 | 64 | 2 | 3 | 1 | 3 | 2 | | | | | | | | and. | |
| P29 | | 100,0 | 50,0 | 63,3 | 2 | 2 | 3 | 3 | 2 | | | | | | | | | and. | |
| NiEr | ohne | P07 | 81,8 | 66,7 | 62,5 | 2 | 2 | 3 | 3 | 2 | | | | | | | | | |
| | | P13 | 72,7 | 33,3 | 50,2 | 2 | 1 | 1 | 2 | 2 | | | | | | | | | |
| | medial | P18 | 63,6 | 25,0 | 48,3 | 2 | 2 | 2 | did | 2 | TR | | | | | | | | |
| | | P16 | 63,6 | 58,3 | 45,5 | 2 | 2 | 3 | 3 | 2 | | TR | | | | | | | |
| | methodisch | P17 | 81,8 | 33,3 | 59,5 | 1 | 2 | 1 | did | 2 | | | | | PS | MRCA | MRCA | | |
| | | P20 | 81,8 | 8,3 | 77,0 | - | 2 | 2 | did | 1 | | | | | n. k. | | TR, ohne | | |
| | | P22 | 72,7 | 33,3 | 56,5 | 2 | 2 | 1 | did | 1 | | | | | n. k. | | TR, Topo | ohne | |
| | | P19 | 63,6 | 50,0 | 68,7 | 2 | 3 | 2 | 2 | 1 | | | | | | | MRCA | MRCA | ohne |
| | medial + methodisch | P21 | 72,7 | 83,3 | 52,5 | 2 | 2 | 2 | did | 2 | TR | | | | | | n. k. | | ohne |
| | | P27 | 63,6 | 91,7 | 49,8 | 2 | 2 | 1 | did | 2 | TR | | | | PS | MRCA | | | |
| | | P09 | 63,6 | 33,3 | 61,0 | 2 | 3 | 3 | 2 | 2 | n. k. | n. k. | | | PS | MRCA | TR, and. | | |
| | | P11 | 54,5 | 25,0 | 69,8 | 1 | 2 | 1 | 2 | 2 | TR | TR | | | n. k. | TR, ohne | n. k. | 1 | |
| | | P08 | 63,6 | 58,3 | 49,0 | 1 | did | 3 | 2 | 2 | TR | TR | | TR | | n. k. | n. k. | | |
| | | P15 | 54,5 | 0,0 | 58,5 | 1 | 2 | 1 | did | 2 | TR | TR | TR | TR | n. k. | MRCA | | | |
| P12 | | 90,9 | 33,3 | 50,3 | 2 | 2 | 1 | did | 2 | TR | | | TR | | n. k. | TR, ohne | TR, teilw | TR, ohne | |

6.13. Zusammenfassung der Ergebnisse

Die Ergebnisse zum **medialen Modelleinsatz** lassen sich für die beiden Erfahrungsgruppen wie folgt zusammenfassen:

HoEr: Bei allen Aufgaben zu beiden Teilkompetenzen überwog allgemein das *wissenschaftlich adäquate Ablesen* (TR_1a). *Wissenschaftlich inadäquates Ablesen* (TR_1b) kam mit geringen Anteilen in 2 Aufgaben vor. Bei der Teilkompetenz *Zweck von Modellen* wurden bei beiden Aufgaben vorrangig *MRCA* beschrieben, dies erfolgte überwiegend durch *wissenschaftlich adäquates Ablesen* (TR_1a). Bei der Teilkompetenz *Alternative Modelle* wurden vorwiegend Kladen *wissenschaftlich adäquat abgelesen* (TR_1a). Dies geschah vor allem durch den Vergleich der beiden DpV. Alle HoEr erreichten bei allen Aufgaben die maximale Tree-Reading-Fähigkeit *wissenschaftlich adäquates Ablesen* (TR_1a).

NiEr: Unabhängig von der Teilkompetenz variierte allgemein der Anteil des *wissenschaftlich adäquaten Ablesens* (TR_1a) an allen Sinneinheiten je nach Aufgabe zwischen 36 % und 60 %. Parallel dazu kam es bei allen Aufgaben zum *oberflächlichen Ablesen* (TR_0) und mit 11 % bis 41 % zum *wissenschaftlich inadäquaten Ablesen* (TR_1b). Dieses wurde vor allem bei der Teilkompetenz *Alternative Modelle* vom Großteil der NiEr gezeigt. Bei der Teilkompetenz *Zweck von Modellen* wurden vor allem *MRCA* und Kladen mit *wissenschaftlich adäquatem* (TR_1a), *oberflächlichem* (TR_0) und *wissenschaftlich inadäquatem* Ablesen (TR_1b) beschrieben. Nur die Hälfte bzw. zwei Drittel der NiEr zeigten *wissenschaftlich adäquates* (TR_1a) als maximale Tree-Reading-Fähigkeit, die anderen *lasen maximal oberflächlich ab* (TR_0), vereinzelt kam *wissenschaftlich inadäquates Ablesen* (TR_1b) vor bzw. Transkripte konnten nicht kodiert werden. Bei der Teilkompetenz *Alternative Modelle* wurden je nach Aufgabe unterschiedliche DpV-Bestandteile genannt. Die beiden DpV wurden zum Großteil vergleichend beschrieben. Die maximale Tree-Reading-Fähigkeit lag bei fast allen NiEr bei *wissenschaftlich adäquatem Ablesen* (TR_1a).

Für den **methodischen Modelleinsatz** lassen sich die Ergebnisse für die beiden Erfahrungsgruppen wie folgt zusammenfassen:

HoEr: Bei 3 von 4 Aufgaben überwog allgemein das *Ableiten* (TR_2). Bei Aufgabe *Alt2* überwog *wissenschaftlich adäquates Ablesen* (TR_1a). *Wissenschaftlich inadäquates Ablesen* (TR_1b) kam nur in Aufgabe *Alt2* bei einigen HoEr vor. Bei der Teilkompetenz *Zweck von Modellen* und Aufgabe *Alt1* der Teilkompetenz *Alternative Modelle* wurden Entstehungs- und Verlustereignisse, Konvergenz, Merkmale von *MRCA* und das Sparsamkeitsprinzip abgeleitet. Bei der Aufgabe *Alt2* der Teilkompetenz *Alternative Modelle* wurden vorwiegend Kladen *wissenschaftlich adäquat* (TR_1a) beschrieben. Je nach Aufgabe der Teilkompetenz *Alternative Modelle* wurden die DpV überwiegend vergleichend (Aufgabe *Alt1*) bzw. einzeln erwähnt (Aufgabe *Alt2*). Die maximal erreichte Tree-Reading-Fähigkeit lag bei 3 Aufgaben bei allen HoEr bei *Ableiten* (TR_2). Bei Aufgabe *Alt2* liegt sie bei *wissenschaftlich adäquatem Ablesen* (TR_1a). Bei der Teilkompetenz *Zweck von Modellen* bearbeiteten alle HoEr die Instruktionen *den wissenschaftlichen Konventionen entsprechend*, bei der Teilkompetenz *Alternative Modelle* alle bis auf je 1 HoEr.

NiEr: Unabhängig von der Teilkompetenz variierte der allgemeine Anteil des *Ableitens* (TR_2) zwischen 34 % und 63,5 % der Sinneinheiten bei 3 der Aufgaben, zwischen 23 % und 45 % kam es entsprechend zu *wissenschaftlich adäquatem Ablesen* (TR_1a). Bei Aufgabe *Alt2* überwog das *wissenschaftlich adäquate Ablesen* (TR_1a). Bei allen 4 Aufgaben kam *wissenschaftlich inadäquates Ablesen* (TR_1b) vor, insbesondere bei der Teilkompetenz *Alternative Modelle*, bei der es von besonders vielen NiEr gezeigt wurde. Bei der Teilkompetenz *Zweck von Modellen* und einer Aufgabe der Teilkompetenz *Alternative Modelle* wurden vor allem Entstehungs- und Verlustereignisse und Merkmale von *MRCA* und Konvergenz *abgeleitet* (TR_2). Bei Aufgabe *Alt2* der Teilkompetenz *Alternative Modelle* wurden überwiegend Kladen und Verwandtschaftsverhältnisse *wissenschaftlich adäquat abgelesen* (TR_1a). Bei der Teilkompetenz *Alternative Modelle* wurden DpV in Aufgabe *Alt1* eher einzeln und in Aufgabe *Alt2* eher vergleichend erwähnt. Bei 3 Aufgaben erreichten je 9 bis 10 der NiEr die maximale Tree-Reading-Fähigkeit *Ableiten* (TR_2), die anderen erreichten *wissenschaftlich adäquates Ablesen* (TR_1a) bzw. die Transkripte waren nicht

kodierbar. Bei Aufgabe *Alt2* zeigten fast alle NiEr maximal *wissenschaftlich adäquates Ablesen* (TR_1a). 3 Aufgaben wurden von weniger als der Hälfte der NiEr *den wissenschaftlichen Konventionen entsprechend* bearbeitet. Viele Bearbeitungen waren nicht kodierbar bzw. *wichen von den wissenschaftlichen Konventionen ab*. Bei der Aufgabe *Alt2* bearbeiteten fast alle NiEr die Aufgaben *den wissenschaftlichen Konventionen entsprechend*.

Betrachtet man die **individuellen Lösungen**, konnten NiEr in 4 verschiedene Gruppen eingeteilt werden: Diejenigen, die aufgrund der maximalen Tree-Reading-Fähigkeit und der Aufgabenlösung über alle Aufgaben nicht von HoEr unterschieden werden können (*ohne*, 2 NiEr), diejenigen, die nur bei *medialen* Aufgaben (2 NiEr) bzw. nur bei *methodischen* Aufgaben (4 NiEr) abwichen und diejenigen, die *sowohl bei medialen als auch bei methodischen* Aufgaben abwichen (7 NiEr). Keiner NiEr wich bei allen Aufgaben ab. Es häuften sich bestimmte Abweichungen bei NiEr, die Versuchsversion 1 bearbeiteten.

Die **Triangulation** der visuellen und verbalen Daten für die Aufgabe *Alt1* im medialen Modelleinsatz zeigte eine Übereinstimmung der Fixationen und verbalen Bezügen.

7. Diskussion

Um zu diskutieren, inwiefern der Erfahrungsgrad mit DpV Tree-Reading-Prozesse beeinflusst (Forschungsfrage 1), werden zunächst die allgemeinen Codehäufigkeiten bzw. die maximalen Tree-Reading-Fähigkeiten und Lösungen der Aufgaben für HoEr und NiEr getrennt nach Modelleinsatz betrachtet. Anschließend werden die allgemeinen Codehäufigkeiten bzw. die maximalen Tree-Reading-Fähigkeiten und Lösungen der Aufgaben zwischen dem medialen und methodischen Modelleinsatz verglichen, um zu eruieren, inwiefern der Modelleinsatz Tree-Reading-Prozesse beeinflusst (Forschungsfrage 2). Abschließend werden diese Ergebnisse miteinander in Beziehung gesetzt, um die Interaktion zwischen Erfahrungsgrad und Modelleinsatz zu diskutieren (Forschungsfrage 3). Der interindividuelle Vergleich von Lösungsprozessen anhand verschiedener Fallbeispiele wird darauf aufbauend zur detaillierten Beschreibung von Tree-Reading-Prozessen und zur Identifizierung weiterer Einflussfaktoren diskutiert.

7.1. Diskussionsgrundlage

Zahlreiche Studien wenden die Eyetracking-Methode an, um Prozesse der Informationsverarbeitung oder Strategien bei der Aufgabenlösung zu untersuchen (vgl. Lai et al., 2013) sowie den Einfluss des Erfahrungsgrades auf die visuelle Wahrnehmung zu beschreiben (z. B. Ooms et al., 2012; Jarodzka et al., 2010). In den letzten Jahren hat die Methode vermehrt Einzug in die didaktische Forschung gehalten, indem z. B. Aufgabenlöseprozesse in der Mathematik-, Chemie- oder Physikdidaktik untersucht wurden (vgl. Klein et al., 2022). Als die vorliegende Studie geplant wurde, befand sich diese Entwicklung in der didaktischen Forschung noch in den Anfängen, die spezielle Betrachtung von visuellen Prozessen im Umgang mit DpV beschränkt sich bis heute auf eine Studie von Novick und Kolleg:innen (2012; vgl. Kapitel 3.2.3). Die vorliegende Studie hatte ursprünglich zum Ziel, in einem explorativen Ansatz erstmals kognitive Prozesse beim Tree-Reading zu untersuchen, indem visuelle Daten mit verbalen Daten flankiert werden. Entsprechend sind die Aufgaben für die Erfassung von visuellen Daten ausgelegt. Als methodisches Artefakt der gewählten Repräsentationsform erwiesen sich im Laufe der Datenanalyse die visuellen Daten als nicht eindeutig auswertbar (Kapitel 5.10), sodass sie nun in Form einer exemplarischen qualitativen Analyse die verbalen Daten stützen (Kapitel 6.9).

Die Charakterisierung der *Tree-Reading-Prozesse* von HoEr und NiEr im medialen und methodischen Modelleinsatz von DpV geschieht in der vorliegenden Studie folglich anhand der umfangreichen verbalen Daten. Zur Beurteilung von *Tree-Reading-Prozessen* wurden auf Grundlage des Kompetenzmodells zum Umgang mit Diagrammen (Nitz et al., 2018) übergeordnete *Tree-Reading-Fähigkeiten* nach Anspruch eingestuft (Kapitel 2.2). Hierbei wird angenommen, dass *Ableiten von Informationen* (TR_2) am anspruchsvollsten ist, dann inhaltliches, *wissenschaftlich adäquates Ablesen* (TR_1a) folgt und schließlich *oberflächliches Ablesen* (TR_0) den geringsten Anspruch hat, gemeinsam mit *wissenschaftlich inadäquates Ablesen* (TR_1b), da dieses auf fehlendem bzw. falschem Fachwissen bzw. alternativen Vorstellungen beruht. Basierend auf dem Kompetenzmodell der Modellkompetenz (Upmeier zu Belzen & Krüger, 2010) wurden für die Teilkompetenzen *Zweck von Modellen* und *Alternative Modelle* jeweils 2 Aufgaben mit dem Fokus auf dem medialen und methodischen Einsatz von Modellen generiert und untersucht.

7.2. Einfluss von Erfahrungsgrad und Modelleinsatz auf *Tree-Reading-Prozesse*

Betrachtet man die Ergebnisse hinsichtlich des Erfahrungsgrades, zeigt sich bei HoEr über alle Aufgaben hinweg ein homogenes Bild: HoEr *lesen* sowohl allgemein als auch maximal überwiegend *wissenschaftlich adäquat ab* (medialer Modelleinsatz) bzw. *leiten ab* (methodischer Modelleinsatz). Es kommen wenige bis keine *wissenschaftlich inadäquaten* Äußerungen vor und die Lösungsprozesse entsprechen den wissenschaftlichen Konventionen. Insgesamt lässt sich **Hypothese 1a** (Kapitel 4) also bekräftigen: ***Personen mit hohem Erfahrungsgrad lesen und leiten Informationen aus DpV wissenschaftlich adäquat ab.***

NiEr sind eine weitaus heterogenere Gruppe als HoEr, sowohl bezogen auf die gezeigten *Tree-Reading-Fähigkeiten* je Aufgabe als auch zwischen den Aufgaben. So kommen beim medialen Modelleinsatz *oberflächliches* bzw. *wissenschaftlich inadäquates Ablesen* je nach Aufgabe häufiger vor als *wissenschaftlich adäquates Ablesen*, sowohl bei den allgemeinen als auch den maximalen *Tree-Reading-Fähigkeiten*.

Beim methodischen Modelleinsatz gibt es weniger Variation als beim medialen Einsatz, jedoch sind die Lösungsprozesse von bis zu einem Drittel der NiEr aufgrund sprachlich bzw. inhaltlich unklarer Äußerungen nicht kodierbar und es kommt häufiger zu Abweichungen von den *wissenschaftlichen Konventionen*. Je nach Aufgabe zeigen fast alle NiEr *wissenschaftlich inadäquates Ablesen*. Es gibt jedoch auch mehrere NiEr, deren *Tree-Reading-Performanz*

bezogen auf die hier gewählten Kriterien nicht von HoEr zu unterscheiden ist. **Hypothese 1b**, die voraussagt, *dass Personen mit niedrigem Erfahrungsgrad häufig oberflächliche DpV-Merkmale nutzen und wissenschaftlich inadäquates Ablesen (TR_1b) beim Tree-Reading aufweisen*, wird also für einen Teil der NiEr gestützt.

Da die Aufgabenkonstruktion anhand des Kompetenzmodells der Modellkompetenz erfolgte (Upmeier zu Belzen & Krüger, 2010; Kapitel 5.4.2), definieren die verschiedenen Niveaustufen *a priori* die erforderlichen Tree-Reading-Fähigkeiten: Bei den medialen Aufgaben *wissenschaftlich adäquates Ablesen*, beim methodischen Modelleinsatz bei 3 Aufgaben *Ableiten*. Eine Ausnahme stellt die Aufgabe *Alt2* im methodischen Modelleinsatz dar, für die aufgrund der Struktur des Kompetenzmodells der Modellkompetenz maximal *wissenschaftlich adäquates Ablesen* benötigt wird. Entsprechend der intendierten Tree-Reading-Fähigkeiten zeigten HoEr beim medialen Modelleinsatz überwiegend *wissenschaftlich adäquates Ablesen*. Beim methodischen Modelleinsatz bearbeiteten HoEr 3 Aufgaben hauptsächlich durch *Ableiten* bzw. im Falle der Aufgabe *Alt2* durch *wissenschaftlich adäquates Ablesen*. **Hypothese 2: Der methodische Modelleinsatz erfordert elaboriertere Lösungsprozesse** wird durch die Lösungsprozesse der HoEr gestützt. Diese Anforderungen können jedoch in der vorliegenden Studie von NiEr nicht immer erfüllt werden, stattdessen bearbeiten einige die Aufgaben anhand anderer Tree-Reading-Fähigkeiten und weichen mit ihren Lösungsprozessen häufig von denen der HoEr ab.

7.3. Interaktion von Erfahrungsgrad und Modelleinsatz

Wie für Forschungsfrage 1 beschrieben, sind HoEr dazu in der Lage, unabhängig vom Modelleinsatz die wissenschaftlich erforderlichen Tree-Reading-Fähigkeiten einzusetzen und die Aufgaben *gemäß wissenschaftlicher Konventionen* zu bearbeiten. **Hypothese 3a - Personen mit hohem Erfahrungsgrad zeigen bei beiden Modelleinsätzen die wissenschaftlich erforderlichen Tree-Reading-Fähigkeiten** - wird also ebenfalls gestützt. HoEr können im Sinne der Expertiseforschung folglich *a posteriori* den Expert:innen zugeordnet werden, da sie eine „consistently superior performance on a specified set of representative tasks for a domain“ (Ericsson & Lehmann, 1996, S. 277) zeigen.

Bei NiEr ergibt sich im Vergleich zu HoEr eine komplexere Situation, da diese Gruppe unabhängig vom Modelleinsatz eine sehr heterogene Performanz zeigt. Diese manifestiert sich sowohl zwischen den Teilkompetenzen *Zweck von Modellen* und *Alternative Modelle*, als

auch innerhalb der Teilkompetenzen in den einzelnen Aufgaben. Nicht zuletzt unterscheiden sich die Individuen teilweise stark voneinander. **Hypothese 3b - Personen mit niedrigem Erfahrungsgrad zeigen bei methodischem Modelleinsatz eingeschränkte Tree-Reading-Fähigkeiten** - wird mit den Ergebnissen der vorliegenden Studie differenziert: Mehr als zwei Drittel der NiEr weichen in mindestens 2 methodischen Aufgaben von den den *wissenschaftlichen Konventionen* entsprechenden Lösungsprozessen ab, die Hälfte davon sogar sowohl bei den medialen als auch den methodischen Aufgaben. Dies stützt zunächst die Hypothese. 4 der NiEr sind jedoch in ihrer Tree-Reading-Performanz nicht oder nur minimal von HoEr bezüglich der untersuchten Kriterien zu unterscheiden und so in der vorliegenden Studie durchaus als Expert:in zu bezeichnen. Dies widerspricht der Arbeit von Halverson und Friedrichsen (2013), die ihr Expert:innenlevel als unerreichbar für Nicht-Wissenschaftler:innen definieren. Das *Revised STREAM* von Schramm und Kolleg:innen (2021a) wiederum lässt eine Entwicklung der Lernenden hin zum Expert:innenlevel zu, sofern alle 5 Dimensionen des *Revised STREAM* auf hoher Komplexität ausgeführt werden können (Kapitel 7.4).

7.4. Einordnung in Tree-Reading-Strukturierung

Die *a priori* durch Kriterien festgelegte Gruppe der NiEr zeigte in sich wesentlich heterogenere Lösungsprozesse als das in der Literatur gezeichnete überwiegend defizitäre Bild annehmen lässt (Kapitel 3). So lässt sich die Tree-Reading-Performanz von einigen NiEr nicht von der von HoEr unterscheiden, andere scheinen von Faktoren wie Modelleinsatz, Repräsentationsform, Aufgabenreihenfolge oder Fachbegriffen beeinflusst zu werden, die in den folgenden Kapiteln diskutiert werden. Insbesondere die NiEr, die in beiden Modelleinsätzen von den *wissenschaftlichen Konventionen* abwichen, lassen sich anhand der vorliegenden Kriterien jedoch nicht weiter gruppieren. Es scheinen also noch weitere individuelle Faktoren vorzuliegen, die die Tree-Reading-Performanz bei NiEr beeinflussen. Entsprechend scheint eine verallgemeinernde hierarchische Strukturierung von Tree-Reading-Fähigkeiten, wie sie von Halverson und Friedrichsen (2013) vorgeschlagen wurde, nicht anwendbar zu sein. Vielmehr kann das *Revised STREAM* (Schramm et al., 2021a; Kapitel 2.2, Abbildung 3) zur Erklärung herangezogen werden, das zu folgendem Schluss kommt: „*students show varying degrees of skill in different skill dimensions*“ (Schramm et al., 2021a, S. 9). So kann die individuelle Tree-Reading-Performanz von NiEr sehr unterschiedlich ausfallen, je nachdem welche Dimension in einer Aufgabe erforderlich ist.

Die Aufgaben in der vorliegenden Studie wurden mit dem Fokus auf das Kompetenzmodell der Modellkompetenz (Upmeyer zu Belzen & Krüger, 2010) entwickelt, bevor das *Revised STREAM* veröffentlicht wurde (Schramm et al., 2021a). Anhand der Zuordnung der *ETCI Learning Outcomes* zu den *Revised STREAM* Dimensionen bzw. Leveln (vgl. Schramm et al., 2021a; Tabelle 2) können die Aufgaben der vorliegenden Studie ins *Revised STREAM* annäherungsweise eingeordnet werden. Die zunehmende Komplexität der Dimensionen wird nicht näher erläutert (Schramm et al., 2021a), sodass diese für die Aufgaben der vorliegenden Studie anhand der erforderlichen Tree-Reading-Fähigkeiten antizipiert wird. Es ergibt sich folgende Verteilung (Abbildung 30):

Den **MRCA identifizieren und wissenschaftlich adäquat beschreiben** (Aufgabe *Zweck1* und *Zweck2* medialer Modelleinsatz) wird der Dimension *Determining Monophyletic Groups* im weniger komplexen Bereich des *Learner's Level* zugeordnet, **2 DpV miteinander zu vergleichen** (Aufgabe *Alt1* und *Alt2* medialer Modelleinsatz) der Dimension *Comparing Trees* ebenfalls im weniger komplexen Bereich. Soll eine Merkmalsentstehung anhand von **Konvergenz abgeleitet** werden (Aufgabe *Zweck1* methodischer Modelleinsatz) bzw. das **Merkmal des MRCA** aller gezeigten Taxa anhand des Sparsamkeitsprinzips **abgeleitet** werden (Aufgabe *Zweck2* methodischer Modelleinsatz), entspricht dies *Handling Apomorphies* auf hohem *Learner's Level*. In der vorliegenden Studie ist aufgrund des Versuchssettings auch schon auf diesem Level mentales Modellieren nötig, um verschiedene Szenarien miteinander zu vergleichen, was normalerweise dem *Expert's Level* zugeordnet werden würde (Schramm et al., 2019; Schramm et al., 2021a).

Das zu einer Hypothese passende DpV identifizieren (Aufgabe *Alt1* methodischer Modelleinsatz) ist in 2 Schritten einzuordnen, da es mehrere Komponenten enthält. So müssen zunächst die **Entstehungs- bzw. Verlustereignisse** (E/ V-Ereignisse) des Merkmals in beiden DpV abgeleitet werden, dies entspricht *Handling Apomorphies* auf einem sehr hohem *Learner's Level*. Anschließend wird anhand von *Comparing Trees* ebenfalls auf einem sehr hohem *Learner's Level* das **DpV identifiziert, dass die sparsamste** Erklärung bietet (Aufgabe *Alt1* methodischer Modelleinsatz). Eine genannte **Klade einem DpV zuzuordnen** und anschließend eine weitere Klade im anderen DpV zu **identifizieren** (Aufgabe *Alt2* methodischer Modelleinsatz) kann der Dimension *Determining Monophyletic Groups* auf einem weniger komplexem *Learner's Level* zugeordnet werden.

Da auch **beide DpV dabei verglichen** werden, entspricht dies auch der Dimension *Comparing Trees* auf einem weniger komplexem *Learner's Level*.

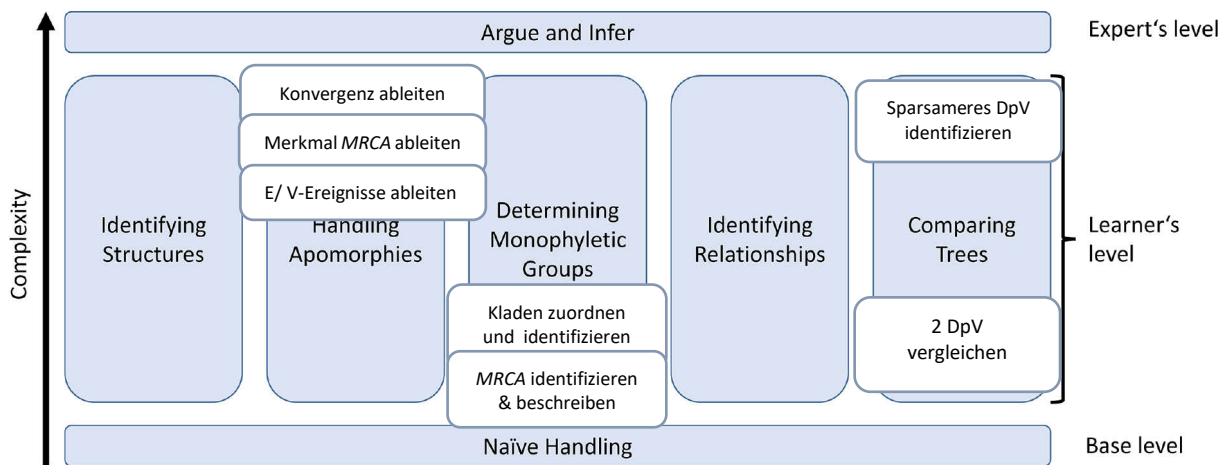


Abbildung 30: Zuordnung der in den Aufgaben erforderlichen Tree-Reading-Fähigkeiten (weiße Kästen) ins Revised Synthetic Tree-Reading Model (STREAM). In Anlehnung an Schramm et al. (2021a, S. 9). E/ V-Ereignisse = Entstehungs- und Verlustereignisse.

Alle NiEr konnten per Definition bereits in der Schule oder Universität zumindest grundlegende Kenntnisse in Tree-Reading erwerben. Es ist also davon auszugehen, dass sie dem *Learner's Level* zuzuordnen sind, sich jedoch je nach Dimension auf unterschiedlichem Niveau befinden. Einige NiEr sind jedoch ausgehend von den erfassten Kriterien unabhängig vom Modelleinsatz nicht von HoEr zu unterscheiden, befinden sich also vermutlich auf hohem Niveau in den von den Aufgaben adressierten Dimensionen. Die NiEr, die nur in einzelnen Aufgaben mit methodischem Modelleinsatz von den *wissenschaftlichen Konventionen abweichen*, sind möglicherweise im *Learner's Level* je nach Dimension zu unterschiedlich komplexer Tree-Reading-Performanz fähig. NiEr, die sowohl im medialen als auch im methodischen Modelleinsatz *wissenschaftlich inadäquate* bzw. von den *wissenschaftlichen Konventionen abweichende* Tree-Reading-Prozesse zeigen, befinden sich vermutlich in mehreren Dimensionen auf einem weniger komplexen Level und sind entsprechend zu weniger anspruchsvollen Tree-Reading-Prozessen in der Lage.

7.6. Einflussfaktoren jenseits von Erfahrungsgrad und Modelleinsatz

Wie im vorherigen Kapitel erläutert, unterstützen die heterogenen interindividuellen Lösungsprozesse von NiEr in der vorliegenden Studie die Annahme, dass Tree-Reading-Fähigkeiten je nach Dimension unterschiedlich gut beherrscht werden (Schramm et al., 2021a). Beim interindividuellen Vergleich der Lösungsprozesse wird deutlich, dass bei einigen Aufgaben bestimmte Lösungsvarianten insbesondere bei NiEr gehäuft vorkommen (Kapitel 6.8). Ordnet man die Proband:innen zusätzlich nach der Versuchsversion, die sie bearbeitet haben, zeigen sich Häufungen bestimmter Lösungsvarianten in einer der Versuchsversionen (Kapitel 6.10). Es ist also davon auszugehen, dass es weitere Faktoren jenseits von Erfahrungsgrad, Modelleinsatz und Tree-Reading-Dimensionen gibt. Im Folgenden wird diskutiert, inwiefern Merkmale der Repräsentationen, die Aufgabenreihenfolge sowie Fachbegriffe und Fachsprache diese Häufungen erklären können und inwiefern das Versuchssetting einen Einfluss auf die Ergebnisse genommen haben könnte.

7.6.1. Merkmale der Repräsentationen

In den medialen Aufgaben zur Teilkompetenz *Zweck von Modellen* soll der *MRCA* von 2 Taxa beschrieben werden. Dieser liegt bei Aufgabe *Zweck1* oberflächlich betrachtet im mittleren Bereich des DpV, während er bei Aufgabe *Zweck2* oberflächlich gesehen der am weitesten links gelegene Knotenpunkt ist, an dem sich Outgroup und übrige Taxa voneinander abspalten. Die Position des *MRCA* scheint die Art der Beschreibung zu beeinflussen, insbesondere bei NiEr. So beschreiben bei der Aufgabe *Zweck1* die Hälfte der NiEr ausschließlich *oberflächlich* bzw. *wissenschaftlich inadäquat*, indem sie meist die Knotenpunkte von links nach rechts abzählen. Möglicherweise sind diese NiEr durch das allgemeine Verständnis von Hierarchien dazu fähig, den *MRCA* korrekt zu identifizieren, ohne das tiefere Verständnis der zugrundeliegenden Prozesse und Theorien zu benötigen (Novick & Catley, 2007). Außerdem kann auch die Fachsprache eine Herausforderung für NiEr sein, wenn sie die Position beschreiben sollen, sodass sie auf *oberflächliche* Beschreibungen zurückgreifen.

Eine Kombination aus *oberflächlichem* Abzählen und *wissenschaftlich adäquaten* Beschreibungen kommt auch bei anderen NiEr und auch bei HoEr vor, möglicherweise mit der Intention, die Aufgabe möglichst ausführlich und umfassend zu bearbeiten. Die Position des *MRCA* scheint also unabhängig vom Erfahrungsgrad eine *oberflächliche* Beschreibung hervorzurufen. Liegt der Knotenpunkt wie bei Aufgabe *Zweck2* nicht inmitten des DpV, wird

die Vorgehensweise des Abzählens nicht angewendet sondern eher *wissenschaftlich adäquat abgelesen*.

Sowohl HoEr als auch NiEr bearbeiteten die Aufgabe vor allem durch die Nennung von *MRCA*, deutlich weniger durch das Beschreiben von Kladen, die durch einen *MRCA* begründet werden, obwohl dies sinnvoll für den Vergleich von DpV ist (Baum & Smith, 2013). Werden Kladen beschrieben, erfolgt dies bei HoEr ausschließlich *wissenschaftlich adäquat*, während NiEr hauptsächlich Kladen *oberflächlich ablesen* und vor allem die Taxa aus dem Aufgabenstamm nennen. Auch hier ist es möglich, dass bei NiEr ein tieferes Verständnis der Bedeutungen von Kladen fehlt (Catley et al., 2012) und aufgrund von oberflächlichen Hierarchien argumentiert wird.

Das Vergleichen von 2 DpV ist der Teilkompetenz *Alternative Modelle* im medialen Modelleinsatz zuzuordnen und erfordert ebenfalls *wissenschaftlich adäquates Ablesen*. Im Gegensatz zu den im vorherigen Absatz beschriebenen Aufgaben zu der Teilkompetenz *Zweck von Modellen* sind die Aufgaben zum Vergleichen von DpV offen formuliert, sodass die Proband:innen die Art und Anzahl ihrer Vergleichskriterien frei wählen können. Daraus resultieren bei beiden Gruppen mindestens dreimal mehr Sinneinheiten als bei der Teilkompetenz *Zweck von Modellen*.

Insgesamt findet bei beiden Aufgaben sowohl bei HoEr als auch bei NiEr überwiegend ein Vergleich zwischen den DpV statt, indem bspw. die Beziehung von Taxa zueinander erst in einem und dann im anderen DpV beschrieben wird. Eher seltener wird nur ein einzelnes DpV ohne Bezug zum anderen beschrieben.

Auch beim Vergleichen zweier DpV ist bei NiEr ein Unterschied zwischen den Aufgaben zu erkennen. Während bei Aufgabe *Alt2* NiEr im ähnlichen Umfang wie HoEr vor allem Kladen verglichen und diese auch zum Großteil *wissenschaftlich adäquat ablesen*, kam es bei der Aufgabe *Alt1* bei NiEr deutlich häufiger zu *wissenschaftlich inadäquatem Ablesen*. Dies ist wahrscheinlich auf die unterschiedliche oberflächliche Topologie in den Aufgaben zurückzuführen. Bei Aufgabe *Alt1* (Abbildung 21) sind beide DpV unbalanciert, da die Entwicklungslinie, die vom ältesten *MRCA* abgeht, speziesärmer ist als die andere (Gregory, 2008). Die DpV suggerieren eine oberflächliche Kontinuität, die aufgrund der Baumform wie eine nach oben bzw. nach unten führende Treppe aussieht. Die unbalancierte Repräsentationsform kann eine Erklärung für den in dieser Aufgabe bei NiEr gehäuft vorkommenden *wissenschaftlich inadäquaten* Vergleich der Anzahl von *MRCA* sein. Dieser

Effekt kann vermutlich abgeschwächt bzw. vermieden werden, indem die unbalancierten DpV an den Knotenpunkten so gedreht werden, dass sie nicht mehr wie Treppen aussehen und keine Kontinuität mehr suggerieren wie bei Aufgabe *Alt2* (Abbildung 22). Ein ähnlicher Effekt wurde bei der Interpretation einzelner leiterförmiger DpV beschrieben. Hier erzeugen unbalancierte und unrotierte DpV bei Studierenden den Eindruck einer *wissenschaftlich inadäquaten* durchgehenden Linie bzw. eines Hauptastes. Dies wird auf das Gestaltgesetz der guten Fortsetzung zurückgeführt (Novick & Catley, 2007; Novick, Shade et al., 2011; Novick, Pickering et al., 2014). Nach Kenntnisstand der Autorin der vorliegenden Studie wurde der hier gezeigte ähnliche Effekt beim Vergleichen baumförmiger DpV bisher noch nicht beschrieben. Die Aussage, dass das *wissenschaftlich adäquate* Vergleichen von DpV anspruchsvoller ist als das Beschreiben eines einzelnen DpV (Halverson, 2011; Halverson & Friedrichsen, 2013; Novick & Catley, 2016; Schramm et al., 2019; Schramm, 2021a) und einen hohen Grad an Erfahrung mit DpV erfordert (Halverson & Friedrichsen, 2013), scheint also auch hier repräsentationsabhängig zu sein. Während HoEr unabhängig von Merkmalen der Repräsentation *wissenschaftlich adäquat ablesend* vergleichen (Halverson & Friedrichsen, 2013; Kummer et al., 2019; Schramm et al., 2019), können NiEr von bestimmten Repräsentationsmerkmalen beeinflusst werden (z. B. Novick & Catley, 2007; Novick et al., 2011; Novick, Pickering et al., 2014).

7.6.2. Aufgabenreihenfolge

Bei den methodischen Aufgaben zur Teilkompetenz *Zweck von Modellen* sind beim *Ableiten* von Merkmalen einige NiEr in ihrer Tree-Reading-Performanz nicht von HoEr zu unterscheiden. Sie *leiten* wie HoEr Entstehungs- und Verlustereignisse korrekt *ab* und erkennen eine konvergente Merkmalsentstehung. Außerdem werden von beiden Erfahrungsgruppen teilweise mehrere Hypothesen genannt und anhand des Sparsamkeitsprinzips gegeneinander abgewogen.

Sortiert man die individuellen Lösungen nach der Versuchsversion, also der Reihenfolge, in der die Aufgaben bearbeitet wurden (Tabelle 16), fällt auf, dass bestimmte *nicht den wissenschaftlichen Konventionen entsprechende* Lösungen der methodischen Aufgaben nur bei NiEr mit Version 1 mit der Aufgabenreihenfolge *Zweck1, Alt1, Zweck2, Alt2* auftreten. 6 von 8 NiEr mit dieser Version weichen in mindestens 2 Aufgaben mit methodischem Modelleinsatz ab, davon 3 NiEr bei allen oder fast allen Aufgaben. In Aufgabe *Zweck1* soll das

Merkmal des *MRCA* aller gezeigten Taxa abgeleitet werden. Kommt diese Aufgabe zuerst, scheint der Fokus für die folgenden Aufgaben im methodischen Modelleinsatz auf dem *MRCA* aller Taxa zu liegen. So begründen NiEr entgegen sparsamerer Szenarien Merkmale bei rezenten Taxa damit, dass sie bereits beim *MRCA* aller Taxa vorlagen und es später ggf. zu Verlustereignissen kam. Insbesondere NiEr neigen zu einem negativen analogen Transfer von Problemlösungen, wenn sich Aufgaben oberflächlich ähneln (Novick, 1988). Es ist also möglich, dass NiEr vom Fokus der Aufgabe *Zweck1* auf den *MRCA* aller Taxa beeinflusst werden. Gegebenenfalls ist ihnen das Prinzip konvergenter Merkmalsentstehung nicht bekannt oder sie beherrschen es nicht (vgl. Morabito et al., 2010), sodass sie auf die ihnen bereits bekannte Bearbeitung der vorausgehenden Aufgabe *Zweck1* zurückgreifen.

In Version 2 mit der Aufgabenreihenfolge *Alt2*, *Zweck2*, *Alt1*, *Zweck1* kommen diese abweichenden Lösungen zwar nicht vor, dafür lösen 3 von 7 NiEr sowohl die Aufgabe *Zweck1* als auch *Alt1* im methodischen Modelleinsatz nicht. Unabhängig von der Versuchsversion scheinen diese Aufgaben für viele NiEr schwer oder gar nicht lösbar zu sein. Möglicherweise ist dies auf den Begriff „Sparsamkeitsprinzip“ zurückzuführen, der in diesen beiden Aufgaben als Kriterium für die Ermittlung des Merkmals genannt wird, dies wurde von manchen NiEr während des Lösungsprozesses geäußert.

7.6.3. Fachsprache und Fachwissen

Die Evolutionsbiologie umfasst eine Vielzahl von fachlichen Konzepten, die auch von vielen Biologielehrkräften oft nicht umfassend verstanden werden (van Dijk & Kattmann, 2010; van Dijk & Reydon, 2010). So wird Lehrkräften hinsichtlich der Evolutionsbiologie teils nicht ausreichendes *Professional Content Knowledge* (PCK) attestiert, um auf *wissenschaftlich inadäquate* Aussagen von Schüler:innen einzugehen (van Dijk & Kattmann, 2010). Erschwerend kommt hinzu, dass Begriffe wie Stammbaum, Vorfahre oder Verwandtschaft(sgrad) in der Alltagssprache und im Unterricht ebenfalls für Verwandtschaft von Individuen verwendet werden (Kapitel 2.1) und so die Konzepte der evolutiven und individuellen Verwandtschaft sprachlich sehr ähnlich, jedoch inhaltlich sehr unterschiedlich sind. In der vorliegenden Arbeit kamen bei NiEr vereinzelt Äußerungen auf Individualebene vor (NiEr P17 beim methodischen Modelleinsatz der Teilkompetenz *Zweck von Modellen*, Kapitel 6.8.3), was für eine Vermischung bzw. Nicht-Unterscheidung der Konzepte sprechen kann.

Das Kategoriensystem, das in der vorliegenden Studie genutzt wurde, war so ausgelegt, dass Aussagen auch als *wissenschaftlich adäquat* bzw. *den wissenschaftlichen Konventionen entsprechend* kodiert wurden, wenn keine Fachbegriffe genutzt, sondern alltagssprachliche Formulierungen verwendet wurden (z. B. Gruppe statt Klade oder Monophylum).²² Da davon auszugehen ist, dass NiEr die Terminologie des Tree-Reading weniger gut beherrschen als HoEr, liegt der Fokus auf dem Inhalt einer Aussage und nicht auf der Fachsprache selbst.²³ Wie im vorherigen Kapitel (7.5.2) erläutert, scheint insbesondere das Sparsamkeitsprinzip für NiEr problematisch zu sein. Obwohl dieser Begriff vor dem Start der Versuche schriftlich erläutert wurde, ist es möglich, dass NiEr dieses zentrale Prinzip nicht oder nicht gut genug kennen oder nicht von der Kurzeinführung vor Versuchsbeginn profitieren konnten. NiEr haben also einerseits das Sparsamkeitsprinzip nicht angewendet, da es ihnen zum Befragungszeitpunkt nicht geläufig war. Andererseits waren sie möglicherweise von dem ihnen nicht bekannten Begriff irritiert und so in ihrem Lösungsprozess beeinträchtigt. Es kann nicht abschließend geklärt werden, inwiefern dies die Lösungsprozesse beeinflusst hat und ob die NiEr die Aufgaben *den wissenschaftlichen Konventionen entsprechend* bearbeitet hätten, wenn der Begriff Sparsamkeitsprinzip nicht explizit genannt worden wäre.

Fachsprache beinhaltet nicht nur verbale Äußerungen, sondern wird umfassender als eine Kombination aus verschiedenen Repräsentationen gesehen, die allesamt zur Kommunikation von und über Fachinhalte genutzt werden (vgl. Nitz, 2012, S. 40). Bezogen auf die vorliegende Arbeit kann also geschlossen werden, dass sowohl DpV als auch Tree-Reading als Interpretation und Vergleich verschiedener DpV gleichermaßen als Teil von Fachsprache bezeichnet werden können. Es ergibt sich folglich für NiEr der doppelte Anspruch, stark abstrahierte DpV, die zur fachsprachlichen Kommunikation von Hypothesen dienen, in verbale Kommunikation zu übersetzen.

²² Eine klare Abgrenzung zwischen Fach- und Alltagssprache ist jedoch ohnehin kaum möglich (vgl. Rincke & Markic, 2018, S. 33; Nitz, 2012, S. 4 f). In der vorliegenden Arbeit werden die in Abschnitt 2.1.1 genannten Begriffe als Fachbegriffe definiert.

²³ Auch Fachexpert:innen unterhalten sich nicht durchgehend mit Fachtermini (vgl. Rincke & Markic, 2018), sodass eine Kodierung ausschließlich anhand der Verwendung von Fachtermini realitätsfern wäre.

7.8. Visuelle Daten

In der vorliegenden Studie bieten die Aufgaben der Teilkompetenz *Alternative Modelle* die Möglichkeit, verbale und visuelle Daten beim Vergleichen zweier DpV zu triangulieren. So kann untersucht werden, inwiefern verbale und visuelle Daten je nach Erfahrungsgrad korrelieren und inwiefern beide DpV fixiert werden. Es zeigt sich exemplarisch, dass sowohl bei HoEr als auch bei NiEr die Fixationen mit den verbalen Daten übereinstimmen. Wird ein DpV einzeln oder vergleichend beschrieben, wird dieses auch überwiegend fixiert. Werden beide DpV anhand verschiedener Kriterien direkt miteinander verglichen, wechseln auch die Fixierungen zwischen beiden hin und her. Dies stützt sowohl die *Eye-Mind-Assumption* als auch die *Immediacy Assumption* (Just & Carpenter, 1980), da die Proband:innen die fixierten Bereiche auch unmittelbar verbal beschreiben. So werden die kognitiven Prozesse sich gegenseitig ergänzend abgebildet (Stieff et al., 2011). Die Komplexität der DpV und der damit verbundene vermutlich erhöhte *Cognitive Load* führen vermutlich sowohl bei HoEr als auch bei NiEr dazu, dass verbale und visuelle Daten eng verbunden sind (Rayner, 1998).

7.9. Kontrollvariablen

Hinsichtlich der Kontrollvariable Tree-Reading-Fähigkeiten (BETTSI) schneiden HoEr wie erwartet signifikant besser ab als NiEr (Tabelle 9) und beantworten mehr als 70 % der Fragen korrekt. Bei NiEr variieren die getesteten Tree-Reading-Fähigkeiten stark, teilweise werden nur etwa die Hälfte der Fragen korrekt beantwortet. Trägt man den *BETTSI*-Score gegen die Einteilung in Untergruppen nach Lösungsprozess auf (Tabelle 18), haben NiEr, die nicht oder nur bei Aufgaben im methodischen Modelleinsatz abweichen, bis auf eine Ausnahme mit HoEr vergleichbare Scores. NiEr, die in Aufgaben im medialen oder methodischen Modelleinsatz abweichen, weisen hingegen meist niedrigere Scores unter 70 % auf. Es kann angenommen werden, dass der *BETTSI*-Score ein Prädiktor für das Abschneiden in Aufgaben mit medialem Modelleinsatz ist, jedoch nicht in den Aufgaben mit methodischem Modelleinsatz. Eine Ausnahme bildet NiEr P12, die mit 91 % den höchsten *BETTSI*-Score aller NiEr hat, jedoch die meisten Abweichungen in allen Aufgaben. Die Gründe hierfür sind *a posteriori* nicht ermittelbar.

Das dreidimensionale Denken ist bei HoEr signifikant stärker ausgeprägt als bei NiEr, allerdings erreichen HoEr im Schnitt maximal zwei Drittel der Punkte (3DW, Tabelle 9), sodass der Großteil aller Proband:innen „durchschnittlich“ zu räumlichem Denken in der Lage ist.

Betrachtet man die Modellkompetenz-Niveaus, werden HoEr bei den Teilkompetenzen *Eigenschaften von Modellen* und *Alternative Modelle* signifikant höhere Niveaus zugeordnet als NiEr (Tabelle 10). Wie in Kapitel 3.2.2 erläutert, ist dies darauf zurückzuführen, dass Wissenschaftler:innen Modelle wesentlich häufiger methodisch einsetzen als Schüler:innen bzw. Studierende.

7.10. Methodenkritik

7.10.1. Versuchssetting

Bei der Einordnung der Ergebnisse muss die Rolle des Versuchssettings berücksichtigt werden. Das ursprüngliche Studiendesign sah eine stärkere Gewichtung der Untersuchung der Blickdaten vor. Entsprechend wurde das Versuchsdesign darauf folgendermaßen abgestimmt: Die Aufgaben wurden auf einem Bildschirm gezeigt und die Proband:innen sollten sich für eine störungsarme Erfassung der Augenbewegungen während der Bearbeitung der Instruktionen möglichst nicht bewegen. Es durften keinerlei Notizen gemacht werden oder der Bildschirm

berührt werden. In Kombination mit dem *CTA*, das ohnehin einen hohen *Cognitive Load* erzeugen kann (Bojko, 2013), ergab sich für die Proband:innen eine ungewohnte und somit potentiell herausfordernde Versuchssituation. Zusätzlich wurden bei Aufgaben mit methodischem Modelleinsatz Merkmale im Text genannt und nicht im DpV eingezeichnet, um die DpV für die Erhebung der visuellen Daten möglichst einheitlich zu gestalten. Dies erhöht wiederum den *Cognitive Load*, da die Textinformation auf das DpV übertragen werden muss, ohne dabei Hilfsmittel wie Notizen im DpV anwenden zu können. Sind Merkmale im DpV eingetragen, kann dies Studierenden das Tree-Reading erleichtern (Novick et al., 2010).

Um negativen Auswirkungen des Versuchssettings entgegenzuwirken, konnten sich die Proband:innen in einem Probelauf an das Setting gewöhnen. Außerdem konnten sich Proband:innen bei jeder neuen Aufgabe in der ersten Aufgabe in das bzw. die DpV und den Aufgabenstamm einarbeiten (Kapitel 5.6). Trotz der Aufforderung, die Aufgaben genau zu befolgen, antizipierten einige Proband:innen bereits während der Aufgabe zur Einarbeitung die mediale, aber auch die Aufgabe im methodischen Modelleinsatz anhand des Aufgabenstamms. Dies führte dazu, dass einige Proband:innen in der Aufgabe zum Einarbeiten unbewusst die mediale oder methodische Aufgabenstellung beantworteten, ihre Lösung später in der eigentlichen Aufgabe jedoch nicht oder nur teilweise wiederholten. Um die Fragestellungen der vorliegenden Studie beantworten zu können und auch eine einheitliche Datenauswertung zu gewährleisten, flossen diese antizipierten Lösungen nicht in die Auswertung mit ein. Nichtsdestotrotz muss berücksichtigt werden, dass wahrscheinlich diese Maßnahmen zur Aufgabengewöhnung und somit zur Verringerung des *Cognitive Load* einen Einfluss auf die Lösungsprozesse einiger Proband:innen hatten.

Für Folgestudien mit dem Fokus auf verbalen Daten ergibt sich, dass unabhängig davon, ob die Aufgaben digital oder analog präsentiert werden, die Proband:innen dabei die Möglichkeit haben sollten, Merkmale in DpV einzutragen, Klade zu markieren oder alternative DpV zu skizzieren. Insbesondere bei der Teilkompetenz *Zweck von Modellen* kann so der *Cognitive Load* verringert werden, da das anspruchsvolle mentale Modellieren, bspw. beim Untersuchen von Entstehungs- und Verlustereignissen anhand verschiedener mentaler Modelle (Hochpöchler et al., 2013), durch das Ergänzen und Erweitern des Modellobjekts vereinfacht werden kann.

Bei DpV liegen bedingt durch die stark abstrahierte Repräsentationsform inhaltlich relevante DpV-Bestandteile wie Knotenpunkte und Taxa nah beieinander. In der vorliegenden Studie

wurden außerdem noch Aufgabenstamm und Aufgabe gemeinsam mit dem bzw. den DpV auf einem 22-Zoll-Monitor präsentiert. Versieht man DpV-Bestandteile wie Taxa oder Knotenpunkte mit AOI, liegen diese aufgrund der grafischen Dichte der DpV nah beieinander, sodass Fixationen nicht eindeutig zuzuordnen sind. In Folgestudien mit Fokus auf visuelle Prozesse beim Tree-Reading sollten die DpV in ausreichender Größe dargestellt werden, sodass eine eindeutige Differenzierung von DpV-Bestandteilen möglich ist.

7.10.2. Weitere Einflussfaktoren

Die in Kapitel 7.5 erläuterten Einflussfaktoren sollten ebenfalls beim Design von Folgestudien berücksichtigt werden. Die Repräsentationsform sollte zwischen den Aufgaben einer Teilkompetenz möglichst wenig variieren. Bspw. können grundsätzlich balancierte DpV eingesetzt werden. In der vorliegenden Studie kann nicht ausgeschlossen werden, dass die Verwendung des Begriffs Sparsamkeitsprinzip trotz kurzer Einführung einzelne NiEr in ihren Lösungsprozessen beeinflusste. Da Lösungsprozesse unabhängig von der Verwendung von Fachsprache kodiert wurden (Kapitel 5.8.1) und die methodischen Aufgaben auch ohne Verwendung des Begriffs Sparsamkeitsprinzip lösbar gewesen wären, sollte in Folgestudien daher Einsatz von Fachwörtern sorgfältig abgewogen werden.

Die Teilnahme an der Studie erfolgte freiwillig. Für der Akquise der Proband:innen wurde der Inhalt der Studie mit der Untersuchung von „Blickbewegungen beim Lesen von evolutiven Stammbäumen mithilfe von Eyetracking“ grob umrissen. Die Adressat:innen der Akquise entschieden sich folglich vermutlich eher dann für die Teilnahme, wenn sie sich für das Thema Evolution oder die Methode des *Eyetrackings* interessierten. Die NiEr der Kohorte zeigen aber dennoch viele Varianten von Tree-Reading-Performanz, sodass ein zu erwartender Selektionseffekt (Bortz & Döring, 2006, S. 557) nicht allzu sehr ins Gewicht zu fallen scheint.

8. Implikationen

Förderung von Tree-Reading-Fähigkeiten

Die vorliegende Studie stützt die Annahme, dass Tree-Reading-Fähigkeiten nicht wie von Halverson und Friedrichsen (2013) postuliert in ihrer Gesamtheit in aufeinander aufbauende Level eingeteilt werden. Vielmehr ist von einer Mehrdimensionalität der Tree-Reading-Fähigkeiten auszugehen, in der eine Person je nach Dimension zu verschiedenen anspruchsvollen Tree-Reading-Prozessen in der Lage ist (vgl. Schramm et al., 2021a; Kapitel 7.4). Das bedeutet für den Biologieunterricht, dass nicht wie von Halverson und Friedrichsen (2013) angenommen ein Level nach dem anderen unterrichtet werden sollte, sondern wie von Schramm und Kolleg:innen (2021a) vorgeschlagen das parallele Unterrichten aller Dimensionen in aufsteigender Schwierigkeit stattfinden sollte. Zu Beginn solcher Einheiten bietet es sich an, den Stand der Lernenden in den einzelnen Dimensionen zu diagnostizieren, um differenziert unterrichten zu können. Dafür kann der von Schramm und Kolleg:innen (2021a) entwickelte Fragenbogen genutzt werden. Zahlreiche Materialien für Lehrkräfte finden sich in englischer Sprache auf der Seite des Projekts *Understanding Evolution* des University of California Museum of Paleontology (UCMP; <https://evolution.berkeley.edu/>).

Förderung von Modellierkompetenz

In der vorliegenden Studie wurde demonstriert, wie DpV eingesetzt werden können, um den medialen und auch den methodischen Modelleinsatz in den Teilkompetenzen *Zweck von Modellen* und *Alternative Modelle* zu adressieren. Wie in Kapitel 5.4.1 erläutert, können auch die anderen Teilkompetenzen mittels DpV behandelt werden, indem z. B. der Hypothesencharakter von DpV betont und damit die Teilkompetenz *Eigenschaften von Modellen* angesprochen wird. Außerdem können DpV datenbasiert erstellt oder verändert werden (Tree-Building), um die Teilkompetenzen *Testen von Modellen* und *Ändern von Modellen* zu behandeln. Wie in Kapitel 1 dargelegt, kann dies an gesellschaftsrelevanten Themen wie der SARS-CoV-2-Pandemie oder den Auswirkungen der Klimakrise auf die Biodiversität veranschaulicht werden.²⁴ Insbesondere für das abduktive Schließen (Niveau IIIa; Krüger & Upmeier zu Belzen, 2021) eignen sich DpV, um das Entstehen von Merkmalen in der Vergangenheit auf hohem Niveau zu erklären.

²⁴ Für diese Themen gibt es ebenfalls zahlreiche Materialien auf der Website des Projekts *Understanding Evolution* (UCMP; <https://evolution.berkeley.edu/>).

Fachsprache

Abgesehen vom Stand der Lernenden innerhalb der Dimensionen ist das Beherrschen der Fachsprache essentiell. Die Verwendung von Begriffen wie Stammbaum, Vorfahre und Verwandtschaft in den Themen Evolution und Genetik sollte explizit besprochen werden, um Lernende dafür zu sensibilisieren, dass es sich um 2 unterschiedliche Konzepte mit unscharfer sprachlicher Trennung handelt. Wie in Kapitel 7.5.3 erläutert, ist unter Fachsprache jedoch nicht nur die Verbalisierung von Tree-Reading mitsamt der speziellen Fachbegriffe zu verstehen, sondern auch DpV selbst als Kommunikationsmittel und Informationsträger. DpV eignen sich also in besonderem Maße, um den Umgang mit Fachsprache zu fördern.

Vermittlung im Biologieunterricht: Zeitpunkt und Qualität

Auch wenn seit der Einführung der neuen Berliner Rahmenlehrpläne für die gymnasiale Oberstufe das Thema Evolution und erstmals sowohl Tree-Reading als auch Tree-Building fest verankert sind (Kapitel 2.6), stellen sich mehrere Fragen: Kann phylogenetische Verwandtschaft als abstraktes Konstrukt zur Erklärung allen Lebens auf der Erde verstanden und verinnerlicht werden, wenn es erst in der 10. Jahrgangsstufe konkret eingeführt und in der 11. Jahrgangsstufe bzw. erst kurz vor dem Abitur vertieft wird? Sollte Evolution nicht wie z. B. von der Nationalen Akademie der Wissenschaften Leopoldina (2017) gefordert, bereits ab dem Grundschulalter vermittelt werden?²⁵ Und wie gut kann der Rahmenlehrplan bzw. ein Schulcurriculum tatsächlich von der einzelnen Lehrkraft umgesetzt werden?

Wie die vorliegende Studie zeigt, unterscheiden sich die individuellen Tree-Reading-Performanzen bei Studierenden des *Master of Education* individuell sehr stark, sodass davon auszugehen ist, dass das Evolutionsverständnis bei angehenden Lehrkräften auch am Ende eines fachdidaktischen Studiums nicht zwangsläufig gut ausgeprägt ist. Um bestmöglich die Tree-Reading-Fähigkeiten, Fachsprache und Modellierkompetenz mit DpV von Lernenden zu fördern, ist es jedoch unabdingbar, die Lehrkräfte in ebendiesen Punkten gezielt im Studium auszubilden. Entsprechend sollte nicht nur die fachwissenschaftliche Ausbildung sondern gerade die fachdidaktische Ausbildung den Fokus stärker auf die Vermittlung von Evolution legen und hier DpV als Modelle *von* und *für* Evolution einsetzen.

²⁵ Das Projekt Evokids stellt wissenschaftlich fundierte Materialien für den Einsatz ab dem Grundschulalter bereit: <https://evokids.de/>.

9. Fazit

Die vorliegende Arbeit beschäftigte sich mit dem Einfluss von Erfahrungsgrad und Modelleinsatz auf Tree-Reading-Prozesse. In Abgrenzung zur Mehrheit der Studien zu Tree-Reading ist hierbei zu betonen, dass der Fokus nicht darauf lag, *ob* eine Tree-Reading-Aufgabe gelöst wurde, sondern *inwiefern* dies geschah. Hierfür wurde charakterisiert, ob Proband:innen Informationen *oberflächlich*, *wissenschaftlich adäquat* bzw. *inadäquat* ablesen oder Informationen *ableiteten* und wie sie diese Informationen nutzten, um eine Tree-Reading-Aufgabe zu lösen.

Die Haupteckenerkenntnis der Studie ist, dass Tree-Reading-Prozesse bei NiEr wesentlich heterogener ausgeprägt sind, als es bspw. die Tree-Reading-Level von Halverson und Friedrichsen (2013) vermuten lassen. Vielmehr lassen sich Tree-Reading-Prozesse von NiEr besser mit dem *Revised STREAM* (Schramm et al., 2021a) erklären, dass Tree-Reading-Fähigkeiten in verschiedene Dimensionen einteilt und bei NiEr unterschiedliche Level je nach Dimension zulässt.

Eine weitere Erkenntnis ist, dass Tree-Reading-Prozesse von NiEr nicht nur vom Modelleinsatz beeinflusst werden, sondern auch die Repräsentationsform, die Aufgabenreihenfolge und die Fachsprache einen Einfluss haben. Außerdem ist davon auszugehen, dass das ungewohnte Versuchsetting ebenfalls Tree-Reading-Prozesse beeinflusst hat, da die Aufgaben ohne Hilfsmittel gelöst werden mussten und so der *Cognitive Load* insbesondere für NiEr vermutlich erhöht wurde.

Ursprünglich sollte der Fokus der Studie auf den visuellen Daten liegen, um explorativ kognitive Prozesse beim Tree-Reading zu untersuchen. Da Fixationen nicht eindeutig einzelnen DpV-Bestandteilen zugeordnet werden konnten, wechselte der Fokus zu den verbalen Daten. Diese erwiesen sich als so aussagekräftig, dass mithilfe der angewendeten Kriterien die Heterogenität von Tree-Reading-Prozessen bei NiEr detailliert herausgearbeitet werden konnte. Abgesehen vom angenommenen Einfluss des Versuchsettings auf NiEr hat sich zwar der Fokus, nicht aber die Aussagekraft der vorliegenden Studie verändert. Trotz der kleinen Stichprobe konnten verschiedene Varianten von Tree-Reading-Prozessen sowie Einflussfaktoren identifiziert werden.

Auch wenn in der Studie HoEr mit ihren homogenen Tree-Reading-Prozessen den Goldstandard bilden, um Tree-Reading-Prozesse von NiEr beschreiben zu können, so liegt doch die Aufmerksamkeit auf den Tree-Reading-Prozessen der NiEr. Denn je besser

Herausforderungen für NiEr beschrieben werden können, desto besser und individualisierter können NiEr im Tree-Reading und somit im Evolutionsverständnis gefördert werden. Dies wiederum kann zu einem besseren Verständnis von aktuellen Themen wie Pandemien, Artensterben, dem Menschen als Evolutionsfaktor oder persistierendem pseudowissenschaftlich begründetem Rassismus führen. Entsprechend kann das in der Einleitung vorgestellte Zitat von Metrick und Weitzmann (1998, S. 25) für die vorliegende Arbeit angepasst werden:

„*Tree thinking*“ represents a prime example of how one picture may be worth a thousand [...]” ***studies to better understand and teach evolution.***

10. Ausblick

Bezugnehmend auf das angepasste Zitat am Ende des vorherigen Kapitels bietet die vorliegende Studie einen ersten tiefen Einblick in das *inwiefern* von Tree-Reading-Prozessen und eröffnet zahlreiche Ansätze, um bestimmte Aspekte näher zu beleuchten:

In Kapitel 7.4 wurde erläutert, inwiefern die Aufgaben der vorliegenden Studie in das *Revised STREAM* (Schramm et al., 2021a) eingeordnet werden können. Umgekehrt könnte dieses empirisch überprüfte Modell als Rahmen für eine systematische qualitative Untersuchung von Tree-Reading-Prozessen dienen, indem für verschiedene Ausprägung jeder Dimension Aufgaben erstellt und qualitativ untersucht werden. Der Tree-Reading-Fragebogen, anhand dessen das *Revised STREAM* entwickelt wurde (Schramm et al., 2021a) kann hier parallel eingesetzt werden, um Tree-Reading-Fähigkeiten von Studierenden zu diagnostizieren und mit ausführlichen qualitative Daten zu den einzelnen Ausprägungen der Dimensionen zu triangulieren. So kann ein umfassendes Bild von Tree-Reading-Prozessen entstehen, um daraus wiederum Implikationen für Schule und Studium zu schärfen

In der vorliegenden Studie wurde unterschieden, ob eine Äußerung *oberflächlich* oder *wissenschaftlich adäquat abgelesen* wurde. Die verbalen Daten lassen hier dank des trennscharfen Kodiermanuals eine eindeutige Unterscheidung zu. Es bleibt jedoch offen, ob *oberflächliche Beschreibungen* immer implizieren, dass das Prinzip der gemeinsamen Abstammung nicht bekannt ist und die Äußerung aufgrund des generellen Verständnisses von Hierarchien erfolgt (Catley et al., 2013) oder ob es sich dabei um mangelnde Kenntnisse der Fachsprache in Bezug auf die Beschreibung von DpV-Bestandteilen handelt. Hieraus ergeben sich 2 Folgefragen: Inwiefern lässt sich von einer *oberflächlichen* oder einer *wissenschaftlich adäquaten* Beschreibung eines Vorfahren darauf schließen, ob das Prinzip der gemeinsamen

Abstammung verstanden wurde? Inwiefern hängt das Beherrschen von Fachsprache mit Tree-Reading-Fähigkeiten zusammen?²⁶ Ähnlich wie bei Blacquiere und Hoese (2016) könnte hier eine qualitative Folgestudie mit dem Fokus ausschließlich auf dem Prinzip des gemeinsamen Vorfahren durchgeführt werden, in der Studierende sowohl *MRCA* beschreiben als auch das Prinzip der gemeinsamen Abstammung erklären sollen.

Die vorliegende Studie zeigt mehrere zu beachtende Punkte für Folgestudien zu visuellen Daten auf: Einerseits müssen DpV in ausreichender Größe dargestellt werden, um Fixationen einzelner DpV-Bestandteile trennscharf analysieren zu können. Andererseits modellieren DpV stark abstrahiert einen ganz bestimmten Ausschnitt aus dem *Tree of Life*, sodass eine Unterscheidung in relevante und irrelevante DpV-Bestandteile (wie bspw. bei Gegenfurtner et al., 2011) nicht zielführend ist. Eher könnte die im vorherigen Absatz beschriebene Folgestudie zur Relevanz von Fachsprache noch um folgende Frage erweitert werden: Inwiefern spiegelt sich in visuellen Daten wider, ob ein gemeinsamer Vorfahre *oberflächlich* oder *wissenschaftlich adäquat beschrieben* wird?

Exemplarisch konnte in der vorliegenden Studie gezeigt werden, dass sowohl HoEr als auch NiEr bei der Teilkompetenz *Alternative Modelle* beide DpV verbal vergleichen und parallel dazu visuell fixieren. Das Vergleichen von 2 DpV ist zwar kognitiv anspruchsvoller als das Ablesen von Informationen aus einem DpV (Halverson, 2011; Kim et al., 2000). Während jedoch bei der Teilkompetenz *Alternative Modelle* beide zu vergleichende Modelle als Modellobjekte dargestellt werden, wird bei der Teilkompetenz *Zweck von Modellen* im methodischen Modelleinsatz eine Metaebene in Form eines mentalen Modells erforderlich: Das Modellobjekt wird mit verschiedenen mentalen Modellen von möglichen Verlust- und Entstehungsszenarien abgeglichen, um das Sparsamste zu ermitteln. Die Teilkompetenz *Zweck von Modellen* erfordert also abstraktere, tiefere kognitive Prozesse im Vergleich zur Teilkompetenz *Alternative Modelle*. Folgestudien können hier den Fokus in Richtung DpV als Modelle *von und für* evolutive Verwandtschaftsverhältnisse vertiefen mit besonderem Blick auf das mentale Modellieren.

Vor dem Hintergrund des Kompetenzmodells der Modellierkompetenz (Krüger & Upmeyer zu Belzen, 2021) bieten DpV eine fachspezifische Möglichkeit, um Niveau IIIa, das

²⁶ Dabei ist zu beachten, dass auch DpV als Repräsentationen eine Form der Fachsprache sind (vgl. Nitz, 2012, S. 40). Es müssen bei TR also 2 Formen der Fachsprache beherrscht werden.

erkenntnisgenerierende Erklären, zu charakterisieren und dessen Anwendung im Biologieunterricht zu untersuchen.

Abschließend lässt sich feststellen, dass sich Studierende des *Master of Education* bestens für Studien zu Tree-Reading-Prozessen eignen, die auf die Verbesserung des Evolutionsunterrichts in Schule und Universität abzielen. Einerseits haben sie selbst Biologieunterricht genossen, andererseits stehen sie kurz davor, selbst Biologie zu unterrichten. So können Daten anhand einer hochrelevanten Kohorte erhoben werden, um daraus Implikationen für die Förderung des Evolutionsverständnisses in Schule und Universität abzuleiten.

11. Danksagung

Ich danke Prof. Annette Upmeyer zu Belzen dafür, dass sie mir den Einstieg in die Fachdidaktik ermöglichte und ich so viel Freiheit bei der Auswahl und der Gestaltung meines Promotionsthemas hatte. Liebe Annette, ich habe von dir gelernt, jeden Satz auseinander zu nehmen, die Kausalität zu hinterfragen und auf den Punkt zu schreiben. Ich danke dir von Herzen dafür, dass ich auch in den belastenden Phasen immer dein Verständnis und deine Unterstützung sicher hatte. Und ich danke dir dafür, dass ich so eine prägende Zeit in Texas und meine 2 SWS unter anderem im Zoo und im Museum für Naturkunde verbringen durfte!

Der ganzen Fachdidaktik und Lehr-/ Lernforschung Biologie der HU Berlin kann ich nur sagen: Wenn ich euch besuche, ist es wie nach Hause kommen! Ich habe mich all die Jahre sehr wohl gefühlt und erinnere mich gern an viele fachliche und unfachliche Gespräche in der Mensa, an der Kaffeemaschine und im Türrahmen. Sabine, Johannes, Joé, Anna, Ronja, Maria, Daniela, Hauke, Kerstin, Jeane, Sandra, Alex, Juliane, Christiane, Meta und alle, die ich noch vergessen habe: Ihr alle habt meine Zeit besonders gemacht!

Ich danke all meinen Proband:innen dafür, dass sie stundenlang regungslos vor einem Bildschirm gesessen haben und so spannende Dinge gesagt haben! Und vielen Dank an Marvin, Hanna und Sascha fürs unermüdliche Transkribieren, an Anna für die Zweitkodierung und Sabine für die Abrundung der Diskussion!

Meiner Familie und allen Freund:innen danke ich für die Unterstützung und das Verständnis über all die Jahre!

Der größte Dank geht an meinen Mann Philipp, du hast mich so oft wieder aufgebaut, hast dir geduldig Fachsimpelei über Evolution und Modelle angehört, hast verinnerlicht, dass Kompetenz wichtig ist (Kompetenzkompetenz!) und hast vor allem in der Abschlussphase den Großteil der Betreuung von Pauli übernommen. Ohne dich hätte ich das alles nicht geschafft!

Es war nie selbstverständlich, dass ich bis zu diesem Punkt komme und doch irgendwie immer.

12. Literaturverzeichnis

- Ainsworth, S., & Saffer, J. (2010). Can children read trees? In International Society of the Learning Sciences (Chair), *International Conference of the Learning Sciences*, Chicago.
- Ainsworth, S., & Saffer, J. (2013). Can Children Read Evolutionary Trees? *Merrill-Palmer Quarterly*, 59(2), 221.
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education (2014). *Standards for Educational and Psychological Testing*. Washington, D.C.: American Educational Research Association.
- Atran, S. (1998). Folk biology and the anthropology of science: Cognitive universals and cultural particulars. *Behavioral and Brain Sciences*, 21(4), 547-69; discussion 569-609.
- Ax, P. (1984). *Das phylogenetische System: Systematisierung d. lebenden Natur aufgrund ihrer Phylogenese*. Stuttgart, New York: Fischer.
- Baum, D. A., & Offner, S. (2008). Phylogenies & Tree-Thinking. *The American Biology Teacher*, 70(4), 222–229.
- Baum, D. A., & Smith, S. D. (2013). *Tree thinking: An introduction to phylogenetic biology*. Greenwood Village, Colo.: Roberts.
- Baum, D. A., DeWitt Smith, S., & Donovan, S. S. S. (2005). The Tree-Thinking Challenge. *Science*, 310(5750), 979–980. <https://doi.org/10.1126/science.1117727>
- Beniermann, A. (2019). *Evolution – von Akzeptanz und Zweifeln: Empirische Studien über Einstellungen zu Evolution und Bewusstsein*. Wiesbaden: Springer Fachmedien Wiesbaden.
- Blacquiere, L. D., & Hoese, W. J. (2016). A valid assessment of students' skill in determining relationships on evolutionary trees. *Evolution: Education and Outreach*, 9(1), 979.
- Blacquiere, L. D., Fawaz, A., & Hoese, W. J. (2020). Who's related to whom? Use published phylogenies and make customized Tree-Thinking assessments. *Evolution: Education and Outreach*, 13(1), 979.
- Bojko, A. (2013). *Eye tracking the user experience: A practical guide to research*. Brooklyn New York: Rosenfeld Media.
- Borrmann, J. R., Reinhardt, N., Krell, M., & Krüger, D. Perspektiven von Lehrkräften über Modelle in den Naturwissenschaften: Eine generalisierende Replikationsstudie. In *Erkenntnisweg Biologiedidaktik* (pp. 57–72).
- Bortz, J., & Döring, N. (2006). *Forschungsmethoden und Evaluation: für Human- und Sozialwissenschaftler* (4th ed.). Berlin Heidelberg: Springer Berlin Heidelberg.
- Brennan, R. L., & Prediger, D. J. (1981). Coefficient Kappa: Some Uses, Misuses, and Alternatives. *Educational and Psychological Measurement*, 41(3), 687–699.
- Brennecke, J. S. (2014). *Schülervorstellungen zur evolutionären Anpassung: qualitative Studien als Grundlage für ein fachdidaktisches Entwicklungskonzept in einem botanischen Garten* (Dissertation). Justus-Liebig-Universität Gießen, Gießen.
- Bucklin, C. J., Daniel, K. L., Nolen, Z., Mead, L., Jenkins, K., & Baum, D. (2019, April). *Developing a Tree-Thinking Diagnostic Test to Assess Individual Elements of Representational Competence*. Paper presented at the annual meeting of the American Educational Research Association, Toronto, CAN.
- Catley, K. M., & Novick, L. R. (2008). Seeing the Wood for the Trees: An Analysis of Evolutionary Diagrams in Biology Textbooks. *BioScience*, 58(10), 976–987.

- Catley, K. M., Novick, L. R., & Funk, D. J. (2012). The Promise and Challenges of Introducing Tree Thinking into Evolution Education. In K. S. Rosengren (Ed.), *Evolution challenges: Integrating research and practice in teaching and learning about evolution* (pp. 93–118). Oxford: Oxford Univ. Press.
- Catley, K. M., Phillips, B. C., & Novick, L. R. (2013). Snakes and Eels and Dogs! Oh, My! Evaluating High School Students' Tree-Thinking Skills: An Entry Point to Understanding Evolution. *Research in Science Education*, 43(6), 2327–2348.
- Catley, K. M., Phillips, B. C., & Novick, L. R. (2013). Snakes and Eels and Dogs! Oh, My! Evaluating High School Students' Tree-Thinking Skills: An Entry Point to Understanding Evolution. *Research in Science Education*, 43(6), 2327–2348.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (Second Edition). New York: Academic Press.
- Cook, M. P. (2006). Visual representations in science education: The influence of prior knowledge and cognitive load theory on instructional design principles. *Science Education*, 90(6), 1073–1091.
- Crisp, M. D., & Cook, L. G. (2005). Do early branching lineages signify ancestral traits? *Trends in Ecology & Evolution*, 20(3), 122–128.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334.
- Darwin, C. R. (1837). *Notebook B*.
- Dees, J., Bussard, C., & Momsen, J. L. (2018). Further Effects of Phylogenetic Tree Style on Student Comprehension in an Introductory Biology Course. *CBE Life Sciences Education*, 17(2).
- Dees, J., Friermuth, D., & Momsen, J. L. (2017). Effects of Phylogenetic Tree Style on Student Comprehension in an Introductory Biology Course. *The American Biology Teacher*, 79(9), 729–737.
- Dobzhansky, T. (1973). Nothing in Biology Makes Sense except in the Light of Evolution. *The American Biology Teacher*, 35(3), 125–129.
- Dresing, T., & Pehl, T. (Eds.) (2018). *Praxisbuch Interview, Transkription & Analyse: Anleitungen und Regelsysteme für qualitativ Forschende* (8. Auflage). Marburg: Eigenverlag.
- Eddy, S. L., Crowe, A. J., Wenderoth, M. P., & Freeman, S. (2013). How should we teach Tree-Thinking? An experimental test of two hypotheses. *Evolution: Education and Outreach*, 6(1).
- Eldredge, N., & Cracraft, J. (1980). *Phylogenetic patterns and the evolutionary process: Method and theory in comparative biology*. New York: Columbia University Press.
- Ericsson, K. A. & Lehmann, A. C. (1996) Expert and Exceptional Performance: Evidence of Maximal Adaptation to Task Constraints.
- Ericsson, K. A., & Simon, H. A. (1993). Verbal reports as data. *Psychological Review*, 87(3), 215–251.
- Field, A. (2009). *Discovering Statistics Using SPSS*. (3rd Edition). London: Sage Publications Ltd.
- France, B. (2019). Illuminating Scientists' Modeling Competence. In *Towards a Competence-Based View on Models and Modeling in Science Education* (pp. 81–97). Springer, Cham.
- France, B., Compton, V. J., & Gilbert, J. K. (2011). Understanding modelling in technology and science: The potential of stories from the field. *International Journal of Technology and Design Education*, 21(3), 381–394.

- Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise Differences in the Comprehension of Visualizations: A Meta-Analysis of Eye-Tracking Research in Professional Domains. *Educational Psychology Review*, 23(4), 523–552.
- Gilbert, S. W. (1991). Model building and a definition of science. *Journal of Research in Science Teaching*, 28(1), 73–79.
- Gittler, G. (1990). *Dreidimensionaler Würfeltest (3DW)*. Weinheim: Beltz.
- Göhner, M., & Krell, M. Qualitative Inhaltsanalyse in naturwissenschaftsdidaktischer Forschung unter Berücksichtigung von Gütekriterien: Ein Review. In (Vol. 26, pp. 207–225). <https://doi.org/10.1007/s40573-020-00111-0>
- Goldsmith, D. W. (2003). The Great Clade Race: Presenting Cladistic Thinking to Biology Majors & General Science Students. *The American Biology Teacher*, 65(9), 679–682.
- Graf, D., & Hamdorf, E. (2011). Evolution Verbreitete Fehlvorstellungen zu einem zentralen Thema. In *Evolutionsbiologie* (pp. 25–41). Spektrum Akademischer Verlag.
- Gregory, T. R. (2008). Understanding Evolutionary Trees. *Evolution: Education and Outreach*. (1), 121–137.
- Gropengießer, H., & Marohn, A. (2018). Schülervorstellungen und Conceptual Change. In D. Krüger, I. Parchmann, & H. Schecker (Eds.), *Theorien in der naturwissenschaftsdidaktischen Forschung* (pp. 49–67). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Grosslight, L., Unger, C., & Jay, E. (1991). Understanding models and their use in science: Conceptions of middle and high school students and experts. *Journal of Research in Science Teaching*, 28(9), 799–822.
- Grünkorn, J., Upmeier 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*, 36(10), 1651–1684.
- Halverson, K. L. (2009). *Investigating the development and use of phylogenetic representations by college undergraduates in a plant systematics course*. Dissertation.
- Halverson, K. L. (2010). Using Pipe Cleaners to Bring the Tree of Life to Life. *The American Biology Teacher*, 72(4), 223–224.
- Halverson, K. L. (2011). Improving Tree-Thinking One Learnable Skill at a Time. *Evolution: Education and Outreach*, 4(1), 95–106.
- Halverson, K. L., & Friedrichsen, P. (2013). Learning Tree Thinking: Developing a New Framework of Representational Competence. In D. F. Treagust & C.-Y. Tsui (Eds.), *Models and Modeling in Science Education: Vol. 7. Multiple Representations in Biological Education* (Vol. 7, pp. 185–201). Dordrecht: Springer.
- Halverson, K. L., Pires, C. J., & Abell, S. K. (2011). Exploring the complexity of tree thinking expertise in an undergraduate systematics course. *Science Education*, 95(5), 794–823.
- Hammann, M., & Scheffel, L. (2005). Stammbaumtraining durch Vergleichen. *Unterricht Biologie*, 29(310), 38–44.
- Hartig, J., & Klieme, E. (2006). Kompetenz und Kompetenzdiagnostik. In *Leistung und Leistungsdiagnostik* (pp. 127–143). Springer, Berlin, Heidelberg.
- Hellström, N. P. (2012). Darwin and the Tree of Life: the roots of the evolutionary tree. *Archives of Natural History*, 39(2), 234–252. <https://doi.org/10.3366/anh.2012.0092>

- Hochpöchler, U., Schnotz, W., Rasch, T., Ullrich, M., Horz, H., McElvany, N., & Baumert, J. (2013). Dynamics of mental model construction from text and graphics. *European Journal of Psychology of Education, 28*(4), 1105–1126.
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford, New York: Oxford University Press.
- Hussy, W., Schreier, M., & Echterhoff, G. (2010). *Forschungsmethoden in Psychologie und Sozialwissenschaften - für Bachelor. Springer-Lehrbuch*. Berlin, Heidelberg: Springer Berlin Heidelberg.
- Jarodzka, H., Scheiter, K., Gerjets, P., & van Gog, T. (2010). In the eyes of the beholder: How experts and novices interpret dynamic stimuli. *Learning and Instruction, 20*(2), 146–154.
- Jenkins, K. P., Mead, L., Baum, D. A., Daniel, K. L., Bucklin, C. J., Leone, E. A., . . . Naegle, E. (2022). Developing the BETSI: A Tree-Thinking Diagnostic Tool to Assess Individual Elements of Representational Competence. *Evolution; International Journal of Organic Evolution*. Advance online publication.
- Jördens, J., Asshoff, R., & Kullmann, H. (2011). Stammbäume lesen und verstehen. In *Evolutionsbiologie* (pp. 377–403). Spektrum Akademischer Verlag.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329–354.
- Kambach, M. (2018). *Experimentierprozesse von Lehramtsstudierenden der Biologie. Eine Videostudie*. Logos Verlag Berlin.
- Kattmann, U. (2013). Diagramme. In Gropengießer, H., Harms, U., & Kattmann, U. (Eds.). *Fachdidaktik Biologie* (9, völlig überarb. Aufl.). Hallbergmoos: Aulis Verl. in der Stark Verlagsges.
- Kim, J., Hahn, J., & Hahn, H. (2000). How Do We Understand a System with (So) Many Diagrams? Cognitive Integration Processes in Diagrammatic Reasoning. *Information Systems Research, 11*(3), 284–303.
- Klein, P., Graulich, N., Kuhn, J., & Schindler, M. (Eds.) (2022). *Eye-Tracking in der Mathematik- und Naturwissenschaftsdidaktik: Forschung und Praxis*. Berlin: Springer Spektrum. <https://doi.org/10.1007/978-3-662-63214-7>
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (Ed.) (2004b). *Einheitliche Prüfungsanforderungen in der Abiturprüfung Biologie: (Beschluss der Kultusministerkonferenz vom 01.12.1989 i.d.F. vom 05.02.2004)*. München: Luchterhand.
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (Ed.) (2004b). *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss* (1. Auflage). München: Luchterhand.
- KMK (Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland) (Ed.) (2020). *Bildungsstandards im Fach Biologie für die Allgemeine Hochschulreife: Beschluss der Kultusministerkonferenz vom 18.06.2020* (1. Auflage). Köln: Carl Link Verlag.
- Knoop, V., & Müller, K. (2009). *Gene und Stammbäume*. Spektrum Akademischer Verlag.
- Kong, Y., Anderson, T., & Pelaez, N. (2016). How to Identify and Interpret Evolutionary Tree Diagrams. *Journal of Biological Education, 1–12*.
- Kozma, R., & Russell, J. (2005). Students becoming Chemists: Developing Representational Competence. In J. K. Gilbert (Ed.), *Visualization in science education* (pp. 121–145). Dordrecht: Springer.
- Krell, F. T., & Cranston, P. S. (2004). Which side of the tree is more basal? *Systematic Entomology, 29*, 279–281.

- Krell, M., & Krüger, D. (2016). Testing Models: A Key Aspect to Promote Teaching Activities Related to Models and Modelling in Biology Lessons? *Journal of Biological Education*, 50(2), 160–173.
- Krüger, D., & Upmeyer zu Belzen, A. (2021). Kompetenzmodell der Modellierkompetenz – Die Rolle abduktiven Schließens beim Modellieren. *Zeitschrift Für Didaktik Der Naturwissenschaften*, 24(260), 21.
- Krüger, D., Kauertz, A., & Upmeyer zu Belzen, A. (2018). Modelle und das Modellieren in den Naturwissenschaften. In D. Krüger, I. Parchmann, & H. Schecker (Eds.), *Theorien in der naturwissenschaftsdidaktischen Forschung* (pp. 141–157). Berlin, Heidelberg: Springer Berlin Heidelberg.
- Kuckartz, U. (2010). *Einführung in die computergestützte Analyse qualitativer Daten* (3rd ed.). Wiesbaden: VS Verlag für Sozialwissenschaften.
- Kummer, T. A. (2017). *Assessing and Improving Student Understanding of Tree-Thinking* (Dissertation). Brigham Young University, Utah.
- Kummer, T. A., Whipple, C. J., Bybee, S. M., Adams, B. J., & Jensen, J. L. (2019). Development of an Evolutionary Tree Concept Inventory. *Journal of Microbiology & Biology Education*, 20(2).
- Lai, M.-L., Tsai, M.-J., Yang, F.-Y., Hsu, C.-Y., Liu, T.-C., Lee, S. W.-Y., . . . Tsai, C.-C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review*, 10, 90–115.
- Lenzer, S., & Nehring, A. (2022). Eye-Tracking als Methode hypothesenprüfender Forschung in den Naturwissenschaftsdidaktiken – Neue Einblicke in die Effekte von Demonstrationsexperimenten. In *Eye-Tracking in der Mathematik- und Naturwissenschaftsdidaktik* (pp. 145–158). Springer Spektrum, Berlin, Heidelberg.
- Liu, J., & Novick, L. R. (2021). Intentional Diagram Design: Using Gestalt Perceptual Grouping in Cladograms to Tackle Misconceptions. In A. Basu, G. Stapleton, S. Linker, C. Legg, E. Manalo, & P. Viana (Eds.), *Springer eBook Collection: Vol. 12909. Diagrammatic Representation and Inference: 12th International Conference, Diagrams 2021, Virtual, September 28–30, 2021, Proceedings* (1st ed., Vol. 12909, pp. 431–438). Cham: Springer International Publishing; Imprint Springer.
- MacDonald, T. (2014). *Evolutionary tree design: An exploratory study of the influence of linear versus branching format on visitors' interpretation and understanding across age groups*. University of Kansas.
- Mahr, B. (2008). Ein Modell des Modellseins: Ein Beitrag zur Aufklärung des Modellbegriffs. In E. Knobloch & U. Dirks (Eds.), *Modelle* (pp. 187–218). Frankfurt am Main: Peter Lang.
- Mahr, B. (2009). Die Informatik und die Logik der Modelle. *Informatik Spektrum*, 32(3), 228–249.
- Mayring, P. (2003). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (8. Aufl.). UTB für Wissenschaft Pädagogik: Vol. 8229. Weinheim: Beltz.
- Mayring, P. (2015). *Qualitative Inhaltsanalyse: Grundlagen und Techniken* (12., überarb. Aufl.). Beltz Pädagogik. Weinheim: Beltz.
- Meir, E., Perry, J., Herron, J. C., & Kingsolver, J. (2007). College Students' Misconceptions About Evolutionary Trees. *The American Biology Teacher*, 69(7), 71–76.
- Metrick, A., & Weitzman, M. L. (1998). Conflicts and Choices in Biodiversity Preservation. *Journal of Economic Perspectives*, 12(3), 21–34.
- Morabito, N. P., Catley, K. M., & Novick, L. R. (2010). Reasoning about evolutionary history: Post-secondary students' knowledge of most recent common ancestry and homoplasy. *Journal of Biological Education*, 44(4), 166–174.

- Morel, B., Barbera, P., Czech, L., Bettisworth, B., Hübner, L., Lutteropp, S., . . . Stamatakis, A. (2021). Phylogenetic Analysis of SARS-CoV-2 Data Is Difficult. *Molecular Biology and Evolution*, *38*(5), 1777–1791.
- Nationale Akademie der Wissenschaften Leopoldina (Ed.) (2017). Evolutionsbiologische Bildung in Schule und Hochschule: Bedeutung und Perspektiven [Special issue]. Halle (Saale): Deutsche Akademie der Naturforscher Leopoldina e.V Nationale Akademie der Wissenschaften: Deutsche Akademie der Naturforscher Leopoldina e.V Nationale Akademie der Wissenschaften.
- Nee, S. The great chain of being: Our persistence in placing ourselves at the top of the Great Chain of Being suggests we have some deep psychological need to see ourselves as the culmination of creation. *Nature*, *(435)*, 429.
- Nehm, R. H., & Ha, M. (2011). Item feature effects in evolution assessment. *Journal of Research in Science Teaching*, *48*(3), 237–256.
- Nehm, R. H., & Ridgway, J. (2011). What Do Experts and Novices “See” in Evolutionary Problems? *Evolution: Education and Outreach*, *4*(4), 666–679. <https://doi.org/10.1007/s12052-011-0369-7>
- Nersessian, N. J. (2002). The cognitive basis of model-based reasoning in science. *The Cognitive Basis of Science*, 133–153.
- Nitz, S. (2012). *Fachsprache im Biologieunterricht: Eine Untersuchung zu Bedingungsfaktoren und Auswirkungen* (Dissertation). Christian-Albrechts-Universität zu Kiel.
- Nitz, S., Meister, S., Schwanewedel, J., & Upmeyer zu Belzen, A. (2018). Kompetenzraster zum Umgang mit Liniendiagrammen: Ein Beispiel für Diagnostik im Lehr-Lern-Labor. *Mathematischer Und Naturwissenschaftlicher Unterricht (MNU)*, *71*(6), 393–400.
- Novick, L. R. (1988). Analogical transfer, problem similarity, and expertise. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *14*(3), 510–520.
- Novick, L. R., & Catley, K. M. (2007). Understanding phylogenies in biology: The influence of a Gestalt perceptual principle. *Journal of Experimental Psychology: Applied*, *13*(4), 197–223.
- Novick, L. R., & Catley, K. M. (2013). Reasoning about evolution's grand patterns: College student's understanding of the tree of life. *American Educational Research Journal*, *50*(1), 138–177.
- Novick, L. R., & Catley, K. M. (2014). When Relationships Depicted Diagrammatically Conflict With Prior Knowledge: An Investigation of Students' Interpretations of Evolutionary Trees. *Science Education*, *98*(2), 269–304.
- Novick, L. R., & Catley, K. M. (2016). Fostering 21st-Century Evolutionary Reasoning: Teaching Tree Thinking to Introductory Biology Students. *CBE Life Sciences Education*, *15*(4).
- Novick, L. R., & Fuselier, L. C. (2019). Perception and conception in understanding evolutionary trees. *Cognition*, *192*, 104001.
- Novick, L. R., Catley, K. M., & Funk, D. J. (2010). Characters Are Key: The Effect of Synapomorphies on Cladogram Comprehension. *Evolution: Education and Outreach*, *3*(4), 539–547.
- Novick, L. R., Catley, K. M., & Funk, D. J. (2011). Inference is bliss: Using evolutionary relationship to guide categorical inferences. *Cognitive Science*, *35*(4), 712–743.
- Novick, L. R., Pickering, J., MacDonald, T., Diamond, J., Ainsworth, S., Aquino, A. E., . . . Scott, M. (2014). Depicting the tree of life in museums: guiding principles from psychological research. *Evolution: Education and Outreach*, *7*(1), 669.

- Novick, L. R., Shade, C. K., & Catley, K. M. (2011). Linear versus branching depictions of evolutionary history: Implications for diagram design. *Topics in Cognitive Science*, 3(3), 536–559.
- Novick, L. R., Stull, A. T., & Catley, K. M. (2012). Reading Phylogenetic Trees: The Effects of Tree Orientation and Text Processing on Comprehension. *BioScience*, 62(8), 757–764.
- O'Hara, R. J. (1988). Homage to Clio, or, toward an historical philosophy for evolutionary biology. *Systematic Zoology*, 37(2), 142–155.
- O'Hara, R. J. (1997). Population thinking and tree thinking in systematics. *Zoologica Scripta*, 26(4), 323–329.
- Omland, K. E., Cook, L. G., & Crisp, M. D. (2008). Tree thinking for all biology: the problem with reading phylogenies as ladders of progress. *BioEssays*, 30(9), 854–867.
- Orquin, J. L., & Holmqvist, K. (2018). Threats to the validity of eye-movement research in psychology. *Behavior Research Methods*, 50(4), 1645–1656.
- Oswald, E. D., & Roth, E. (1987). *Der Zahlen-Verbindungs-Test (ZVT)*. Hogrefe Verlag für Psychologie.
- Passmore, C., Gouvea, J. S., & Giere, R. (2014). Models in science and in learning science: Focusing scientific practice on sense-making. In M. R. Matthews (Ed.), *International Handbook of Research in History, Philosophy and Science Teaching* (1st ed., pp. 1171–1202). Niederlande: Springer Netherlands.
- Phillips, B. C., Novick, L. R., Catley, K. M., & Funk, D. J. (2012). Teaching Tree Thinking to College Students: It's Not as Easy as You Think. *Evolution: Education and Outreach*, 5(4), 595–602.
- Purves, W. K., Sadava, D. E., Orians, G. H., & Heller, H. C. (2004). *Biologie*. (7. Auflage). München, Heidelberg: Elsevier Spektrum Akademischer Verlag.
- Rayner, K. (1998). Eye movements in reading and information processing. 20 years of research. *Psychological Bulletin*, 124(3), 372–422.
- Rincke, K., Markic, S. (2018). Sprache und das Lernen von Naturwissenschaften. In: Krüger, D., Parchmann, I., Schecker, H. (Eds.) *Theorien in der naturwissenschaftsdidaktischen Forschung*. Springer, Berlin, Heidelberg
- Scholtz, G., & Scholtz, F. (2019). Phylogenetische Analyse von Kunstwerken: Ottavio Leonis Porträtserie Galileo Galileis von 1624. In J. Kaasch & M. Kaasch (Eds.), *Biologie und Kunst: Beiträge zur 25. Jahrestagung der DGGTB in Bonn 2016* (pp. 179–192). Verlag für Wissenschaft und Bildung.
- Schramm, T., & Schmiemann, P. (2019). Teleological pitfalls in reading evolutionary trees and ways to avoid them. *Evolution: Education and Outreach*, 12(1), 1271.
- Schramm, T., Jose, A., & Schmiemann, P. (2021a). Modeling and Measuring Tree-Reading Skills in Undergraduate and Graduate Students. *CBE Life Sciences Education*, 20(3).
- Schramm, T., Jose, A., & Schmiemann, P. (2021b). Seeing the Woods for the Trees Again: Analyzing Evolutionary Diagrams in German and US University-Level Textbooks. *Education Sciences*, 11(8), 367.
- Schramm, T., Schachtschneider, Y., & Schmiemann, P. (2019). Understanding the tree of life: an overview of Tree-Reading skill frameworks. *Evolution: Education and Outreach*, 12(1).
- Senatsverwaltung für Bildung, Jugend und Sport Berlin, & Ministerium für Bildung, Jugend und Sport des Landes Brandenburg (2021). *Rahmenlehrplan für die gymnasiale Oberstufe: Teil C Biologie*. Berlin.
- Senatsverwaltung für Bildung, Jugend und Sport Berlin, Ministerium für Bildung, Jugend und Sport des Landes Brandenburg, & Ministerium für Bildung, Wissenschaft und Kultur Mecklenburg-Vorpommern (2015). *Teil C - Biologie: Jahrgangsstufe 7-10*.

- Singer, F., Hagen, J. B., & Sheehy, R. R. (2001). The Comparative Method, Hypothesis Testing & Phylogenetic Analysis – An Introductory Laboratory. *The American Biology Teacher*, 63(7), 518–523.
- Stenlund, J. I., Schönborn, K. J., & Tibell, L. A. E. (2021). Zooming in Time—Exploring Students' Interpretations of a Dynamic Tree of Life. *Journal of Science Education and Technology*, 30(1), 125–138.
- Stieff, M., Hegarty, M., & Deslongchamps, G. (2011). Identifying Representational Competence With Multi-Representational Displays. *Cognition and Instruction*, 29(1), 123–145.
- Thewissen, J. G. M., Madar, S. I., & Hussain, S. T. (1998). Whale ankles and evolutionary relationships. *Nature*, 393(6685), 531.
- Treagust, D. F., Chittleborough, G., & Mamiala, T. L. (2002). Students' understanding of the role of scientific models in learning science. *International Journal of Science Education*, 24(4), 357–368.
- Trier, U., Krüger, D., & Upmeier zu Belzen, A. (2014). Students' versus scientists' conceptions of models and modelling. In D. Krüger & M. Ekborg (Eds.), *Research in biological education* (pp. 103–115). Westermann.
- Ubben, I., Nitz, S., Daniel, K. L., & Upmeier zu Belzen, A. (2018). Assessing Representational Competence with Eye Tracking Technology. In K. L. Daniel (Ed.), *Models and Modeling in Science Education. Towards a Framework for Representational Competence in Science Education* (Vol. 11, pp. 229–245). [S.I.]: Springer International Publishing.
- Ubben, I., Nitz, S., Rousseau, M., & Upmeier zu Belzen, A. (2015). Modelle von und für Evolution in Schulbüchern. In U. Gebhard, M. Hammann, & B. Knälmann (Chairs), *der Biologie im VBIO*. Symposium conducted at the meeting of Universität Hamburg, Hamburg, Germany.
- Ubben, I., Salisbury, S. L., & Daniel, K. L. (2019). Combining Visual and Verbal Data to Diagnose and Assess Modeling Competence. In A. Upmeier zu Belzen, D. Krüger, & J. van Driel (Eds.), *Models and Modeling in Science Education. Towards a Competence-Based View on Models and Modeling in Science Education* (Vol. 12, pp. 99–115). Cham: Springer International Publishing.
- Upmeier zu Belzen, A., & Krüger, D. (2010). Modellkompetenz im Biologieunterricht [Model competence in biology education]. *ZfDN*, 16, 41–57.
- Upmeier zu Belzen, A., & Krüger, D. (2019). Modelle als methodische Werkzeuge begreifen und nutzen: Empirische Befunde und Empfehlungen für die Praxis. In J. Groß, M. Hammann, P. Schmiemann, & J. Zabel (Eds.), *Biologiedidaktische Forschung: Erträge für die Praxis* (pp. 129–146). Berlin, Heidelberg: Springer Spektrum.
- van Dijk, E., & Kattmann, U. (2010). Evolution im Unterricht: Eine Studie über fachdidaktisches Wissen von Lehrerinnen und Lehrern. *Zeitschrift Für Didaktik Der Naturwissenschaften*, 16, 7–21.
- van Driel, J. H., & Verloop, N. (2002). Experienced teachers' knowledge of teaching and learning of models and modelling in science education. *International Journal of Science Education*, 24(12), 1255–1272.
- van Gog, T. (2012). Expertise. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning*. Boston, MA.
- van Gog, T., Paas, F., & van Merriënboer, J. J. G. (2005). Uncovering expertise-related differences in troubleshooting performance: combining eye movement and concurrent verbal protocol data. *Applied Cognitive Psychology*, 19(2), 205–221.
- van Someren, M. W., Barnard, Y. F., & Sandberg, J. A. C. (1994). *The think aloud method: A practical guide to modelling cognitive processes*. London: Academic Press.
- Viiri, J., Kekule, M., Isoniemi, J., & Hautala, J. (2017). Eye-tracking the Effects of Representation on Students' Problem Solving Approaches. In Finnish Mathematics and Science Education Research Association (Chair), *Proceedings of the annual FMSEERA symposium 2016*.

- Weigand, A. M. (2015). Baum ist nicht gleich Baum. Wie man phylogenetische Stolpersteine umgeht - ein kleiner Leitfaden. *Laborjournal*. (4), 24–26.
- Weinert, F. E. Vergleichende Leistungsmessung in Schulen - eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Ed.), *Leistungsmessungen in Schulen* (pp. 17–31). Weinheim: Beltz.
- Wilson, K. J., Long, T. M., Momsen, J. L., & Bray Speth, E. (2020). Modeling in the Classroom: Making Relationships and Systems Visible. *CBE Life Sciences Education*, 19(1).

Online-Ressourcen

- Jütte, P. (2021). Der Corona-Stammbaum - auf der Suche nach Viren.
<https://www.daserste.de/information/wissen-kultur/w-wie-wissen/Der-Corona-Stammbaum-100.html>.
Abgerufen am 17.12.2022.
- Lenzen, M. (2015). Was ist Sozialdarwinismus? CC BY-NC-ND 3.0 DE.
<https://www.bpb.de/themen/rechtsextremismus/dossier-rechtsextremismus/214188/was-ist-sozialdarwinismus/> Abgerufen am 20.12.2022.
- Nextstrain.org. Latest global SARS-CoV-2 analysis (GISAID data). CC BY 4.0.
<https://nextstrain.org/ncov/gisaid/global/6m>. Abgerufen am 17.12.2022.

13. Anhang

13.1. Anhang 1: Basic Evolutionary Tree-Thinking Skill Inventory (BETTSI)

BETTSI: Basic Evolutionary Tree-Thinking Skills Inventory

Fragen zum Umgang mit evolutiven Stammbäumen

Deutsche Übersetzung einer Vorversion von 2016, im englischen Original veröffentlicht von Bucklin et al. (2019) und Jenkins et al. (2022)

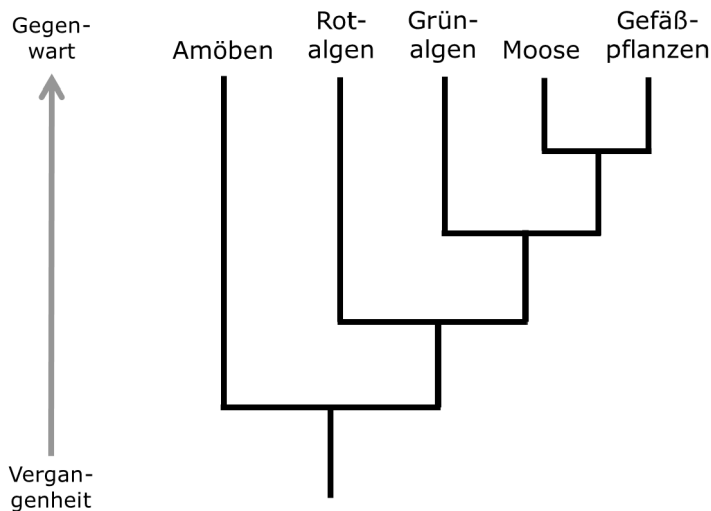
In der Biologie werden evolutive Verwandtschaftsverhältnisse zwischen Organismengruppen (Taxa) mithilfe von evolutiven Stammbäumen dargestellt. Mit Ihren Antworten zu den folgenden Fragen untersuchen wir Ihre Fähigkeit, evolutive Stammbäume zu lesen und evolutive Problemstellungen zu lösen.

Bearbeiten Sie nun bitte die 11 Fragen (Vorder- und Rückseite).

Jede Frage hat genau **eine** richtige Antwort.²⁷

²⁷ Korrekte Antworten sind grau hinterlegt.

1)



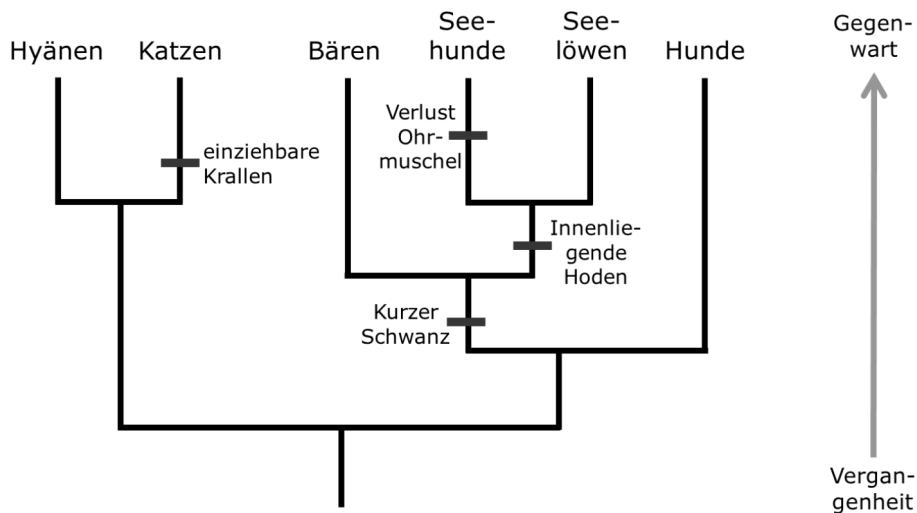
Welche der folgenden Aussagen über evolutive Verwandtschaft ist in Bezug auf den abgebildeten Stammbaum zutreffend?

- Grünalgen sind näher mit Rotalgen verwandt als mit Moosen.
- Grünalgen sind näher mit Moosen verwandt als mit Rotalgen.
- Grünalgen sind gleich nah mit Rotalgen und Moosen verwandt.
- Grünalgen sind mit Rotalgen verwandt aber nicht mit Moosen.
- Die Organismen sind untereinander nicht verwandt.

2) Drei Studierende diskutieren über die richtige Interpretation eines weiteren Aspektes des Stammbaums aus Aufgabe 1. Wer von ihnen hat Recht?

- Student A besteht darauf, dass Gefäßpflanzen das am höchsten entwickelte Taxon sind, weil sie sich als letzte entwickelt haben und komplexer als die anderen Spezies sind.
- Student B sagt, dass Amöben das am höchsten entwickelte Taxon sind, weil sie älter als die anderen Taxa sind.
- Student C sagt, dass von mehreren rezenten Taxa keines höher entwickelt ist als ein anderes, weil alle rezenten Taxa sich über den gleichen Zeitraum von ihrem letzten gemeinsamen Vorfahren aus entwickelt haben.
- Keiner der Studierenden hat Recht.
- Ich weiß nicht, wie ich den Stammbaum interpretieren soll.

3)



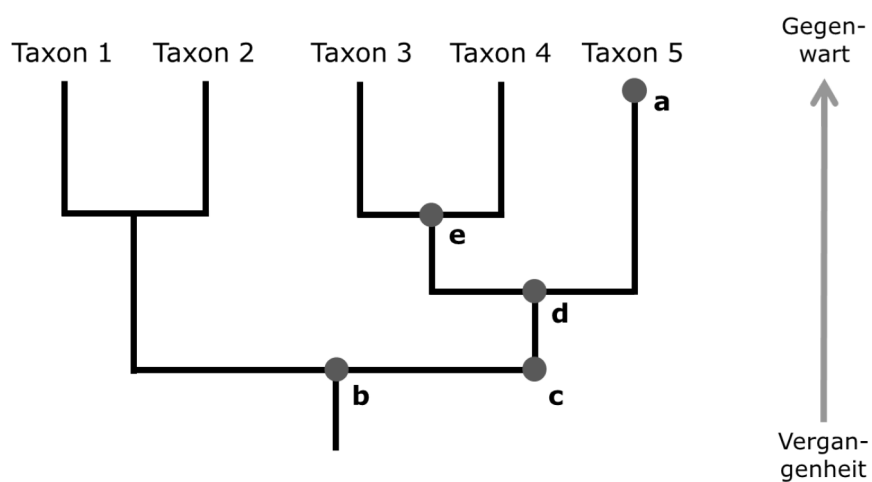
Für den abgebildeten Stammbaum wird angenommen, dass die Stammform einen langen Schwanz, Ohrmuscheln, äußere Hoden und feststehende Krallen hatte. Welche Merkmale haben Seelöwen basierend auf dem abgebildeten Stammbaum und der Annahme, dass alle evolutiven Veränderungen der Merkmale abgebildet sind?

- Langer Schwanz, Ohrmuscheln, außenliegende Hoden, feststehende Krallen
- Kurzer Schwanz, keine Ohrmuscheln, außenliegende Hoden, feststehende Krallen
- Kurzer Schwanz, keine Ohrmuscheln, innenliegende Hoden, feststehende Krallen
- Kurzer Schwanz, Ohrmuscheln, innenliegende Hoden, feststehende Krallen
- Langer Schwanz, Ohrmuscheln, innenliegende Hoden, einziehbare Krallen

4) Zwei Studierende diskutieren über die evolutiven Verwandtschaftsverhältnisse von Seelöwen, Seehunden und Hunden im Stammbaum aus Aufgabe 3. Wer von ihnen hat Recht?

- Student A sagt, dass Seehunde und Seelöwen gleich nah mit Hunden verwandt sind, weil die Entwicklungslinien von Seehunden und Seelöwen denselben gemeinsamen Vorfahren haben.
- Student B sagt, dass Seelöwen näher mit Hunden verwandt sind als Seehunde es sind, weil es weniger Merkmalsunterschiede zwischen Seelöwen und Hunden gibt und Seelöwen im gezeigten Stammbaum neben den Hunden abgebildet sind.
- Beide Studierenden haben Recht.
- Keiner der Studierenden hat Recht.
- Ich weiß nicht, wie ich den Stammbaum interpretieren soll.

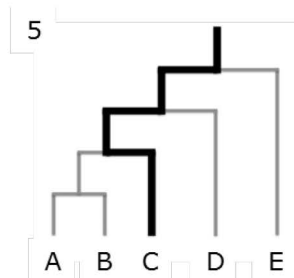
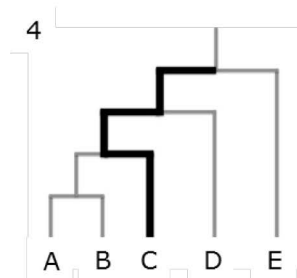
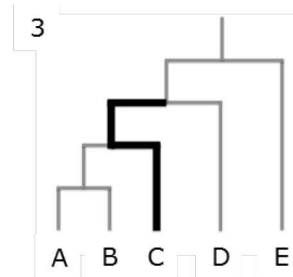
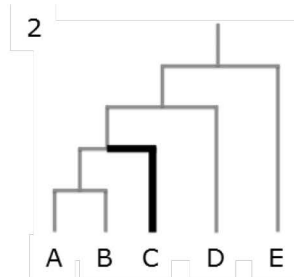
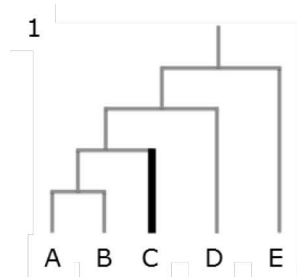
5)



Welche der fünf Markierungen mit den Buchstaben a bis e entspricht dem letzten gemeinsamen Vorfahren von Taxon 3 und Taxon 5?

- Markierung a
- Markierung b
- Markierung c
- Markierung d
- Markierung e

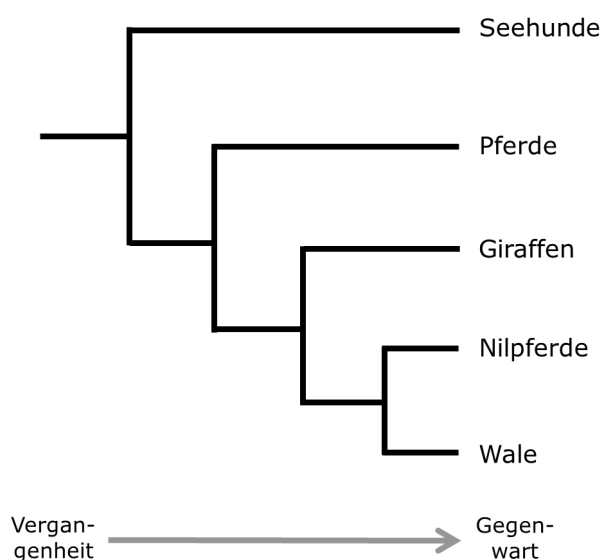
6)



Eine Evolutionslinie bezieht sich auf die gesamte evolutive Entstehungsgeschichte einer Spezies oder eines Taxons. In welchem der evolutionen Stammbäume wird die Evolutionslinie des Taxons C korrekt durch die fettgedruckte Linie dargestellt?

- Stammbaum 1
- Stammbaum 2
- Stammbaum 3
- Stammbaum 4
- Stammbaum 5

7)



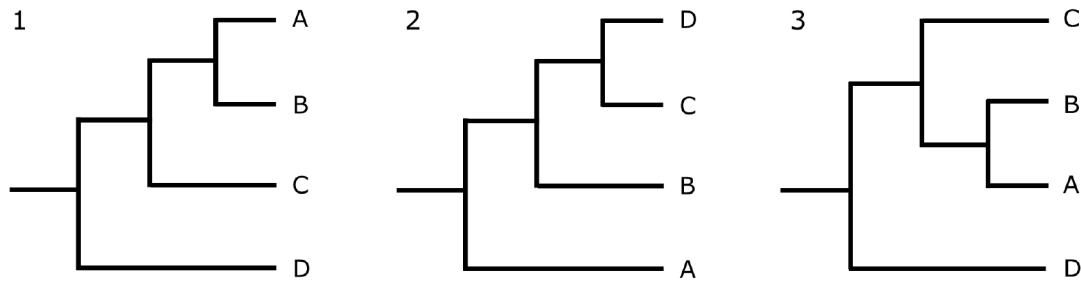
Welche der folgenden Aussagen trifft auf den dargestellten evolutiven Stammbaum zu?

- Seehunde sind näher mit Pferden verwandt als mit Walen.
- Seehunde sind näher mit Walen verwandt als mit Pferden.
- Seehunde sind gleich nah mit Pferden und Walen verwandt.
- Seehunde sind mit Walen verwandt, aber nicht mit Pferden.
- Die Organismen sind untereinander nicht verwandt.

8) Stellen Sie sich vor, Sie könnten in der Zeit zurückreisen und den letzten gemeinsamen Vorfahren von Giraffen und Nilpferden untersuchen. Was würde er sein?

- Eine Giraffe
- Ein Nilpferd
- Ein Pferd
- Eine Spezies, die nicht als eine der oben genannten klassifiziert werden kann
- Giraffen und Nilpferde haben keinen gemeinsamen Vorfahren

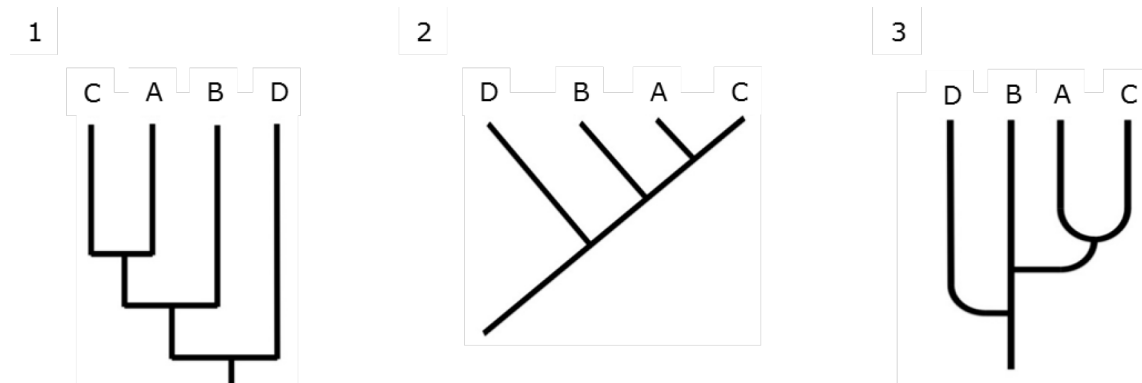
9)



Welcher der drei evolutiven Stammbäume stellt ein anderes evolutives Verwandtschaftsverhältnis dar als die anderen?

- Stammbaum 1
- Stammbaum 2
- Stammbaum 3
- Alle drei Stammbäume stellen das gleiche evolutive Verwandtschaftsverhältnis dar.
- Alle drei Stammbäume stellen unterschiedliche evolutive Verwandtschaftsverhältnisse dar.

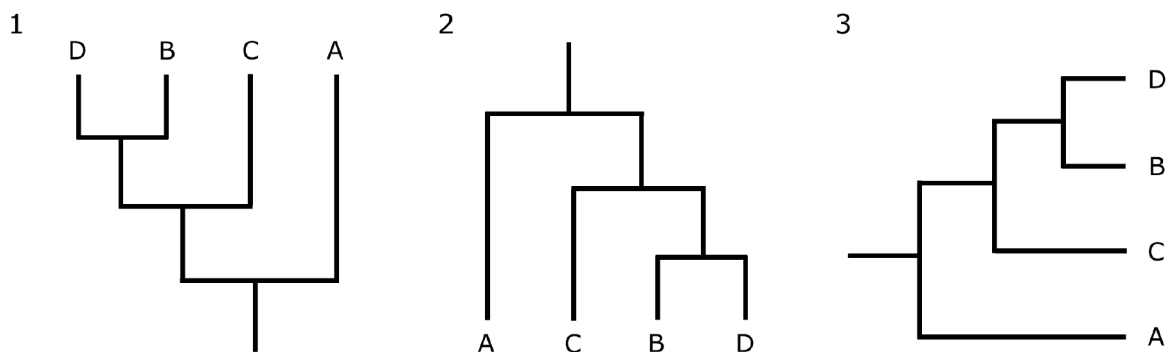
10)



Welcher der drei evolutiven Stammbäume stellt ein anderes evolutives Verwandtschaftsverhältnis dar als die anderen?

- Stammbaum 1
- Stammbaum 2
- Stammbaum 3
- Alle drei Stammbäume stellen das gleiche evolutive Verwandtschaftsverhältnis dar.
- Alle drei Stammbäume stellen unterschiedliche evolutive Verwandtschaftsverhältnisse dar.

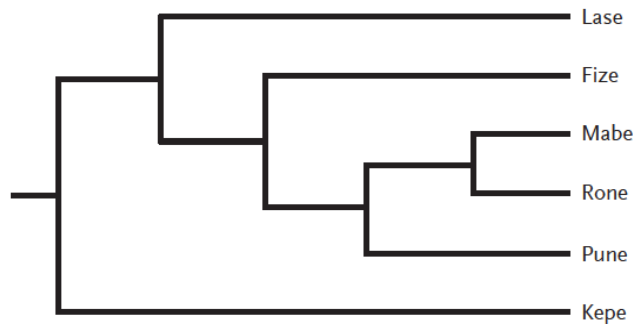
11)



Welcher der drei evolutiven Stammbäume stellt ein anderes evolutives Verwandtschaftsverhältnis dar als die anderen?

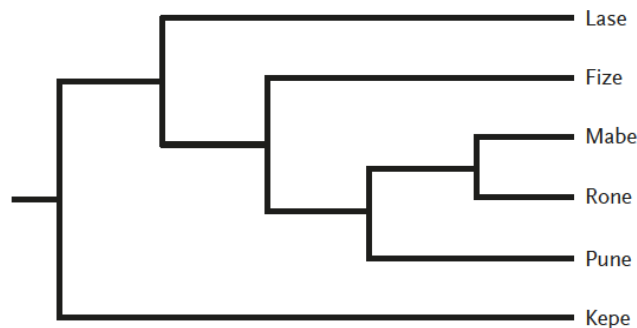
- Stammbaum 1
- Stammbaum 2
- Stammbaum 3
- Alle drei Stammbäume stellen das gleiche evolutive Verwandtschaftsverhältnis dar.
- Alle drei Stammbäume stellen unterschiedliche evolutive Verwandtschaftsverhältnisse dar.

Vielen Dank für Ihre Teilnahme ☺

13.2. Anhang 2: Aufgabe *Zweck1**Zweck1* Einarbeitung

Die Analyse verschiedener genetischer Marker ergab diesen Stammbaum. Eine zusätzliche morphologische Analyse zeigt, dass nur Lase, Pune und Kepe Streifen haben. Alle anderen gezeigten Taxa haben nur Punkte.

Was sehen Sie auf der Abbildung? Sagen Sie alles, was Sie denken.

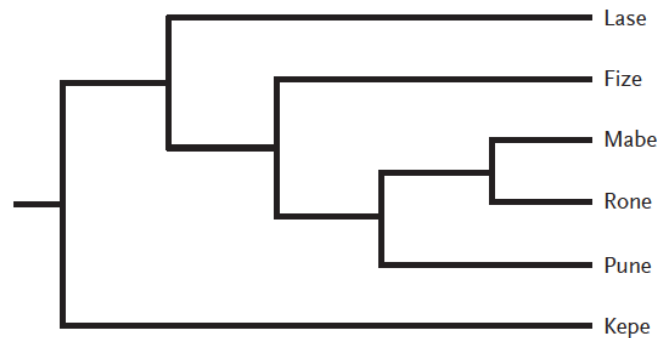
Zweck1 medialer Modelleinsatz

Die Analyse verschiedener genetischer Marker ergab diesen Stammbaum. Eine zusätzliche morphologische Analyse zeigt, dass nur Lase, Pune und Kepe Streifen haben. Alle anderen gezeigten Taxa haben nur Punkte.

Welche Stelle markiert den letzten gemeinsamen Vorfahren von Rone und Fize? Beschreiben Sie die Stelle.

Lösung: MRCA von Fize, Mabe, Rone und Pune (Knotenpunkt oder Linie davor).

Zweck1 methodischer Modelleinsatz



Die Analyse verschiedener genetischer Marker ergab diesen Stammbaum. Eine zusätzliche morphologische Analyse zeigt, dass nur Lase, Pune und Kepe Streifen haben. Alle anderen gezeigten Taxa haben nur Punkte.

Welche Hypothese kann über das Muster des letzten gemeinsamen Vorfahren aller gezeigten Taxa unter Anwendung des Sparsamkeitsprinzips aufgestellt werden?

Zu prüfende

Hypothese a

Hypothese „kein Muster“:

Der letzte *MRCA* aller gezeigten Taxa hatte kein Muster. Dann sind Streifen dreimal unabhängig voneinander entstanden (Lase, Kepe bzw. Pune). Bei Pune haben sich Streifen nachträglich aus Punkten entwickelt. Punkte sind einmal entstanden (Fize, Pune und (Mabe + Rone)).

→ 4 Ereignisse

Zu prüfende

Hypothese b

Hypothese „Punkte“:

Der letzte *MRCA* aller gezeigten Taxa hatte Punkte. Dann sind dreimal Streifen entstanden (Lase, Pune, Kepe) → 3 Ereignisse

Zu prüfende

Hypothese c

Hypothese „Streifen“:

Der letzte *MRCA* aller gezeigten Taxa hatte Streifen. Dann sind Punkte zweimal unabhängig voneinander entstanden (Fize und (Mabe + Rone))

→ 2 Ereignisse

Fazit

Hypothese „Streifen“ ist am sparsamsten.

Alternative

Hypothese

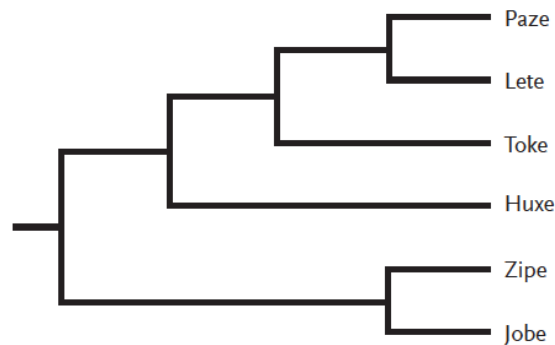
Hypothese „Punkte Alternativ“:

Entstehungsereignis von Streifen beim *MRCA* von Fize, Mabe, Rone, Pune und eine spätere erneute Merkmalsänderung zu Streifen bei Pune.

→ 2 Ereignisse

13.3. Anhang 3: Aufgabe Zweck2

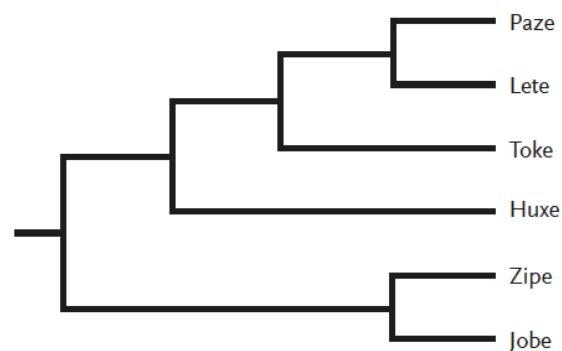
Zweck2 Einarbeitung



Die Analyse verschiedener genetischer Marker ergab diesen Stammbaum. Eine zusätzliche morphologische Analyse zeigt, dass nur Paze, Lete und Zipe Beinpanzer haben. Weder die anderen Taxa noch zahlreiche gefundene Fossilien weisen Beinpanzer auf.

Was sehen Sie auf der Abbildung? Sagen Sie alles, was Sie denken.

Zweck2 medialer Modelleinsatz

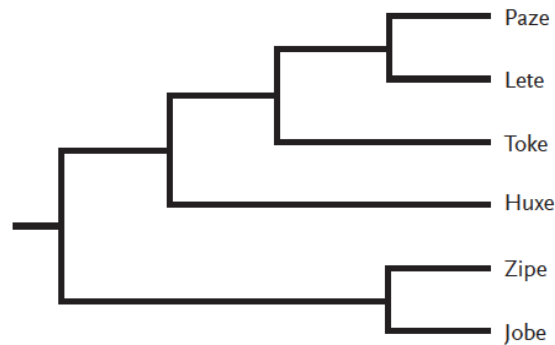


Die Analyse verschiedener genetischer Marker ergab diesen Stammbaum. Eine zusätzliche morphologische Analyse zeigt, dass nur Paze, Lete und Zipe Beinpanzer haben. Weder die anderen Taxa noch zahlreiche gefundene Fossilien weisen Beinpanzer auf.

Welche Stelle markiert den letzten gemeinsamen Vorfahren von Jobe und Huxe? Beschreiben Sie die Stelle.

Lösung: MRCA aller gezeigten Taxa (Knotenpunkt oder Linie kurz davor).

Zweck2 methodischer Modelleinsatz



Die Analyse verschiedener genetischer Marker ergab diesen Stammbaum. Eine zusätzliche morphologische Analyse zeigt, dass nur Paze, Lete und Zipe Beinpanzer haben. Weder die anderen Taxa noch zahlreiche gefundene Fossilien weisen Beinpanzer auf.

Welche Hypothese kann über die Entstehung des Merkmals Beinpanzer bei Paze, Lete und Zipe aufgestellt werden? Begründen Sie Ihre Antwort.

Zu prüfende

Hypothese „Beinpanzer Synapomorphie“

Hypothese a

Hätte der *MRCA* von Paze, Lete und Zipe Beinpanzer gehabt, wäre folgendes passiert: 1. Entstehung Beinpanzer, 2. Verlust bei Toke, 3. Verlust bei Huxe, 4. Verlust bei Jobe

→ 4 Ereignisse

Zu prüfende

Hypothese „Konvergenz“:

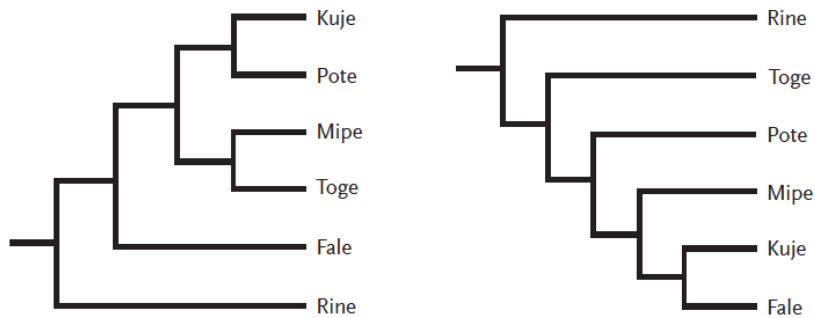
Hypothese b

Beinpanzer haben sich beim *MRCA* von Paze und Lete und bei Zipe jeweils unabhängig entwickelt.

→ 2 Ereignisse

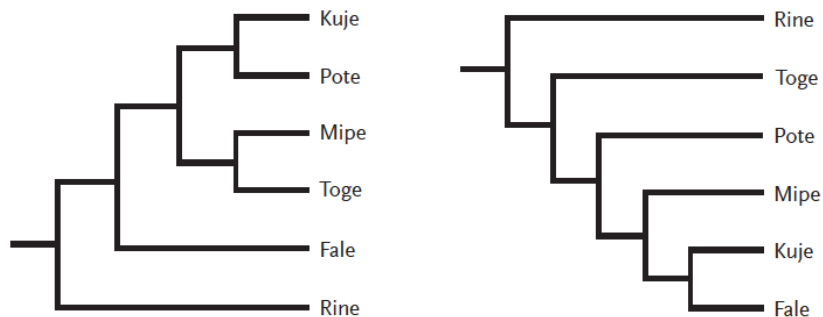
Fazit

Eine konvergente Entwicklung des Merkmals Beinpanzer ist am sparsamsten. Beinpanzer sind 2x unabhängig voneinander entstanden.

13.4. Anhang 4: Aufgabe *Alt1**Alt1* Einarbeitung

Nur Kuje und Mipe haben Schneidezähne. Fossilfunde führten zu der Hypothese, dass Vorfahren von Fale auch Schneidezähne hatten. Bei rezenten Fale sind diese nicht vorhanden.

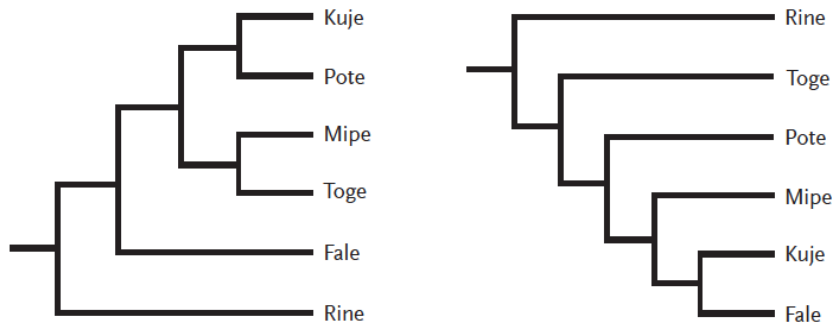
Was sehen Sie auf den Abbildungen? Sagen Sie alles, was Sie denken.

Alt1 medialer Modelleinsatz

Nur Kuje und Mipe haben Schneidezähne. Fossilfunde führten zu der Hypothese, dass Vorfahren von Fale auch Schneidezähne hatten. Bei rezenten Fale sind diese nicht vorhanden.

Welche Unterschiede und Gemeinsamkeiten weisen die Stammbäume auf? Vergleichen Sie sie miteinander.

Alt1 methodischer Modelleinsatz



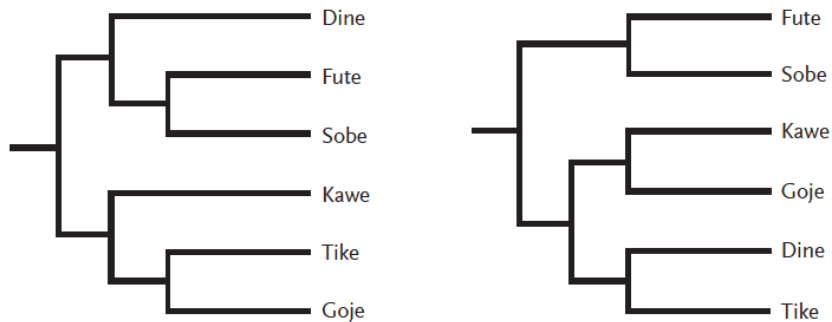
Nur Kuje und Mipe haben Schneidezähne. Fossilfunde führten zu der Hypothese, dass Vorfahren von Fale auch Schneidezähne hatten. Bei rezenten Fale sind diese nicht vorhanden.

Welcher der beiden Stammbäume unterstützt die genannte Hypothese unter Anwendung des Sparsamkeitsprinzips? Begründen Sie Ihre Antwort.

| | |
|---|--|
| <i>Passung DpV 1 (links) mit Hypothese</i> | Hätte der <i>MRCA</i> von Kuje, Mipe und Fale Schneidezähne gehabt, hätten sie sich jeweils bei Fale, Toge und Pote zurückentwickelt → 1 Entstehungs- und 3 Verlustereignisse, also insgesamt 4 Ereignisse |
| <i>Passung DpV 2 (rechts) mit Hypothese</i> | Mipe, Kuje und Fale bilden ein Monophylum. Nur in diesem Monophylum kommen Schneidezähne vor. Bei Fale haben sie sich zurückentwickelt → 1 Entstehungs- und 1 Verlustereignis, also insgesamt 2 Ereignisse |
| <i>Fazit</i> | DpV 2 erklärt die Hypothese am sparsamsten. |

13.5. Anhang 5: Aufgabe Alt2

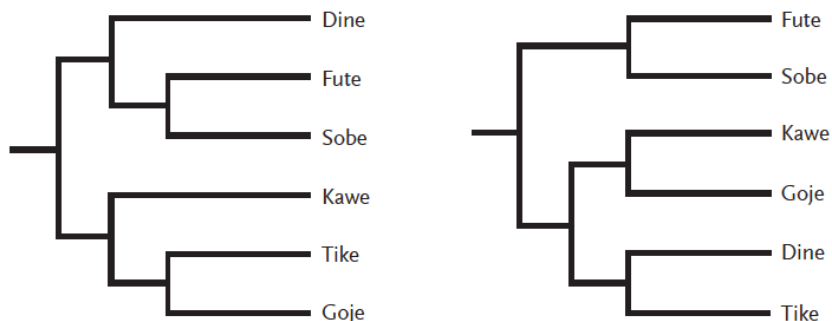
Alt2 Einarbeitung



Einer der Stammbäume beruht auf morphologischen, der andere auf genetischen Daten. Die morphologischen Daten deuten darauf hin, dass nur Kawe und Goje ein Monophylum mit ihrem letzten gemeinsamen Vorfahren bilden. Die genetischen Daten führen zu einer anderen Hypothese über dieses Monophylum.

Was sehen Sie auf den Abbildungen? Sagen Sie alles, was Sie denken.

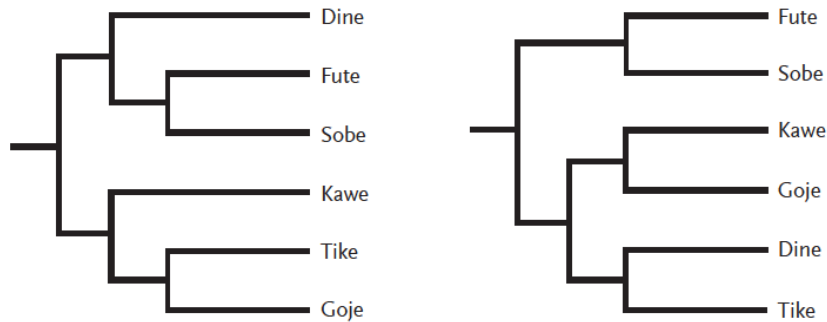
Alt2 medialer Modelleinsatz



Einer der Stammbäume beruht auf morphologischen, der andere auf genetischen Daten. Die morphologischen Daten deuten darauf hin, dass nur Kawe und Goje ein Monophylum mit ihrem letzten gemeinsamen Vorfahren bilden. Die genetischen Daten führen zu einer anderen Hypothese über dieses Monophylum.

Welche Unterschiede und Gemeinsamkeiten weisen die Stammbäume auf? Vergleichen Sie sie miteinander.

Alt2 methodischer Modelleinsatz



Einer der Stammbäume beruht auf morphologischen, der andere auf genetischen Daten. Die morphologischen Daten deuten darauf hin, dass nur Kawe und Goje ein Monophylum mit ihrem letzten gemeinsamen Vorfahren bilden. Die genetischen Daten führen zu einer anderen Hypothese über dieses Monophylum.

Welche Hypothese über das Monophylum, das Kawe und Goje enthält, kann mithilfe der genetischen Daten aufgestellt werden? Begründen Sie Ihre Antwort.

Passung DpV 2

*(rechts) mit Morpho-
Hypothese*

Im rechten DpV bilden nur Kawe und Goje mit ihrem *MRCA* ein Monophylum → morphologisch

*Passung DpV 1 (links)
mit Morpho-
Hypothese*

Im linken DpV ist Tike Teil des Monophylums;
Ausschlussverfahren: Wenn der rechte DpV der morphologische ist, ist der linke der genetische

*Ableiten der Gen-
Hypothese*

Laut den genetischen Daten bildet Tike ein Monophylum mit Kawe, Goje + *MRCA*

13.6. Anhang 6: Transkriptionsregeln nach Dresing und Pehl (2018)

Die Transkription der verbalen Daten des lauten Denkens (CTA) erfolgt nach Dresing und Pehl (2018) mithilfe eines einfachen Transkriptionssystems:

Benennung des Transkripts:

Das Transkript wird identisch zur zugehörigen Videodatei im Muster „18_01_01_V2_Exp_P01_Alt1_3_3“ benannt. Beginn des Transkripts ist die erste verbale Äußerung der_in des Proband_in.

Absätze und Zeitmarken

Nach jedem gesprochenen Satz wird mit der Transkriptionsfunktion in MAXQDA ein Absatz und somit eine Zeitmarke eingefügt.

Transkriptionsregeln

- Wörtliche Transkription mit Übertragung von Dialekten ins Hochdeutsche. Z.B. „Ick och“ wird zu „Ich auch“.
- Ausschreiben von Verkürzungen. Z.B. „Ich seh da n Problem“ wird zu „Ich sehe da ein Problem“
- Beibehalten der grammatischen Struktur, auch wenn sie nicht der Standardsprache entspricht (um Satzstruktur beizubehalten).
- Werden Sätze oder Wörter abgebrochen, wird dies mit einem / direkt hinter dem abgebrochenen Wort(fragment) dargestellt. Z.B. „Das ist ein Mono/ ein Monophylum“, „Das Monophylum besteht aus/ es enthält drei Taxa“. Dadurch wird die Erklärungsstruktur erhalten.
- Pausen werden entsprechend ihrer ungefähren Länge in Klammern notiert. Ein (.) bis drei (...) Sekunden mit der entsprechenden Anzahl von Punkten. Längere Pausen unter Angabe der Länge in Sekunden T: (T).
- Unverständliche Passagen oder Begriffe werden mit (unv.) transkribiert.
- Äußerungen des Überlegens oder Bestätigens werden transkribiert. Überlegende Äußerungen wie ähm, hmm werden als „ähm“ transkribiert. Bestätigungen oder andere Äußerungen werden in Klammern interpretiert: mhm (bestätigend).
- Die sprechende Person wird mit P (Proband_in) oder V (Versuchsleiterin) markiert.
- Zahlen von null bis zwölf werden ausgeschrieben, ab 13 in Ziffern angegeben.

- Englische (Fach)Begriffe werden der deutschen Rechtschreibung und Grammatik entsprechend transkribiert, z.B. *paraphyletic Group*, *Most Recent Common Ancestor*.

13.7. Anhang 7: Kodierleitfaden

Struktur der Aufgaben

Die Proband*innen haben je zwei Aufgaben zu den Teilkompetenzen *Alternative Modelle* bzw. *Zweck von Modellen* bearbeitet, die in jeweils 3 Teile gegliedert sind (Einarbeitung, medialer Modelleinsatz, methodischer Modelleinsatz). Insgesamt bearbeitete jede:r Proband:in 12 Aufgabenteile, je 6 pro Teilkompetenz. Jeder Aufgabenteil entspricht einem Transkript. Der hierfür relevante letzte Teil der Bezeichnung des Transkripts setzt sich folgend zusammen:

Teilkompetenz_Aufgabe_Aufgabenteil_Folie (Folie ist eine Bezeichnung, die nur für den Einsatz am ET relevant ist und hier keine weitere Bedeutung hat). Beim Aufgabenteil steht 1 für Aufgabenverständnis, 2 für medial, 3 für methodisch. „Alt1_2_3“ entspricht also der Teilkompetenz *Alternative Modelle* Aufgabe 1 medialer Modelleinsatz.

- Aufgaben Teilkompetenz *Alternative Modelle* 1 bzw. 2 (*Alt1* bzw. *Alt2*)
 - Einarbeitung (Alt1_1_3 bzw. Alt2_1_3)
 - Medialer Modelleinsatz (Alt1_2_3 bzw. Alt2_2_3)
 - Methodischer Modelleinsatz (Alt1_3_3 bzw. Alt2_3_3)
- Aufgaben Teilkompetenz *Zweck von Modellen* 1 bzw. 2 (*Zweck1* bzw. *Zweck2*)
 - Einarbeitung (Zw1_1_3 bzw. Zw2_1_3)
 - Medialer Modelleinsatz (Zw1_2_3 bzw. Zw2_2_3)
 - Methodischer Modelleinsatz (Zw1_3_3 bzw. Zw2_3_3)

Gesamtbezeichnung der Transkripte: Datum der Versuchsdurchführung, Proband:innennummer, Versuchsversion, Teilkompetenz, Aufgabe, Aufgabenteil, Folie; 18_01_05_P01_V2_Alt2_3_3 entspricht also dem 05.01.2018, an dem Proband:in Nr. 01 Version 2 bearbeitet hat und zwar die Aufgabe 2 zur Teilkompetenz *Alternative Modelle* im methodischen Modelleinsatz.

Allgemeine Kodierregeln:

- Das Kodiersystem ist in **fünf Teile** gegliedert:
 1. DpV-Bestandteile (Fachlich: Verwandtschaft, Merkmale, Zeit, Oberflächenmerkmale)
 2. Tree-Reading-Fähigkeiten (Wissenschaftlich: Anwendung von DpV)
 3. Bezug (nur bei Teilkompetenz *Alternative Modellen*)
 4. Modellkompetenz
 5. Sonstiges

- Jeder Teil des Kodiersystems wird separat auf das Transkript (1 Kodierzyklus pro Teil) angewendet. D.h. man geht das Transkript für jeden Teil durch und kodiert nur gemäß der jeweiligen „Brille“. So soll vermieden werden, dass Subkategorien vergessen werden. Sinneinheiten können je nach Subkategorie variieren:
 - DpV-Bestandteile, Tree-Reading-Fähigkeiten und Bezug (bei Teilkompetenz *Alternativen Modelle*) haben die gleichen Sinneinheiten. D.h. jede Information, die im Teil DpV-Bestandteile kodiert wird, wird auch in den anderen beiden Teilen kodiert.
 - Aussagen, die sich auf ein oder mehrere DpV beziehen aber nicht in die Teile DpV-Bestandteile und Tree-Reading-Fähigkeiten kodiert werden, werden bei Bezug kodiert („der rechte Stammbaum ist der morphologische“).
 - Modellkompetenz und Sonstiges werden jeweils unabhängig von anderen Teilen kodiert und können einzelne Aussagen bis ganze Abschnitte des Transkripts umfassen.
- **Sinneinheiten:**
 - Sinneinheiten mit dem Cursor vom ersten bis zum letzten Zeichen markiert. Kein Leerzeichen am Ende mitmarkieren.
 - *Äh, also, so*, und ähnliches werden nicht mitkodiert, wenn sie an eine Sinneinheit angrenzen, wenn:
 - sie durch ein Komma abgetrennt sind und keine sinnstiftende Eigenschaft haben „Also, klar es gibt auch ein Phylum, (.) äh, ein, eine (.) eine, ja, monophyletische Gruppe, [...]“ (P23, Alt2_3_3)
 - ein *und* o.ä. nicht die darauffolgende Sinneinheit einleitet im Sinne einer Verknüpfung zweier Aussagen, sondern ohne Zusammenhang dort steht: „(14) Und, auf jeden Fall gibt es da zwei Gruppen“ (fi)
 - Abbrüche oder Wiederholungen von Wörtern
 - Vor einer Sinneinheit:
„Die, (.) äh, (...) die Topologie ist natürlich immer eine Hypothese,“ (P23, 2_3_3). Hier wird erst kodiert, wenn der Satz einmal am Stück gesagt wird. Die abgebrochenen Teile werden nicht kodiert. Das gilt auch für abgebrochene Aussagen, die noch nicht allein kodiert werden können: „Die beiden/ äh (..) Die beiden Bäume zeigen beide vier Gruppen“ (fi)
 - In einer Sinneinheit:
Ändern Wiederholungen oder Abbrüche von Sätzen nicht den Inhalt der Sinneinheit, werden sie mitkodiert: „die beiden bilden äh bilden dann eine Gruppe“ (fi); „Mipe und To/ Toge bi/ äh sind eine Gruppe“ fi
 - Einleitende Hauptsätze werden mitmarkiert („Ja, also kann man sagen, dass Kawe und Goje (.) enger mit Dine und Tike verwandt sind als mit Fute und Sobe (..)“ P29, Alt2_3_3)
 - Pausen kürzer als 5 sec werden jeweils mit der vorhergehenden Aussage markiert.
 - Teilt eine Pause ≥ 5 sec eine Aussage, wird nur der Teil der Aussage vor oder nach der Pause kodiert, der den Kern der Aussage enthält („Deshalb bilden (10) Kuje und Mipe eine Gruppe“ (fi))
 - Je nach Subkategorie können unterschiedlich lange Sinneinheiten kodiert werden (siehe Kodierregeln).
 - Wenn nicht eindeutig erkennbar ist, worauf sich etwas bezieht („eins, zwei, drei“ P29, Alt1_3_3) nicht kodieren

- Abgebrochene Aussagen ohne eindeutig erkennbaren Inhalt nicht kodieren ([Beginn des Transkripts] „Zwei Bäume zu sehen bekommen“ (P23, Alt2_1_3))
 - Einzelne Ausdrücke wie „äh (..)“ oder „so (..)“ oder „also,“ ohne direkte Zugehörigkeit zum Hauptsatz am Rande einer Sinneinheit werden nicht mit kodiert.
- Das gesamte Kategoriensystem existiert doppelt: Einmal als **Entspricht Instruktion** und einmal **Entspricht anderer Instruktion_PRÄ**. Falls in einem Transkript, das zur Aufgabenverständnis- oder medialen Instruktion gehört, schon die methodische Instruktion antizipiert und bearbeitet wird, werden die betreffenden Sinneinheiten in **Entspricht anderer Instruktion_PRÄ** kodiert. So kann in der anschließenden Analyse sichergestellt werden, dass sich bestimmte Aussagen auf eine bestimmte Instruktion beziehen. Dies gilt nicht, wenn im methodischen Modelleinsatz verglichen und beschrieben wird. Dies wird in **Entspricht Instruktion** kodiert.
 - Beispiel *Alt1*: Wenn in Aufgabenverständnis oder medial bereits Merkmale zugeordnet werden, die über die Informationen aus dem Aufgabenstamm hinausgehen. Wenn anschließende Beschreibungen/ Vergleiche sich auf diese Merkmale beziehen:
 - Beispiel *Alt2*: In Aufgabenverständnis oder medial werden die Datenquellen (morphologisch und genetisch) den DpV zugeordnet, obwohl dies in der Instruktion nicht verlangt wird. Diese Zuordnung wird in **Entspricht anderer Instruktion_PRÄ** kodiert. Wenn anschließende Beschreibungen oder Vergleiche der Untermauerung/Rückversicherung zur Zuordnung dienen. Im Ankerbeispiel beziehen sich alle grau unterlegten Aussagen auf die Zuordnung der Datenquelle. Nach der 37 sec-Pause erfolgt eine Rückversicherung der Zuordnung, daher wird auch diese in PRÄ kodiert.
 - „Na dann müsste ja der rechte der morphologische sein (.) und der linke müsste der genetische (.) Stammbaum sein (...) wenn ich das jetzt richtig verstanden habe (4) Weil der rechte ist ein Ast, wo sich halt Kawe und Goje befinden (..) nebeneinander (..) und beim (.) linken, wahrscheinlich dem genetischen Stammbaum, (.) oder der auf den genetischen Daten beruht, (.) ist da noch Tike mit dabei (...) ja (4) Was sehen Sie auf den Abbildungen, ja (12) Mehr fällt mir nicht ein (37) Aber selbst beim rechten (.) Stammbaum (...) da sind ja Kawe, Goje, Dine und Tike (...) naja, die gehen ja mit den anderen auch alle auf einen gemeinsamen Vorfahren (.) zurück,“ (P11, Alt2_1_3)
 - Explizite Aufgabenlösung methodisch PRÄ: Wird in Aufgabenverständnis oder medial eine explizite Aussage dazu gemacht, welches DpV sparsamer ist bzw. dass im genetischen DpV Tike zum Monophylum gehört, wird dies in PRÄ kodiert. Wird Tike nur als Teil eines Vergleichs/einer Beschreibung zugeordnet und dieser Vergleich ist nicht Teil der Zuordnung der Datenquelle, wird dies nicht als PRÄ kodiert.
 - in Aufgabenverständnis oder medial:
 - „Deshalb würde ich sagen, dass im genetischen Baum Tike noch mit dazu gehört“ (fi) → PRÄ
 - „Im linken Stammbaum bilden Goje und Tike eine Gruppe, dann kommt Kawe dazu“ (fi; kein Bezug zu Datenquelle) → Entspricht Instruktion

- Wird in PRÄ kodiert, wird in „Bezug_PRÄ“ als Kommentar vermerkt, welcher Aspekt der methodischen Aufgabe adressiert wurde (Zuordnung Datenquelle, Aufstellen Hypothese; Zuordnung Merkmale, sparsameres DpV auswählen).
- Die Kodierregeln für PRÄ entsprechen bis auf die oben genannten Ergänzungen denen der „Entspricht Instruktion“-Teile
- Codes innerhalb eines Teils sind distinkt, können also pro Sinneinheit nur einmal vergeben werden. Eine Sinneinheit kann entweder nur in Entspricht Instruktion oder Entspricht anderer Instruktion_PRÄ vergeben werden.
- Die globale Lösung einer TA (korrekt, teilweise korrekt, falsch, unvollständig) wird als Memo für das ganze Transkript in der Dokumentenliste vermerkt, nicht als Code.
- Sprechpausen:
 - Pausen < 5 sec
 - am Anfang eines Transkripts mit dem ersten Satz mitkodieren
 - innerhalb eines Satzes mit dem vorhergehenden Teil des Satzes mitkodieren („Also, die Hypothese ist, dass die Vorfahren von Fale auch Schneidezähne hatten (..)“ „und welcher der beiden Stammbäume unterstützt die genannte Hypothese [...]“ (P11_Alt_1_3_3))
 - Füllwörter nach Pausen <5 sec werden grundsätzlich mit dem folgenden Satz(abschnitt) kodiert.
 - Pausen > 5 sec
 - werden extra kodiert.
- Werden zwei Pausen nur von einem Füllwort unterbrochen, werden sie inkl. Füllwort als eine lange Pause kodiert (Zusammengesetzte „Superpausen“: (...) hm (26) hm (15) als Summe der Pausendauer in Sekunden). Dies gilt auch für Pausen <5 sec.
- Einsatz von Fachwörtern/implizite Fachbegriffe: Wenn aus dem Kontext ohne großen Interpretationsaufwand deutlich wird, dass die Person z.B. von einer Klade spricht, aber einen anderen/falschen oder gar keinen Fachbegriff nennt, wird dies trotzdem als Klade kodiert. Den genannten Terminus bitte als Kommentar.
- Die jeweiligen Aufgabenblätter (siehe Anhang) sollten während der Kodierung miteinbezogen werden, um das Gesprochene nachzuvollziehen und ggf. einen falschen Gebrauch von Fachtermini festzustellen bzw. Aussagen zuzuordnen. Bsp. Alternative Modelle: Geht aus dem Abgleich mit dem Aufgabenblatt zweifelsfrei hervor, über welches DpV gesprochen wird, kann auch bei Nichtnennung eines Bezugs dieser kodiert werden.
- In den meisten Kategorien gibt es die Möglichkeit, Sinneinheiten als „unklar“ zu kodieren. Diese Codings werden anschließend genauer betrachtet und ggf. einem bestehenden Code zugeordnet oder für eine induktive Kategorienbildung genutzt.

a) DpV-Bestandteile

Klade

Sinneinheiten: Jede einzelne Information wird in diesem Teil des Kategoriensystems einzeln kodiert. Beispiel 1 (eine Graustufe bedeutet eine Sinneinheit): Kuje und Pote bilden eine Klade und Mipe und Toge auch. Beispiel 2: Erst spaltet sich Rine ab, dann Toge, dann Pote, dann Mipe und dann Kuje und Fale.

Länge der Sinneinheit (eine Graustufe bedeutet eine Sinneinheit, Wechsel der Graustufe dient nur der besseren Unterscheidung, hat keine inhaltliche Bedeutung): eine Sinneinheit umfasst den (Teil)Satz, in dem sie inhaltlich eingebettet ist. Also z.B. den Bezug (beide DpV, ein DpV). Aussagen ohne direkt zu kodierenden Inhalt („Dann haben wir allerdings Unterschiede innerhalb dieses/ (.) dieses Monophylums, was aus diesen fünf Taxa besteht (..)“) werden nicht mitmarkiert:

Also, wie gesagt, gemeinsam haben beide Stammbäume, dass Rine die basale Abzweigung ist, (.) äh, dass es also ein Monophylum gibt, was aus Kuje, Pote, Mipe, Toge und Fale (.) gebildet wird äh in beiden Stammbäumen (.) Dann haben wir allerdings Unterschiede innerhalb dieses/ (.) dieses Monophylums, was aus diesen fünf Taxa besteht (..) Einmal ist Fale, also im linken Stammbaum, ist Fale äh die basale Abzweigung innerhalb dieses Taxas (..) Taxons und äh einmal ist es Toge (..) Außerdem gibt es äh noch den Unterschied, dass im linken Stammbaum jeweils zwei terminale Taxa ein Schwestergruppenverhältnis bilden (.) Äh nämlich einmal Kuje und Pote (.) und einmal Mipe und Toge (.) Und dass diese beiden äh Schwestergruppen auch auf einen gemeinsamen Vorfahren zurück gehen (.) Während im (4) rechten Stammbaum (.) eben nur ein äh Schwestergruppenverhältnis zwischen Kuje und Fale besteht (..) Und äh (.) äh Pote und Mipe kein Schwestergruppenverhältnis (.) bilden (4) (P29, Alt1_2_3).

Wenn bei einem Vergleich zweier DpV eine Information für ein DpV genau genannt wird („und die andere Klade Kawe, Tike, Goje (..)“), beim anderen DpV aber nicht weiter ausgeführt wird („und äh, die beiden Kladen werden so überhaupt nicht gefunden auf der/ auf der rechten Seite.“ (P23, Alt2_2_3), wird nur die erste Aussage in 1. DpV-Bestandteile kodiert, bei den Tree-Reading-Fähigkeiten wird jedoch die erste als mehrere_DpV kodiert, die zweite nicht. In 3. Bezug werden beide Aussagen den DpV zugeordnet.

„Dine und Tike sind dieses Mal ein/ bilden ein Schwestergruppenverhältnis,“ ⊥ „wie auch in dem (.) in dem linken Stammbaum (..)“ (P29, Alt2_1_3) Hier wird auch die zweite Sinneinheit als Klade kodiert.

| | |
|---------------------|---|
| Code: | Klade |
| Definition | <p>Eine Klade/eine monophyletische Gruppe umfasst eine Stammlinie (<i>ancestral lineage</i>) und alle Taxa, die von ihr abstammen (Baum & Smith, 2012; Wiesemüller et al., 2003). Alle Taxa einer Klade sind gleich nah mit den anderen Taxa im DpV verwandt. Eine Klade kann aus einem Taxon und seinem <i>MRCA</i> bestehen aber auch aus x Taxa, solange sie von einem <i>MRCA</i> abstammen. In dieser Studie kann aufgabenbedingt eine Klade also aus 1 bis 6 Taxa bestehen. Verschieden umfangreiche Kladen können ineinander verschachtelt (<i>nested</i>) sein (z.B. Gregory, 2008). Das Bestimmen von Kladen (Halverson, 2011; Halverson & Friedrichsen, 2013; Novick & Catley, 2016; Schramm et al., 2019), verschachtelten Kladen (Novick & Catley, 2016) sowie der Vergleich von Kladen zwischen DpV (Halverson, 2011; Lemke et al., 2019) sind grundlegende TR-Fähigkeiten.</p> <p>Die Verschachtelung wird hier durch die Leserichtung kodiert. Die Topologie (Verzweigungsmuster) eines oder mehrerer DpV kann von den Taxa zur Wurzel hin beschrieben werden (Taxa zu Wurzel), von der Wurzel zu den Taxa (Wurzel zu Taxa) oder ohne erkennbare Leserichtung (Keine Leserichtung). Vorhergehende und nachfolgende Aussagen über Kladen sind hierfür zu berücksichtigen.</p> |
| Kodierregeln | <p>Mögliche andere Bezeichnungen einer Klade: Gruppe, Monophylum, monophyletische Gruppe, Schwestertaxa. Eine Klade kann durch den zu ihr führenden Ast (Stammlinie) oder durch den unmittelbar angrenzenden Knotenpunkt (<i>MRCA</i>) definiert werden. Somit kann eine Klade auch aus nur einem Taxon bestehen, wenn es das einzige Taxon ist, das aus einer Stammlinie hervorgeht. Wird auch kodiert, wenn die Stammlinie bzw. der <i>MRCA</i> nicht explizit genannt wird (jeweils eigener Code). Wenn für Taxa eine paraphyletische Beziehung postuliert wird (sie bilden nicht exklusiv eine Klade, sondern weitere Taxa sind Teil der Klade, zu der sie gehören), wird dies auch als einzelne Klade (mit Kommentar Paraphylum) kodiert.</p> <p>Die zu kodierende Sinneinheit ist immer im Kontext der vorhergehenden und nachfolgenden Aussagen zu beurteilen.</p> <p>Vorfahr-Regelung bei Zweck von Modellen → siehe Subcode Vorfahr Zweck von Modellen</p> |
| Sinneinheit | Siehe globale Beschreibung Code Verwandtschaftsgrad |

| Subcodes ²⁸ | Beschreibung | Kodierregeln | Ankerbeispiele ²⁹ |
|----------------------------------|--|--|---|
| Inhaltlich | Inhaltlich steht im Gegensatz zu oberflächlich. | Siehe „oberflächlich“ zur Abgrenzung. _mit MRCA nur, wenn der Vorfahre explizit als solcher genannt wird. Bezeichnungen wie „sind auf einem Ast“ werden in oberflächlich“ kodiert | |
| Taxa zu Wurzel _ohne MRCA | Beschreibt die Leserichtung ausgehend von den Taxa in Richtung Wurzel. | Wenn eine Schwestergruppe/Klade zu einer anderen Klade mit mehr als einem Taxon beschrieben wird. Wenn zwei Schwestertaxa mit zwei anderen Schwestertaxa zu einer Klade zusammengefasst werden und somit der nächste Knotenpunkt in Richtung Wurzel adressiert wird. Die vorher initial beschriebenen Schwestergruppen, die dann zueinander in Beziehung gesetzt werden, werden auch in diese Leserichtung eingeteilt. Daher immer erst den Zusammenhang der Beschreibung lesen und dann kodieren. | „Also Pote bspw. (..) äh ist jetzt äh (.) die Schwestergruppe z/ im/ im rechten T/ äh Stammbaum die Schwestergruppe zu einem Taxon, was aus Mipe, Kuje und Fale besteht (..)“ (P29, Alt1_2_3) „und dann kommt das Kawe dazu“ (P23, Alt2_1_3) |
| _mit MRCA | | | <i>[Mipe und Toge bilden ein Schwestergruppenverhältnis im (.) linken Stammbaum (..) und Kuje und Pote, (..)]</i> „(..) die zusammen auch auf einen gemeinsamen Vorfahren zurückgehen (...)“ P29, Alt1_1_3 |

²⁸ induktiv hinzugefügte Subcodes sind mit (ind) gekennzeichnet.

²⁹ Vorerst fiktive Ankerbeispiele sind mit (fi) gekennzeichnet.

| | | | |
|---|---|--|--|
| Wurzel zu Taxa _ohne MRCA | Beschreibt die Leserichtung ausgehend von der Wurzel zu den Taxa. | Wenn ausgehend von der Wurzel nacheinander Kladen beschrieben werden Indikator: Abspaltung, Auftrennung, abgehen | „Also in beiden Stammbäumen ist auf jeden Fall Rine die basale Abzweigung, \perp ³⁰ Dann gibt es also ein Monophylum, das aus allen übrigen Taxa besteht (...)“ (P29, Alt1_1_3) |
| _mit MRCA | | | „bei dem Vorfahren spaltet sich dann Rine ab \perp und ein Monophylum, das aus allen übrigen Taxa besteht (...)“ (fi) |
| keine Leserichtung _ohne MRCA | | Wenn keine Leserichtung erkennbar ist, weil z.B. über eine einzelne Klade gesprochen wird. Wenn eine Gruppe als paraphyletisch identifiziert wird. | „Auf dem linken Baum sind die in einem Monophylum mit Tike, (.)“ (P05, Alt2_1_3) „und beim (.) linken, wahrscheinlich dem genetischen Stammbaum, (.) oder der auf den genetischen Daten beruht, (.) ist da noch Tike mit dabei (...)“ (P11 Alt2_1_3) |
| _mit MRCA | | Wenn zwar das Verb eine Richtung vermuten lässt, es sich aber um eine einzelne Klade handelt (z.B. auf einen Vorfahren <i>zurückgehen</i> , siehe Ankerbeispiel) | „Kuje und Mipe sind in zwei Schwestergruppen (.)“ (P08, Alt1_1_3) „Da haben zwar Kuje, Mipe und Fale alle einen gemeinsamen Vorfahren“ [...] „Jedoch gehören sie keinem (.) ja, Monophylum an.“ (P08, Alt1_2_3) (2 Sinneinheiten mit MRCA und _ohne MRCA, jeweils mit Kommentar Paraphylum) „dann gehen ja Dine und Tike auch auf den gemeinsamen Vorfahren zurück“ (P11_Alt2_3_3) |

³⁰ \perp bedeutet den Beginn einer neuen Sinneinheit und wird bei Ankerbeispielen anstelle von Graustufen verwendet.

| | | | |
|----------------|---|---|---|
| Oberflächlich | Oberflächlich steht im Gegensatz zu inhaltlich. Oberflächliche Aussagen zu DpV stellen keine Verbindung von Bestandteilen zum Inhalt her. | Wenn Gruppen beschrieben werden, ohne mindestens ein beteiligtes Taxon konkret zu nennen (Bsp. 1a, 3c). Wenn Abspaltungen nicht mit Taxa kombiniert werden (Bsp. 2) oder nummeriert werden (Bsp. 2b). Wenn die räumliche Nähe von Taxa beschrieben wird (Bsp. 3a und b). | |
| Taxa zu Wurzel | Beschreibt die Leserichtung ausgehend von den Taxa in Richtung Wurzel. | <p>Wenn z.B. zwei (nicht genau benannte) Schwestertaxa mit zwei anderen Schwestertaxa zu einer Gruppe zusammengefasst werden und somit ein Knotenpunkt (wenn nicht als <i>MRCA</i> i. w. S. benannt) in Richtung Wurzel adressiert wird.</p> <p>Wenn vorher initial Schwestergruppen beschrieben wurden, die dann zueinander in Beziehung gesetzt werden, werden sie auch in diese Leserichtung eingeteilt.</p> <p>Bsp. 1b: Pote wird als nicht zugehörig zur Klade aus Kuje + Fale gesehen, sondern oberflächlich in Relation dazu gesetzt. Wird nicht bei Topologie kodiert, weil es sich hier um die relative Position eines Taxons zu anderen konkreten Taxa handelt.</p> | <p>Bsp. 1a: „Und jeweils zwei Taxa <u>└</u> und noch zwei andere Taxa <u>└</u> sind dann zusammen in einer Vierergruppe“ (fi, 3x dieser Code)</p> <p>Bsp. 1b: „und Pote (.) ist da ein bisschen weiter weg [von Kuje und Fale als Klade], ist/ ist weiter voneinander entfernt (.)“ (P13, Alt1_2_3)</p> <p>„<u>O</u>ben stehen Kuje und Pote (.)“ (P11, Alt1_2_3) [und die andere Gruppe ist Mipe und Toge (4) und die gehen alle auf einen (4) Vorfahren zurück]</p> |
| Wurzel zu Taxa | | Wenn eine Leserichtung von Wurzel zu Taxa erkennbar ist. Wenn Abspaltung nummeriert werden. Wenn Taxa/Kladen „alleine“ stehen und nicht in Zusammenhang mit anderen Kladen gebracht werden. | <p>Bsp. 2a: „Sondern es gibt immer wieder (..) Einzelabspaltungen“ (P11, Alt 1_2_3)</p> <p>Bsp. 2b: „die [Position von Fale] ist im rechten Stammbaum (...) die zweite Abspaltung sozusagen (..)“ (P29, Alt 1_1_3)</p> <p>„Also auf dem linken Kladogramm (...) steht Rine alleine (.)“ (P11, Alt1_2_3)</p> |

| | | | |
|----------------------------------|--|--|---|
| <p>keine Leserichtung</p> | | <p>Wenn Aussage zu „Nähe“ oder räumlicher Position getroffen wird, ohne einen Bezug zu Verwandtschaftsgrad herzustellen. Wenn Äste beschrieben werden ohne Bezug zu Evolutionslinie oder Vorfahr. Indikatoren: nebeneinander, auf einem Ast, näher zusammen.</p> | <p>Bsp. 3a: „Weil der rechte ist ein Ast, wo sich halt Kawe und Goje befinden (..) nebeneinander (..)“ (P11 Alt2_1_3)</p> <p>Bsp. 3b: „Mipe, Kuje und Fale (...) sind auf jeden Fall näher beieinander (..)“ (P11, Alt1_1_3)</p> <p>Bsp. 3c: „und (..) der größte Unterschied ist wahrscheinlich, dass oben/ (..) äh der obere Ast auf der rechten Seite (.) nur zwei Taxa beinhaltet (.) \perp und der untere vier (..)“ (P11, Alt2_2_3)</p> <p>Bsp. 3d: „sehe ich, dass sich Kuje und (.) Mipe (..) auf einem anderen Ast befinden (..) als Fale (..)“ (P11, Alt1_1_3)</p> |
|----------------------------------|--|--|---|

| | | | |
|---|--|---|---|
| Code: | Verwandtschaftsgrad | | |
| Definition | Die relative Verwandtschaft mehrerer Taxa (Sind Taxon A und B näher miteinander verwandt als Taxon A und C?) wird bestimmt, indem man feststellt, welche zwei Taxa mehr <i>MRCAs</i> miteinander teilen (z.B. Baum & Smith, 2013). Davon abzugrenzen ist die alternative Vorstellung, dass durch Knotenzählen der Verwandtschaftsgrad bestimmt werden kann, also wie viele Knoten zwischen zwei Taxa liegen (z.B. Blacquiere & Hoese, 2016; Halverson, 2011; Halverson & Friedrichsen, 2013; Gregory, 2008). Auch kommt die Bestimmung des Verwandtschaftsgrades durch räumliche Nähe der Taxa im DpV (Baum et al., 2005; Gregory, 2008; Meir et al., 2007; Nee, 2005; O’Hara, 1988; Blacquiere & Hoese, 2016; Halverson, 2011; Halverson & Friedrichsen, 2013) oder physische Ähnlichkeit bzw. Vorwissen (Halverson, 2011; Halverson & Friedrichsen, 2013; Blacquiere & Hoese, 2016; Schramm et al., 2019) als alternative Vorstellung vor. | | |
| Kodierregeln | Bei Formulieren wie „sind näher miteinander verwandt“, „sind nah miteinander verwandt“. Nicht bei „teilen sich einen gemeinsamen Vorfahren“ → Klade | | |
| Subcode | Beschreibung | Kodierregeln | Ankerbeispiele |
| fachlich | | | |
| relativ | s.o. | Auch wenn nur implizit darauf geschlossen werden kann, wie der Verwandtschaftsgrad ermittelt wurde. Für jede Gegenüberstellung von Taxa einzeln (siehe Ankerbeispiel) | „aber Tike und Goje sind (.) näher miteinander verwandt <u>└</u> als (.) Kawe (.) und Goje“ (P08, Alt2_3_3) |
| (ind) durch_Merkmale | | Wenn ein Grund für den unterschiedlichen Verwandtschaftsgrad genannt wird, z.B. allgemein morphologische oder genetische Unterschiede oder konkrete Merkmale, die sich auf die konkret in der Aufgabe genannten Merkmale bezieht. | „der [gemeinsame Vorfahre] sich zu Kawe halt schon genetisch stärker unterschieden hat u/ stärker unterschieden hat, als zwischen den beiden (4) (P13, Alt2_3_3) [da sich die Aussage auf ein genetisch erstelltes DpV bezieht] |
| Alternative Vorstellungen/ Leseweisen³¹ | | | |
| AVoLe_unklar | Für zu definierende AVoLe | | |
| Knotenzählen | AVoLe: Anzahl der Knoten zwischen zwei Taxa wird als Maß für Verwandtschaft angesehen. | Wenn Anzahl oder Menge (mehr oder weniger, gleich viele) von Knoten/ Vorfahren angeben | „Das heißt, (.) ja, die unterscheiden sich beide in den/ In den/ In den Vorfahren/ |

³¹ Im Kodierleitfaden wird noch von Alternativen Vorstellungen/ Leseweisen (AVoLe) gesprochen, dies ist äquivalent zur Bezeichnung *wissenschaftlich inadäquates Ablesen (TR_1b)*.

| | | | |
|--|---|---|---|
| | | | In der verschiedenen Anzahl der Vorfahren,“ (P20, Alt1_2_3) |
| Räumliche Nähe | AVoLe: Räumliche Nähe der Taxa wird als Maß für Verwandtschaft angesehen. | | „Und (.) Kuje/ also ja, Mipe, Kuje und Fale (...) sind auf jeden Fall näher beieinander (.) als auf der linken Seite“ (P11, Alt1_1_3) |
| Physische Ähnlichkeit | AVoLe: Physische Ähnlichkeit (bekannter Taxa) wird als Maß für Verwandtschaft angesehen. | | |
| Taxa stammen voneinander ab | | Wenn ein Taxon als Vorfahre eines anderen angegeben wird. | „Fale ist dabei direkter Vorfahr (.)“ (P20, Alt1_1_3) |
| Outgroup = Vorfahr | | Wenn die Outgroup als Vorfahre aller anderen Taxa gesehen wird. | |
| Kein Knotenpunkt = kein weiterer Vorfahr (Ind) | AVoLe | | „Links, äh, (.) links ist es so, dass, äh, Fale aus dem Vorfahr/ Aus einem Vorfahr direkt entstanden ist.“ (P20, Alt1_2_3) |
| Knotenpunkte haben unterschiedliche Bedeutung (ind) | AVoLe: manche Knotenpunkte werden als <i>MRCA</i> gesehen, andere als Zwischenstufen oder ohne Bedeutung (oberflächliche Aufgabelung) | Wenn Knotenpunkte nicht einheitlich als <i>MRCA</i> genutzt werden. | „an einer Gabelung, die direkt auf den Vorfahren (.) zuführt (...)“ (P11, Alt2_2_3) „im linken, äh, Fute und Sobe haben nochmal einen Zwischenvorfahren, äh, denselben wie Dine auch (.) hat.“ (P21, Alt2_2_3) |
| Gleiche Position Knoten = gleicher <i>MRCA</i> (ind) | AVoLe: oberflächlich gleich positionierte Knotenpunkte bei zwei DpV werden als dieselben <i>MRCA</i> gesehen. | | „aber, äh, Rine hat den gleichen Vorfahren, (.) äh, wie/ Also in beiden Stammbäumen den gleichen Vorfahren,“ (P21, Alt1_1_3) |
| Anagenese | AVoLe: Eine Spezies geht direkt in eine andere über und „verschwindet“ als eigene Spezies → keine Koexistenz möglich | | |

| | | | |
|---------------------|---|---------------------|---|
| Code: | (Ind) Outgroup | | |
| Definition | Ein oder mehrere Taxa, die <i>a priori</i> festgelegt werden: Gehören nicht zur <i>ingroup</i> (zu untersuchende Taxa, die eine Klade bilden) und dienen zur Wurzelung des DpV (Baum & Smith, 2013) | | |
| Kodierregeln | Nur kodieren, wenn ein Taxon explizit als Outgroup/ Außengruppe genannt wird | | |
| Subcode | Beschreibung | Kodierregeln | Ankerbeispiele |
| fachlich | s.o. | s.o. | „Also bei den einen/ bei beiden ist Rine die Außengruppe“ (P05, Alt1_1_3“ |

| | | | |
|---------------------|--|---------------------|--|
| Code: | Vorfahre (Teilkompetenz <i>Alternative Modelle</i>) | | |
| Definition | Ein Knotenpunkt oder Ast stellt den <i>MRCA</i> bzw. einen gemeinsamen Vorfahren dar. | | |
| Kodierregeln | Nur kodieren, wenn explizit ein Vorfahre verortet wird und der Fokus auf dem Vorfahren und nicht auf den Taxa liegt. | | |
| Subcode | Beschreibung | Kodierregeln | Ankerbeispiele |
| fachlich | s.o. | s.o. | „Der Vorfahre von Kuje, Mipe und Fale im linken Stammbaum ist der letzte gemeinsame Vorfahre von der großen Gruppe ab Fale“ (fi) |
| oberflächlich | | | „Der Vorfahre von Kuje, Mipe und Fale ist der zweite Knotenpunkt von links“ (fi) |

| | | | |
|--|---|---|--|
| Alternative Vorstellungen/ Leseweisen | | | |
| (ind) Abspaltungen_zählen | AVoLe: Entwicklung einzelner Taxa in zeitlicher Reihenfolge | Unterschied zu ältere und jüngere Taxa: Nummerieren von Abspaltungsereignissen mit Beimuessung von zeitlicher Bedeutung | „Die Pote wäre auf dem rechten Stammbaum (.) als dritte Art abgespalten worden ↓ und hier als eins, zwei, drei, vierte Art,“ (P12, Alt1_2_3) |
| (ind)_nur_MRCA_für_ein_Taxon | AVo: Ein Knotenpunkt ist nur der <i>MRCA</i> für ein abgehendes Taxon | | "der Vorfahre von Mipe hatte äh. Schneidezähne, hat die an Mipe weitergegeben und auch an den Vorfahren von Kuje und Fale" (P16, Alt 1_3_3) |

| | | | |
|---------------------|---|---------------------|-----------------------|
| Code: | Vorfahre (Teilkompetenz Zweck von Modellen) | | |
| Definition | Ein Knotenpunkt oder Ast stellt den <i>MRC</i> A bzw. einen gemeinsamen Vorfahren dar. | | |
| Kodierregeln | <p>Bezieht sich auf die Bearbeitung der medialen Instruktion von Zweck von Modellen „Welche Stelle markiert den letzten gemeinsamen Vorfahren von x und y. Beschreiben Sie die Stelle“.</p> <p>Vorfahr-Regelung (nur bei Zweck von Modellen): Vorfahr steht über Klade. Wenn ein Vorfahr inhaltlich anhand einer Klade verortet wird, wird in den Subcodes von Vorfahre_inhaltlich kodiert.</p> <ul style="list-style-type: none"> Die Subcodes von Klade werden verwendet, sofern <u>weitere</u> Bestandteile des DpV beschrieben werden. Dabei werden Leserichtungen von Kladen und/oder Vorfahren aufgegriffen, sofern sie direkt zusammenhängen. Ankerbeispiel: "Rone und Fize haben einen gemeinsamer Vorfahre, der auch, äh, Mabe und Pune mit einschließt [Vorfahre inhaltlich Taxa zu Wurzel] und äh, (.) die Außengruppe davon wäre die Lase (.) [Klade inhaltlich Taxa zu Wurzel]" (P23 Zweck1_2_3). Es wird eine lange SE Vorfahre kodiert, wenn Kladen <u>direkt</u> zur Beschreibung des Vorfahren genutzt werden: „und der gemeinsame Vorfahre ist eigentlich die Wurzel vom ganzen Baum, (.) weil erst dort äh spalten sich diese beiden langen monophyletischen Gruppen einmal zu Huxe, Toke, Lete, Paze und auf der anderen Seite Zipe und Jobe auf [Vorfahre_inhaltlich ohne Leserichtung] “ (p05, Zw2_2_3). „Fize ist jetzt der zweite von oben, [Klade oberflächlich] ↓Rone der Vierte von oben. [Klade oberflächlich] ↓Rone hat einen gemeinsamen Vorfahren mit Mabe [Vorfahre inhaltlich Ta>Wu] ↓ und die beiden zusammen haben einen gemeinsamen Vorfahren mit Pune. [Vorfahre inhaltlich Ta>Wu]↓ Um jetzt noch Fize mit reinzubekommen, müssen wir nochmal einen Vorfahren weiter zurück gehen, also einen Zweig des Stammbaums nochmal zurück. [Vorfahr oberflächlich Ta>Wu]“ P16, Zw1_2_3 | | |
| Subcode | Beschreibung | Kodierregeln | Ankerbeispiele |
| inhaltlich | <ul style="list-style-type: none"> Wenn ein Vorfahr anhand von Kladen oder Taxa beschrieben wird (spezifische Taxa, „alle Taxa/Lebewesen“, Stammbaum, Baum, Klade, Gruppe im Sinne von „alle Taxa“). E Es wird nichts kodiert, wenn nur die zwei Taxa aus der Instruktion genannt werden. Wird auch kodiert, wenn der Vorfahre als Knotenpunkt/Aufspaltung o.ä. bezeichnet wird, aber in Kombination mit spezifischen Taxa, „alle Taxa/Lebewesen“, Stammbaum, Baum, Klade, Gruppe im Sinne von „alle Taxa“ genannt wird: „Der Knotenpunkt, von dem der Baum ausgeht“ (fi). Wird auch kodiert, wenn nicht explizit das Wort Vorfahr genannt wird, allerdings aus dem Kontext erschlossen werden kann, dass es sich um die Beschreibung des Vorfahren handelt. Ankerbeispiel: „das ist äh der Ast der die monophyletische Gruppe von Fize, Mabe, Rone und Pune verbindet (..)“ (P05, Zweck1_2_3). Wird <u>nicht</u> kodiert, wenn ein Begriff wie „Ausgangspunkt“, „Wurzelknoten“ o.ä. alleine ohne weitere Ausführung genannt wird: „Das ist der Wurzelknoten.“ (fi) Immer per Kommentar vermerken, welche SE die explizite Beantwortung der Instruktion beinhaltet. | | |

| | | | |
|---------------------|---|---|--|
| _Taxa zu Wurzel | Beschreibt die Leserichtung ausgehend von den Taxa in Richtung Wurzel | <p>Wenn ausgehend von den Taxa nacheinander Vorfahren in Richtung Wurzel beschrieben werden. Wenn diese Vorfahren miteinander in Zusammenhang gebracht werden.</p> <p>Der initial beschriebene Vorfahr wird auch in diese Leserichtung eingeteilt. Daher immer Zusammenhang beachten. Greift auch Leserichtungen aus Vorfahre_oberflächlich auf.</p> | <p>Also Mabe, Pune und Rone haben ja den gleichen gemeinsamen, letzten Vorfahre \perp und dieser sozusagen (.) ist aus dem gemeinsamen, letzten Vorfahren von (.) Fize und eben den (..) Mabe, Rone, Pune Komplex entstanden</p> |
| _Wurzel zu Taxa | Beschreibt die Leserichtung ausgehend von der Wurzel in Richtung Taxa | <p>Wenn ausgehend von der Wurzel direkt nacheinander Vorfahren beschrieben werden. Wenn diese Vorfahren miteinander in Zusammenhang gebracht werden.</p> <p>Der initial beschriebene Vorfahr wird auch in diese Leserichtung eingeteilt. Daher immer Zusammenhang beachten.</p> <p>Indikator: Abspaltung, Auftrennung, abgehen Greift auch Leserichtungen aus Vorfahre_oberflächlich auf.</p> | <p>„das ist ein (.) gemeinsamer Vorfahr, der trennt sich auf in die Gruppe/ (.) in die Fünfergruppe und Kepe. \perp Dann kommt es zu einer Auftrennung (.) der Vierergruppe und Lase (..)und (.) das ist der gemeinsame Vorfahre (..) von wem nochmal? (.) Rone? (..) Und Fize. Genau und das ist der gemeinsame Vorfahre (..) von (.) Fize und Rone (...) [Vorfahre inhaltlich WuTa]“ (P07, Zw_1_2_3)</p> <p>„der eine ist der Ausgangs- (.) - vorfahr ist, ja, von dem alle sechs abstammen. \perp Dann die nächste Schnittstelle, wo sich dann ja Lase ab- (.) spalt/ gespalten hat, \perp und dann die (.) dritte Ebene quasi, wo Fize sich abgespalten hat und die anderen drei Arten Mabe, Rone und Pune.“ (P13, Zw1_2_3)</p> |
| _keine Leserichtung | | Ohne unmittelbar vorhergehende oder nachfolgende Beschreibung anderer Vorfahren. | „Der letzte gemeinsame Vorfahre von Rone und Fize äh. ist (...) an der Stelle, wo Fize die Abzweigung/ Also wo die |

| | | | |
|----------------------|---|---|--|
| | | Gegensatz zu oberflächlich: Wenn ein Knotenpunkt/Vorfahre/Stammart o.ä. in Bezug zu gesamtem Stammbaum/Baum/großer Klade/ganze Gruppe genannt wird (Ankerbeispiel 2) | Abzweigung ist zu Fize und dann nach unten zu Pune. (..)“ (P18, Zw1_2_3) „(.) Der letzte gemeinsame Vorfahre von Jobe und Huxe, (..) äh, ist der, (..) äh, (..) ja, der Ursprungs-, äh, -vorfahre von/ von dem gesamten Stammbaum (..)“ (P28, Zweck2_2_3) |
| oberflächlich | <ul style="list-style-type: none"> • Wenn ein Vorfahr ohne inhaltliche Angabe über Abzählen von Knotenpunkten, Ebenen o.ä. beschrieben wird. • Auch wenn nicht explizit das Wort Knotenpunkt genannt wird, allerdings aus dem Kontext erschlossen werden kann, dass ein Knotenpunkt gemeint ist (Abspaltung, Kreuzung, Stelle an der sich etwas aufteilt o.ä.). • Kodieren, wenn eine Ortsangabe gemacht wird: „der ganz links im Stammbaum. Im Gegensatz dazu inhaltlich „von dem der ganze Stammbaum ausgeht“ • Inhaltlich kodieren, sobald der Begriff Knotenpunkt o.ä. in Verbindung mit Stammbaum, Baum, alle Taxa o.ä. als <u>Ausgangspunkt</u> genannt wird → siehe Kodierregeln inhaltlich • Immer per Kommentar vermerken, welche SE die explizite Beantwortung der Instruktion beinhaltet. | | |
| _Taxa zu Wurzel | | Hier gelten bewegungsimplizierende Begriffe im eigentlichen Sinne: zurückgehen, nach links Auch ohne explizite Nennung anderer Vorfahren möglich, wenn durch Angaben wie „dritter Vorfahre“ eine Reihenfolge impliziert wird. Ausnahme „erster Vorfahre von links“ (dann ohne Leserichtung). Greift auch Leserichtungen aus Vorfahre_inhaltlich auf. | [vorher Kladen Ta>Wu] „Um jetzt noch Fize mit reinzubekommen, müssen wir nochmal einen Vorfahren weiter zurück gehen, also einen Zweig des Stammbaums nochmal zurück.“ (P16, Zw1_2_3) |
| _Wurzel zu Taxa | | Auch ohne explizite Nennung anderer Vorfahren möglich, wenn durch Angaben wie „dritter Vorfahre“ eine Reihenfolge impliziert wird. Ausnahme „erster Vorfahre von links“ (dann ohne Leserichtung) | „der gemeinsame Vorfahre beschreibt den, (.) mh (zustimmend), (.) dritten Punkt „geht man den, äh, Stamm oben, äh, äh, horizontal hoch, äh, geht dann nach |

| | | | |
|-------------------------|--|--|---|
| | | <p>Wenn der Verlauf der Linien ohne Zwischenstationen beschrieben wird, wird alles als eine SE kodiert (siehe zweites Ankerbeispiel).</p> <p>Wenn der Verlauf anhand von Knotenpunkte (oder äquivalente Bezeichnung wie Zweigstelle, Stelle) genannt wird, wird jede „Zwischenstation“ als eine SE kodiert (siehe drittes Ankerbeispiel).</p> <p>Greift auch Leserichtungen aus Vorfahre_inhaltlich auf.</p> | <p>rechts, geht dann nach unten (.) und geht dann nach rechts. So, da haben wir einen Vorfahren, dann gibt es, äh/ Das ist der letzte gemeinsame Vorfahre (..)“ (P15, Zw1_2_3)</p> <p>„der ist, wenn man vom Anfang des Stammbaums losgeht, der erste <u>Zweigstelle</u> nach oben (..) ↓ und die nächste <u>Zweigstelle</u> nach (.) unten (.) Das ist der (.) letzte gemeinsame Vorfahre von (.) Fize (.) und (..) Rone (...),“ [P08, Zweck1_2_3)</p> |
| _ohne Leserichtung | | <p>Wenn keine weiteren Vorfahren/Knotenpunkte beschrieben werden. Wenn keine näheren Angaben gemacht werden, welche Gruppe gemeint ist (wie z. B. konkrete Taxanamen oder „alle Taxa“).</p> <p>Sobald</p> | <p>„dass der gemeinsame Vorfahre von Jobe und Huxe, der also ganz links die erste Abzweigung“(fi)</p> |
| _vertikale Leserichtung | | <p>wenn die Leserichtung nicht Ta>Wu oder Wu>Ta entspricht (siehe zweites Ankerbeispiel), bitte als Kommentar abweichende Leserichtung vermerken“</p> | <p>Von oben nach unten: „das ist (..) ja, die/ von oben betrachtet die zweite Abspaltung.“ (p22, Zw1_2_3)</p> |

| Code: | Merkmale | | |
|--------------|--|---|--|
| Definition | Merkmale oder Apomorphien | | |
| Kodierregeln | Merkmale schlägt Kladen: Wenn ein Merkmal für eine ganze Klade genannt wird, wird nicht die Klade kodiert sondern das Merkmal ((fi) „Die Gruppe von Toge, Pote und Mipe weißt Schneidezähne auf“). | | |
| Subcode | Beschreibung | Kodierregeln | Ankerbeispiele |
| fachlich | | | |
| Taxa | Nennung von Merkmalen einzelner oder mehrerer Taxa | <p>Wenn Merkmale genannt werden, die nicht explizit aus dem Aufgabenstamm hervorgehen.</p> <p>Ausnahme: Bekannte Taxamerkmale werden genutzt, um Mechanismen (Konvergenz, Entstehung/Verlust, Parsimonie) zu begründen.</p> <p>Auch wenn diese Merkmale implizit aus dem Kontext geschlossen werden können (siehe Ankerbeispiel e).</p> <p>Wenn Merkmale als Aut- oder Synapomorphien benannt werden. (siehe Ankerbeispiel f)</p> <p>Auch kodieren, wenn Merkmale falsch zugeordnet werden, dann bitte Kommentar (siehe Ankerbeispiel g).</p> | <p>a) „was dazu führen würde, dass die komplette Gruppe ab Mipe Sch/ wahrscheinlich Schneidezähne hätte“ (P06, Alt1_3_3)</p> <p>b) „Also die anderen haben leider keine Schneidezähne, (..)“ (P13, Alt1_3_3)</p> <p>c) „Das heißt, dass (...) die/ (..) <u>zwei der äußersten</u> Taxa (.) Streifen haben (.) <u>und</u> das einzige, was halt (..) ja, interessant ist, ist, dass Fize auch Punkte hat, wo man vielleicht <u>annehmen würde</u>, dass es auch Streifen hat.“ (P06, Zweck1_3_3)</p> <p>d) „und dann (..) die Punkte (4) Mabe, Rone und (.) Fize (.) behalten haben“ (P28, Zweck1_3_3)</p> <p>e) „(.) <i>Genau, wenn Lase, Pune, Kepe Streifen haben (.) ist das Sparsamkeitsprinzip, dass der gemeinsamste Vorfahre von allen auch Streifen gehabt hat, weil Kepe (.) und Lase, die (.) am weitesten entfernten Linien sind und irgendwann hat sich zwischendrin dann (.) das</i></p> |

| | | | |
|---|--|---|--|
| | | | <p><i>Punktmuster evolviert</i>“ (P05, Zweck1_3_3)</p> <p>f) „einmal/ als Autapomorphie von Zipe <u>l</u> und einmal als Synapomorphie (..) von Paze und Lete, (.)“ (P24, Zweck2_3_3)</p> <p>g) „<i>Ich würde mal sagen, der hatte Streifen</i>, weil Streifen kommen überall vor.“ (P11, Zweck1_3_3)</p> |
| MRCA | Nennung von Merkmalen einzelner oder mehrerer MRCA | <p><u>Synonyme</u>: plesiomorphes/ anzestrales/ ursprüngliches Merkmal im Sinne von „Merkmal des MRCA aller dargestellten Taxa“.</p> <p>Wenn dasselbe Merkmal für denselben MRCA zweimal direkt hintereinander genannt wird, wird nur eine SE kodiert.</p> <p>Wenn es ohne unmittelbare Begründung genannt wird, bitte Kommentar, auf welche Aussage es sich bezieht bzw. dass es mehrfach genannt wurde.</p> | <p>„Der gleiche Vorfahre hatte dann wahrscheinlich auch die Schneidezähne“ (P21, Alt1_2_3)</p> <p>„Wenn man im Gegensatz dazu (...) Punkte als plesiomorphes Merkmal annehmen würde (..)“ (P14, Zweck1_3_3)</p> <p>„sodass äh. die Streifen auf jeden Fall ein Urmerkmal sind,“ (P18, Zweck1_3_3)</p> <p>„Deswegen würde ich (.) prinzipiell davon ausgehen, dass Streifen das ursprüngliche Merkmal sind, also der Vorfahre der kompletten Gruppe Streifen hatte“ (P06, Zweck1_3_3)</p> |
| Unspezifisch (ind) | | Wenn keine spezifischen Merkmale genannt werden. | |
| Alternative Vorstellungen/ Leseweisen | | | |
| (ind) Verlust_bei_Nichtgebrauch und Entstehung bei Gebrauch | AVoLe: Merkmale gehen verloren, wenn sie nicht mehr gebraucht werden bzw. entstehen, wenn sie gebraucht werden | | „und je nach dem ob sie gebraucht werden oder nicht, äh hat man sie |

| | | | |
|---------------------------------------|---|---|--|
| | | | noch wie Kuje und (.) Mipe, oder halt nicht (.)“ (Nov22, Alt1_3_3) |
| (ind) Individualebene | AVo: Vererbung auf Individualebene, Auftreten auf Individualebene statt auf Populationsebene. | | „man kann davon ausgehen, dass es rezessiv ist, da es nicht immer bei jedem auftritt, sondern halt zu bestimmten Gelegenheiten.“ (P17, Zweck1_3_3) |
| (Vorwissen) | AVoLe: Merkmale werden nicht vom DpV abgelesen sondern aus Vorwissen geschlossen | | |
| (Antizipation) | AVoLe: Merkmale werden nicht vom DpV abgelesen sondern antizipiert | | |
| Durchgehender Ast=keine Veränderungen | AVoLe: Es finden keine anderen Merkmalsveränderungen statt, wenn kein Lineage Splitting erfolgt. | Nur in Bezug auf Merkmale, sonst Topologie_AVoLe_Ast=rezent Taxon | |
| Höhere und niedrigere Taxa | AVoLe: Taxa mit früher entstandenem oder weniger Merkmalen sind weniger hoch entwickelt als Taxa mit später entstandenem/ mehr Merkmalen. | | |
| Veränderung nur an Knotenpunkten | AVoLe: Merkmalsveränderungen treten ausschließlich an Knotenpunkten auf | | |
| Similarity vs. relatedness | AVo: Taxa, die nah miteinander verwandt sind, müssen sich auch ähnlich sein. Allerdings ist dies nicht der Fall, da Taxa sich mit unterschiedlicher evol. Rate entwickeln (Baum & Smith, 2012, S. 114 | | „Die könnten also sehr, sehr ähnlich sein, (.) äh, sowohl rechts,“ (P21, Alt2_3_3) |

| Code: | Mechanismen | | |
|--------------------------|---|---|---|
| Definition | Evolutive Mechanismen wie konvergente Entwicklung von Merkmalen, Verlust- und Entstehungsereignisse (primär und sekundär), Sparsamkeitsprinzip. Das Sparsamkeitsprinzip ist zwar ein Prinzip der Datenauswertung, dennoch wird es hier eingeordnet, da ihm evolutive Prozesse wie die Anzahl von Entstehungs-/Verlustereignissen zugrunde liegen. | | |
| Kodierregeln | Siehe spezielle Kodierregeln | | |
| Subcode | Beschreibung | Kodierregeln | Ankerbeispiele |
| fachlich | | | |
| (ind) Entstehung_Verlust | Entstehungs- und Verlustereignisse von Merkmalen werden genannt oder beschrieben | <p>Für jedes Ereignis einzeln kodieren (Ankerbsp. a und b). Wenn In Kommentar vermerken, ob Verlust oder Entstehung</p> <p>Indikatoren: Merkmal ist entstanden, es gab eine Mutation, Veränderung, etwas ist verloren gegangen, es gab x Ereignisse.</p> <p>Kodieren, wenn mit Entstehung ein Aufspaltungsereignis genannt wird (Ankerbeispiel e).</p> <p>Nicht kodieren, wenn Bezug zu einzelner Info aus Aufgabenstamm (bei Fale verloren gegangen)</p> | <p>a) „und dann müsste es bei (.) Toge verloren gegangen sein“ ↓ „und bei (..) Pote, (..)“ (P29, Alt1_3_3)</p> <p>b) „und Punkte <u>zweimal</u> entstanden sind, <u>einmal</u> bei Fize und <u>einmal</u> bei Mabe und Rone.“ (P06, Zweck1_3_3)</p> <p>c) „Vier Schritte, <i>wenn ich jetzt das richtig überblicke (4)</i>“ (P29, Alt1_3_3)</p> <p>d) „und wir können jetzt, äh, (.) praktisch sowohl annehmen, dass Zipe das nochmal neu erworben hat, ↓ was jetzt ein Schritt wäre, den man dazurechnen müsste (.)“ (P25, Zweck2_3_3)</p> <p>e) „und hier wiederum kann es zum/ (..) äh zur erneuten Aufteilung/ äh Auftrennung gekommen sein, die eine Merkmalsänderung wieder zurück zum Streifen oder nicht zurück, sondern zum Streifen äh begünstigt hatte (.) und das könnte wieder die Aufteilung sein.“ (P07, Zweck1_3_3)</p> |
| | | Ausnahme: Abzählen von Ereignissen als Zahlen ohne Inhalt wird als ein Ereignis kodiert | „Äh eins, zwei, drei (.)“ (P29, Alt1_3_3) |

| | | | |
|---------------------|--|--|---|
| | | Werden Entstehung_Verlust zusammen mit Konvergenz genannt, wird nur Konvergenz kodiert. | Siehe Konvergenz |
| Sparsamkeitsprinzip | Beim Vergleich von mehreren DpV ist das DpV zu bevorzugen, das die Merkmalsausprägung bestimmter Taxa am sparsamsten, also mit den wenigsten (sekundären) Verlust-/Entstehungsereignissen erklärt. = Parsimonie | Wenn das Sparsamkeitsprinzip als Begründung für eine Entscheidung angegeben wird. Sinneinheit: Wenn die Begründung E/V miteinschließt, wird dies extra kodiert. Wenn dadurch eine andere SE unterbrochen wird, wird wie bei Pausen >5 sec der Teil der anderen SE kodiert, der inhaltlich relevant ist (siehe zweites Ankerbeispiel). Nicht, wenn nur das Wort genannt oder erklärt wird, es sei denn, es besteht ein Bezug zu einer anderen Aussage. Kommentar, worauf sich bezogen wird (Hypothese, DpV, Ereignisse o.ä.) | „Das heißt, im rechten Stammbaum wäre es (..) sparsamer, (.) von dieser Hypothese auszugehen (4)“ (P29, Alt1_3_3) <i>ich (..) sage der letzte gemeinsame Vorfahre (.) hatte, wenn man Sparsamkeitsprinzip anwendet (...)</i> [Sparsamkeitspr.] äh (.) Streifen, (4) [Merkmal MRCA] |
| | | Wird Sparsamkeit in einem Satz mit Entstehung_Verlust oder Konvergenz begründet, wird der Satz in die entsprechenden SE eingeteilt und kodiert. Als Kommentar wie Sparsamkeit begründet bzw. eingebettet wird. | „und dreimal konvergentes Entstehen von Streifen auf Punkten (..) [konv. E.] ↓ würde ich als weniger sparsam einschätzen [Spars.] ↓ als zweimal (.) das (.) Auflösen von Streifen zu Punkten (..) [E/V]“ (P14, Zw1_3_3) „Es ist äh (.) parsimonischer, [Spars.] ↓ davon auszugehen, dass (.) Paze und Lete ihn einmal zusammen erworben haben in ihrem gemeinsamen Vorfahren [E/V] ↓ und Zipe (.) [E/V]“ (P29, Zw2_3_3) |

| | | | |
|------------------------|---|--|--|
| Konvergente Entstehung | Gleiche bzw. ähnliche Merkmale sind nicht auf einen gemeinsamen Vorfahren zurückzuführen, sondern sind unabhängig voneinander entstanden. | <p>Indikatoren: konvergente, unabhängige Entstehung., Homoplasie.</p> <p>Auch kodieren, wenn das falsche Fachwort im korrekten Kontext benutzt wird (siehe zweites Ankerbeispiel).</p> | <p>„Die Beinpanzer sind also unabhängig voneinander entstanden“ (fi)</p> <p>„Die sind unabhängig voneinander entstanden, also ist es eine <u>Homologie</u>“ (fi, wie bei P07)</p> |
| | | <p>Werden Entstehung_Verlust zusammen mit Konvergenz genannt, wird nur Konvergente Entstehung kodiert.</p> <p>Sinneinheiten: Wie bei Entstehung/Verlust für jedes Ereignis eine Sinneinheit (siehe Ankerbeispiel c und d).</p> <p>Wenn Konvergenz direkt hintereinander für dieselben Taxa genannt wird, wird nur einmal kodiert (Ankerbeispiel c, erste SE, Ankerbeispiel d).</p> <p>Beispielhafte Beschreibung siehe P18, Zweck2_3_3</p> | <p>a) „dann (..) müssten Streifen (.) bei (.) Lase (.), <u>1</u> Pune <u>1</u> und Kepe (4) dreimal konvergent entstanden sein (.)“ (P14, Zw1_3_3)</p> <p>b) „und über die Entstehung des Merkmals Beinpanzer, äh, kann, äh, die Hypothese aufgestellt werden, dass sie zweimal unabhängig in der Evolution entstanden sein muss.“ (26, Zweck2_3_3)</p> <p>c) „dass es ein analoges Merk/ oder ein/ ein/ ein konvergentes Merkmal ist, also, (.) äh (.) oder ein Merkmal ist, was mehrmals entstanden sein könnte, also einmal in Zipe (.) <u>1</u> und dann nochmal in Paze <u>1</u> und Lete (.)<u>1</u> oder in einem gemeinsamen Vorfahre von Paze und Lete, (..)“ (P28, Zweck2_3_3)</p> <p>d) „Und unabhängig davon (.) äh Zipe (..) den Beinpanzer erworben hat, also konvergent erworben hat (.) <u>1</u></p> |

| | | | |
|---------------------------------|---|---|--|
| | | | unabhängig von Paze und Lete (.)“ (29, 2_3_3) |
| Evolute Abfolge | Reihenfolge der Entstehung verschiedener Merkmale | | |
| (Ind) genetische Prädisposition | Wenn ein Merkmal aufgrund genetischer Veranlagung begründet wird. | | „meine Hypothese lautet dann, dass der letzte gemeinsame Vorfahre anscheinend äh genetische Veranlagung/ äh Veranlagungen besessen haben muss für Punkte und Streifen, (...)“ (P09, Zweck1_3_3) |
| (ind) nicht näher spezifiziert | | Wenn ein weiterer evolutiver Prozess genannt wird, der nicht näher spezifiziert wird. | „Äh (.) weiß grad nicht mehr genau, wie es heißt (lacht). Aber wenns nicht mehr gebraucht wird“ (P08, Alt1_1_3) → vermutlich Rudiment „Dass die auf andere Art entstanden sind als bei Paze und Lete,“ (O16, Zweck2_3_3) |

| | | | |
|--|--|---------------------|---|
| Code: | Zeitverlauf | | |
| Definition | DpV bzw. in diesem Falle Kladogramme, bilden Zeit relativ ab, die Länge der Äste hat keine zeitliche Bedeutung. Die relative Zeitachse verläuft von der Wurzel zu den Taxa. Zeit schlägt Klade: Wird ein Zeitverlauf beschrieben, wird das Ereignis in dieser Kategorie kodiert und nicht z.B. bei Klade. | | |
| Kodierregeln | Wenn eine Aussage über das Alter von Taxa, Kladen, <i>MRCA</i> gemacht wird. Wenn Aussagen über den Zeitverlauf allgemein gemacht werden. | | |
| Subcode | Beschreibung | Kodierregeln | Ankerbeispiele |
| fachlich | | | „Und (..) dass also das Taxon/ (.) oder die Linie, die zu Fute und Sobe geführt hat, etwas (.) früher äh/ also sozusagen eine, ja (.) basale Abzweigung sind (...) ja (.)“ (P29, Alt2_3_3) |
| Alternative Vorstellungen/ Leseweisen | | | |
| (ind) Abzweigungsreihenfolge | AVoLe | | „also das Sparsamkeitsprinzip, das sehen wir halt eher im (.) rechten Stammbaum, würde ich so sehen, weil dort immer/ wir haben immer eine rezente Art, die sich früher absplattet als die (.) andere.“ (P22, Alt1_3_3) |
| Richtung | AVoLe: Zeitverlauf wird anders verortet, z.B. entlang der Taxa | | „Das oberste Taxon ist das jüngste, das untere das älteste“ (fi) |
| Absolute Zeit | AVoLe: Es wird versucht, aus der Länge der Äste Zeiteinheiten zu generieren. | | „Ein Ast ist doppelt so lang wie der andere, also ist doppelt so viel Zeit vergangen.“ (fi) |
| Ältere und jüngere Taxa | AVoLe: Taxa, die sich später von einem <i>MRCA</i> abgespalten haben, werden als jünger angesehen als z.B. die Outgroup | | „(.) Äh, (.) also, wie gesagt, Kuje, äh, ist in beiden Stammbäumen als Letztes entstanden, (.)“ (P21, Alt1_2_3) |
| Astlänge hat Bedeutung | AVoLe: ein längerer Ast bedeutet, dass sich etwas länger entwickelt hat. Vor allem beim Vergleich mehrerer DpV | | |

| | | | |
|--|--|--|--|
| Code: | Weitere Bestandteile | | |
| Definition | In Kladogrammen spielt die Länge, Richtung oder Biegung von Ästen keine Rolle, da nur die relative Zeit angegeben ist. Sie verbinden die Knotenpunkte (<i>MRC</i> A) miteinander (z.B. Baum & Smith, 2013). | | |
| Kodierregeln | | | |
| Subcode | Beschreibung | Kodierregeln | Ankerbeispiele |
| Taxa _Eigenschaften | Information über die gezeigten Taxa | | „alle Taxa sind rezent“ (fi) |
| _oberflächlich | | | „mit jeweils sechs terminalen Taxa (4)“ (P29, Alt1_1_3) |
| Äste und Topologie _inhaltlich | | Wenn direkt oder aus dem Kontext geschlossen werden kann, dass es um inhaltliche Aussagen geht. Nicht, wenn nur das Wort Topologie ohne Kontext genutzt wird. Wenn vorher schon Topologie inhaltlich beschrieben wurde und sich darauf bezogen wird. | „(..) Links und rechts sind zwei verschiedene Topologien, (.) \perp nicht die gleichen Bäume, (.) sondern haben wirklich andere Taxa, äh, andere, äh <i>Clades</i> “ (P23, Alt1_1_3) |
| oberflächlich | | Wenn allgemeine, unspezifische Aussagen über die Topologie getroffen werden, auch wenn der Fachbegriff Topologie (ohne erkennbaren Zusammenhang) genutzt wird. Wenn unspezifische Aussagen zur Position von einzelnen Taxa gemacht werden. | „ansonsten sind die Abbildungen/ Ja, andersrum, dass bei der linken die erste Abstammung ganz unten ist“ (P18, Alt1_1_3) |
| (Ind) allgemein_ zusammenfassend | | Wenn Aussagen über die Topologie sich auf vorher gemachte Aussagen beziehen und diese zusammenfassen. | „Gemeinsamkeiten habe ich jetzt wenige gefunden,“ (P23, Alt1_2_3) „die Position von (.) Pote (..) ja unterscheidet sich/“ (P29, Alt1_2_3) |

| | | | |
|--|--|---|--|
| | | Wenn Aussagen eher allgemein sind, sich aber auf vorher gemachte fachliche Beschreibungen beziehen. | <p>„sonst sind die beiden Bäume sehr unterschiedlich (.).“ (P23, Alt2_2_3)</p> <p>„es gibt viel tiefer in der Phylogenie nicht viele/ nicht viele Gemeinsamkeiten (...)“ (P23, Alt1_2_3)</p> <p>„[aber haben auch zusammen mit anderen (5)] Arten dann mehr, ja, gemeinsame Vorfahren, (.).“ (P13, Alt2_2_3)</p> <p>„die Morphologie stimmt hier mit der Genetik überhaupt nicht überein (...)“ (P23, Alt2_2_3)</p> <p>„Äh, (.). das ist aber so ziemlich das a/ die einzige Gemeinsamkeit, (4) die ich/ (.). die ich sehen kann (10) Ja, sonst sind die beiden Bäume sehr unterschiedlich (.).“ (P23, Alt2_2_3)</p> |
| Alternative Vorstellungen/ Leseweisen | | | |
| Oberflächliche Topologie | AVoLe: Zwei DpV, die die gleichen Kladen in anderer Orientierung (durch Rotation) zeigen, werden als unterschiedlich angesehen | | <p>„und ja, die Anderen sind etwas dann/ Mehr durcheinander, (.).“ (P21, Alt1_1_3)</p> <p>„Der rechte sieht halt systematischer aus (.).“ (P08, Alt1_2_3)</p> <p>„Der linke Stammbaum zeigt ein etwas, äh, vereinfachtes Bild, (.). wie das zu den Arten gekommen ist (.).“ (P15, Alt1_2_3)</p> |
| Hauptast | AVoLe: Vor allem bei leiterförmigen DpV wird der vermeintlich durchgehende Ast als Hauptast gesehen. | | |
| Anzahl DpV | | | „Es sind wieder zwei Stammbäume (.).“ (P29, Alt1_1_3) |

b) Tree-Reading-Fähigkeiten³²

| Code: | Tree-Reading-Fähigkeiten | |
|---------------------|---|---|
| Definition | <p>Jede Sinneinheit, die in 1. DpV-Bestandteile kodiert wurde, wird auch in dieser Kategorie kodiert. Es geht nun nicht mehr um die exakte Identifikation von Kladen etc., sondern wie diese Bestandteile auf einer Metaebene genutzt werden. Tree-Reading beinhaltet die Interpretation und den Vergleich von DpV (Halverson, 2011). Die Tree-Reading-Fähigkeiten sind unterteilt in AVoLe (A), oberflächlich (0), ablesen (1) und ableiten (2) und vorhersagen (2). Die Zahlen 0-2 sind eine vermutete Einteilung in Niveaus. Es wird unterschieden, ob eine Sinneinheit sich auf mehrere_DpV oder ein_DpV (bzw. _unklar) bezieht. Dem liegt zugrunde, dass es anspruchsvoller ist, mehrere DpV miteinander zu vergleichen als nur ein DpV zu interpretieren (z.B. Schramm et al., 2019).</p> | |
| Kodierregeln | <p>Tree-Reading-Fähigkeit_Mehrere_DpV wird kodiert, sobald die vorhergehende oder nachfolgende Sinneinheit aus 1. DpV-Bestandteile das gleiche Taxon, eine ähnliche Klade etc. vergleichend gegenüberstellt bzw. darauf Bezug genommen wird. Jede Sinneinheit aus 1. DpV-Bestandteile wird dabei einzeln als mehrere_DpV kodiert (Bsp. 1).</p> <p>Tree-Reading-Fähigkeit_Mehrere_DpV wird auch kodiert, wenn eine Aussage für das andere DpV bestätigt wird, ohne sie ausführlich zu nennen, (Bsp. 2).</p> <p>Tree-Reading-Fähigkeit_Mehrere_DpV wird nicht kodiert, wenn ein Unterschied unspezifisch formuliert ist, also keine Information konkret abgelesen oder abgeleitet wird (Bsp. 3).</p> | <p>Bsp. 1: „Wobei (..) im linken Stammbaum Fale (unv) (..) eine basale Abzweigung innerhalb dieses Taxons mit den fünf verbleibenden (..) äh Taxa ist/ terminalen Taxa ist,“ ⊥ „während es (..) im rechten Stammbaum die (..) Schwestergruppe zu Kuje bildet (..) das ist also schonmal ein großer Unterschied (..)“ (P29, Alt1_1_3)</p> <p>Bsp. 2: „Ja, Goje und Takwe/ äh (.) Goje und Kawe sind noch näher miteinander verwandt“ (Ablesen_1_mehrere_DpV) ⊥ wie auch im ersten Stammbaum.“ (Ablesen_1_mehrere_DpV) (P29, Alt2_1_3)</p> <p>Bsp. 3: „Dass, wenn der linke immer noch der genetische ist (..), bedeutet dass, Dine, Fute und Sobe sind ein Monophylum (..)“ (Ablesen_1_mehrere_DpV) ⊥ „Der morphologische sagt aber (...) das ist nicht so“ (nicht in dieser Kategorie kodieren) (P08, Alt2_2_3)</p> |

³² Oberflächlich_0 = oberflächliches Ablesen (TR_0), inhaltlich_1 = wissenschaftlich adäquates Ablesen, AVoLe_A = wissenschaftlich inadäquates Ablesen (TR_1b), Ableiten_2 = Ableiten (TR_2)

| Subcode | Beschreibung | Kodierregeln | Ankerbeispiele |
|----------------------------------|--|--------------|--|
| AVoLe nutzen_A | Wird kodiert, wenn in 1. DpV-Bestandteile eine AVoLe kodiert wurde. Unabhängig davon, ob auf ein oder mehrere DpV angewendet. | | |
| Oberflächliche Merkmale nutzen_0 | Nutzung oberflächlicher Bestandteile/ Charakteristika ohne Bezug zu zugrundeliegender Information | | „Ist der, (..) ist nicht der gleiche Stammbaum, nee. Ist nicht der gleiche Stammbaum, sind zwei verschiedene,“ (P21, Alt1_1_3) „und (..) der größte Unterschied ist wahrscheinlich, dass oben/ (..) äh der obere Ast auf der rechten Seite (.) nur zwei Taxa beinhaltet (.)“ \perp „und der untere vier (.)“ (P11, Alt2_2_3) |
| Ablesen_1 | Vorhandene zugrundeliegende Informationen werden aus DpV abgelesen. Dies beinhaltet Verwandtschaftsverhältnisse, Verwandtschaftsgrade und das implizite Schließen auf Merkmale aus dem Aufgabenstamm (nur x und y haben Punkte → alle anderen haben keine Punkte) | | „während es (..) im rechten Stammbaum die (.) Schwestergruppe zu Kuje bildet (.) das ist also schon mal ein großer Unterschied (..)“ (P29, Alt1_1_3) „der letzte gemeinsame Vorfahre müsste also Streifen haben [ableiten] weil Lase [ablesen] und Kepe auch Streifen haben [ablesen]“ (fi) |
| Ableiten_2 | Wenn Informationen aus Aufgabestamm und/oder DpV genutzt werden, um neue Informationen über vergangene Ereignisse oder Zustände abzuleiten. Wenn Informationen genannt werden, die über die Infos in Aufgabenstamm und DpV hinausgehen. Wenn Merkmale von Taxa oder MRCA, die nicht explizit oder implizit im Aufgabenstamm genannt werden. Wenn bekannte Informationen (z.B. Merkmale von Taxa aus Aufgabenstamm) genutzt werden, um Mechanismen | | „Das heißt, im rechten Stammbaum wäre es (..) sparsamer, (.) von dieser Hypothese auszugehen (4)“ (P29, Alt1_3_3) „so hier wäre es also so, dass äh wenn Schneidezähne beim letzten gemeinsamen Vorfahren des (.) Fünfertaxons entstanden sind, (..) \perp dann hätten wir hier einen evolutiven Schritt (..)“ P29, Alt1_3_3) „und das einzige, was halt (..) ja, interessant ist, ist, dass Fize auch Punkte hat, wo man vielleicht annehmen würde, dass es auch Streifen hat.“ (P06, Zweck1_3_3) |

| | | | |
|----------------------|---|---|--|
| | (Entwicklung von Merkmalen, Konvergenz) zu begründen. | | |
| Vorhersagen_2 | Wenn Informationen aus Aufgabestamm und/oder DpV genutzt werden, um neue Informationen über zukünftige Ereignisse oder Zustände abzuleiten. Wenn Informationen genannt werden, die über die Infos in Aufgabestamm und DpV hinausgehen. Z.B. über die Einordnung einer neu entdeckten Spezies in ein bestehendes DpV. | | |
| DpV als Teil des ToL | | Wenn explizit erwähnt wird, dass DpV aus der TA nur einen Ausschnitt des Tree of Life (ToL) repräsentiert | |

c) Bezug (nur bei Teilkompetenz *Alternative Modelle*)

| Code: | Bezug | |
|---------------------|---|--|
| Definition | Diese Kategorie wird für die Auswertung der visuellen Daten kodiert und ist deshalb unabhängig vom Inhalt. Es soll kodiert werden, welches DpV wann erwähnt wird und ob dies als Vergleich oder einzeln stattfindet. Sie gilt nur für TA Alternative Modelle. Jede Sinneinheit aus 1. DpV-Bestandteile wird hier kodiert. Außerdem werden alle weiteren Aussagen kodiert, die sich auf DpV beziehen (siehe Bsp. 2). Es wird kodiert, auf welches der beiden DpV sich die Sinneinheit bezieht. DpV_1 ist immer das linke, DpV_2 immer das rechte DpV. Es wird unterschieden zwischen <i>_Vergleich</i> und <i>_einzeln</i> . | |
| Kodierregeln | <p><i>_Vergleich</i> wird kodiert, sobald die vorhergehende oder nachfolgende Sinneinheit aus 1. DpV-Bestandteile das gleiche Taxon, eine ähnliche Klade etc. vergleichend im anderen DpV gegenüberstellt bzw. darauf Bezug genommen wird. Jede Sinneinheit aus 1. DpV-Bestandteile wird dabei einzeln als <i>_Vergleich</i> kodiert (Bsp. 1 und 2)</p> <p>Im Gegensatz zu 2. Tree-Reading-Fähigkeiten wird hier ein unspezifischer Vergleich kodiert (Bsp. 3)</p> <p><i>_Vergleich</i> kann für ein DpV ggf. ohne Gegenstück kodiert werden (Bsp. 4). Indikator sind hier Formulierungen wie „auch“ oder „genau wie“.</p> | <p>Bsp. 1: „Wobei (..) im linken Stammbaum Fale (unv) (.) eine basale Abzweigung innerhalb dieses Taxons mit den fünf verbleibenden (..) äh Taxa ist/ terminalen Taxa ist,“ (DpV_1_Vergleich) ↓</p> <p>„während es (..) im rechten Stammbaum die (.) Schwestergruppe zu Kuje bildet (..)“ (DpV_2_Vergleich) (P29, Alt1_1_3);</p> <p>Bsp. 2: „Also ist der rechte der morphologische (...) Genau.“ (DpV2_Vergleich) ↓ „Und der linke ist der genetische, (..)“ (DpV_1_Vergleich) (P08_Alt2_1_3)</p> <p>Bsp. 3: „Dass, wenn der linke immer noch der genetische ist (..), bedeutet dass, Dine, Fute und Sobe sind ein Monophylum (..)“ (DpV1_Vergleich) ↓ „Der morphologische sagt aber (...) das ist nicht so“ (DpV2_Vergleich) (P08, Alt2_2_3)</p> <p>Bsp. 4: „Und das andere (räuspert sich) rechte Kladogramm (...) da steht Rine auch einzeln (4)“ (DpV2_Vergleich) (P11_Alt1_2_3)</p> |

| Subcode | Beschreibung | Kodierregeln | Ankerbeispiele |
|------------|--------------|--|--|
| DpV_1 | | | |
| _Vergleich | s.o. | s.o. | s.o.; „wie auch im ersten Stammbaum.“ (P29, Alt2_1_3) |
| _einzeln | | Wenn vorher oder nachher kein Bezug zum erwähnten Aspekt hergestellt wird (bestimmtes Taxon, Klade o.ä.) | „ Gucke ich mir erstmal den linken Stammbaum an (...) so hier wäre es also so, dass äh wenn Schneidezähne beim letzten gemeinsamen Vorfahren des (.) Fünfertaxons entstanden sind, (..) dann hätten wir hier einen evolutiven Schritt (..) [es folgen 7 Sinneinheiten, die sich auf DpV1 beziehen].“ |
| DpV_2 | | | |
| _Vergleich | s.o. | s.o. | s.o.; „äh, im rechten ist dies nicht der Fall, (.) äh, (...).“ (P21, Alt1_2_3) |
| _einzeln | | | |
| DpV_beide | | Indikator: beide, die zwei, der linke und der rechte | „Also in beiden Stammbäumen ist auf jeden Fall Rine die basale Abzweigung, das ist schonmal gleich (.)“ (P29, Alt1_1_3) |
| DpV_unklar | | | |

d)

Modellkompetenz

- siehe Kodierleitfaden von Krell, M., & Krüger, D. (2016).
- Wird nur kodiert, wenn der Modellcharakter von DpV erwähnt bzw. erläutert wird: „Die, (.) äh, (...) die Topologie ist natürlich immer eine Hypothese,“ (P23, Alt2_3_3)

e) Aufgabenbearbeitung (ehemals „Sonstiges“)

| Code: | Aufgabenbearbeitung | | |
|---------------------------------|---|---|--|
| Definition | | | |
| Kodierregeln | | | |
| Subcode | Beschreibung | Kodierregeln | Ankerbeispiele |
| (ind) Wiedergabe_Instr_St | Wiedergabe von Instruktion oder Informationen aus Aufgabenstamm ohne weitere Verknüpfung (für Triangulation) | Auch paraphrasiert oder fragmentiert | „also Schneidezähne wurden gefunden bei Kuje und (.) Mipe (.) und F/ der Vorfahre von Fale hatte wahrscheinlich auch Schneidezähne.“ (P06, Alt1_3_3) „die Hypothese ist, dass die Vorfahren von Fale auch Schneidezähne hatten (..) und welcher der beiden Stammbäume unterstützt die genannte Hypothese unter Anwendung des Sparsamkeitsprinzips/prinzips (.) Also die Vorfahren (...) hatten (...) noch Schneidezähne“ (P11, Alt1_3_3) |
| (ind) Explizite_Beantwortung | Auf die methodische Instruktion bezogene Antwort (siehe Anhang). Auch wenn sie nicht dem Erwartungshorizont entspricht. | Wenn explizit eine Antwort auf die Instruktion gegeben wird (siehe Anhang). Auch wenn sie nicht dem Erwartungshorizont entspricht (Kommentar zu abweichender Bearbeitungsart) | „Und das wäre die Hypothese, dass äh, (.) äh, Kawe und Goje kein, äh, Monophylum ist, sondern ein Paraphylum und äh, Tike noch dazu gehört, (.) anhand von genetischen Daten (.)“ (P23, Alt2_3_3) |
| (ind) allgemeines zu DpV | Aussagen zur Konstruktion, zu fehlenden Daten etc. | Sinneinheit ist ein ganzes Argument bzw. eine Reihe von Argumenten, die sich aufeinander beziehen. | „Ich habe ein bisschen in die andere Richtung gedacht, weil ich, äh/ Das soll zwar nicht die Genetik immer bevorzugen, aber die meisten Stammbäume, die (unv) genetisch aufgestellt wurden, (.) äh, (.) haben so eine gute Datengrundlage, (.) dass man dann eher die (...) Morphologie als interpretiert/ a/ anhand von der Genetik, ni/ nicht die (..) morph/ die Hypothese/ die/ die morphologische Hypothese hinterfragt, mit der genetischen (4)“ (P23, Alt2_3_3) |
| Abstrakte Aufgabenlösung | Alternative Argumentationswege, sowohl korrekt als auch fachlich fraglich oder unlogisch. | Wenn Aussagen zur Beantwortung der Instruktion getroffen | „Wenn wir hier Schneidezähne nehmen, würde das halt mit dem Linken nicht, äh, übereingehen, weil dann zum Beispiel Kuje und Mipe in einem |

| | | | |
|-----------------------------|---|---|--|
| | | <p>werden, die jenseits der Ableitungen sind, die aus Aufgabenstamm und DpV gemacht werden können. Es wird auch kodiert, wenn die Aussage fachlich falsch oder unlogisch ist.</p> <p>Sinneinheit ist ein ganzes Argument bzw. eine Reihe von Argumenten, die sich aufeinander beziehen.</p> | <p>Monophylum sein müssten. Aber vielleicht gibt es auch noch andere, (.) äh, morphologische Eigenschaften, die halt dann (.) dazu führen. ↓</p> <p>Wobei ich jetzt natürlich auch wieder im Hinterkopf die genetische und die morphologische, (.) äh, (.) morphologische un/ und genetischen Stammbäume habe, sodass das halt links wahrscheinlich so ein genetischer Stammbaum ist, (..) wo ja (...) auch andere Umweltbedingungen, Stichwort Epigenetik, eine Rolle spielen könnten, warum halt F/ Schneidezähne jetzt ausgebildet werden und warum keine ausgebildet werden, bei manchen Arten.“ (2x kodiert, P13, Alt1_3_3)</p> |
| Interaktion mit Testleitung | Frage an Testleitung und Antwort der Testleitung. | Wenn Frage und Antwort direkt aufeinander folgen, werden sie als eine Sinneinheit kodiert. | |
| Definition Fachbegriffe | Wenn Fachbegriffe definiert werden unabhängig davon, ob die Definition korrekt ist. | Fachbegriff und korrekt/inkorrekt als Kommentar | „Jetzt das Sparsamkeitsprinzip, jetzt müssen wir also wieder die (.) notwendigen Sch/ evolutiven Schritte zählen (..)“ (P29, Alt1_3_3) |
| (ind) Pause 5-10 sec | | | |
| (ind) Pause 11-20 sec | | Wenn eine Pause durch Husten oder kurze Fragmente (hm). unterbrochen wird, wird die Pausenzeit vorher und hinterher addiert und kodiert. | „ (16) (hustet) (4)“ P11_Alt_1_3_3 → 16+4=20 |
| (ind) Pause > 20 sec | | | |