

### 25.2.1.2 Trennschärfe

Die Schwierigkeit einer Aufgabe bzw. eines Untersuchungsbereichs steht in einem engen Zusammenhang mit dessen Trennschärfe, wenn sie auch nicht allein dafür ausschlaggebend ist, ob die Aufgabe gut zwischen starken und schwachen Leistungen unterscheidet. Aber extrem schwere oder leichte Aufgaben haben in der Regel eine geringere Trennschärfe als Aufgaben mittlerer Schwierigkeit, die gleichzeitig auch eine gute Trennschärfe aufweisen, wenn die mit der Aufgabe zu treffende Unterscheidung die Erwartung erfüllt, dass Probanden mit einem hohen Gesamtpunktwert dabei erfolgreicher sind als solche mit einem niedrigen Gesamtergebnis.

Wegen der Beziehung der Trennschärfe zum Gesamtergebnis wird diese mitunter auch als Homogenität oder innere Konsistenz bezeichnet und auf diejenigen Items, Aufgaben oder Untersuchungsbereiche bezogen, die hoch mit dem Gesamtpunktwert korrelieren und daher geeignet erscheinen, die beobachtete Varianz gut aufzuklären. Das Homogenitätskonzept greift natürlich nur bei einem homogen konstruierten Untersuchungsverfahren. Das Beobilingua-2dit-Verfahren setzt sich hingegen der Komplexität der Sprachkompetenz entsprechend bewusst aus heterogenen Bereichen zusammen, weshalb die berechneten Trennschärfekoeffizienten nicht überinterpretiert werden dürfen. Während sich Schwierigkeitsindizes schon bei Stichproben geringen Umfangs (ab 25 Pbn) als ziemlich konsistent erweisen, ist die Stabilität eines Trennschärfekoeffizienten erst ab etwa 400 Pbn einigermaßen gesichert.<sup>249</sup> Vor allem bei heterogenen Verfahren ist die Trennschärfe sehr labil, so dass die zum Teil sehr niedrigen Trennschärfen angesichts der weitaus geringeren Probandenzahl nicht unbedingt als Beleg einer mangelhaften Differenzierungsfähigkeit des Verfahrens aufgefasst werden sollten.

Bei der Beurteilung der Trennschärfen des verwendeten Verfahrens müssen mehrere Aspekte berücksichtigt werden, was an einigen Beispielen erläutert werden soll. Allgemein gilt bei quantitativen Verfahren eine Aufgabe als trennscharf, wenn kompetente Schüler sie lösen, inkompetente Schüler dagegen nicht. Analog ist bei der qualitativen Bewertung dieses Beobachtungsverfahrens ein Untersuchungsbereich bzw. der dabei geltende Kriterienkatalog als trennscharf zu definieren, wenn sprachkompetente Schüler möglichst viele Kriterien erfüllen, schwächere Schüler hingegen nur wenige. Aufgaben mit einem Trennschärfekoeffizienten um Null gelten als unbrauchbar, weil sie nicht deutlich genug gute und schwache Leistungen unterscheiden.

---

<sup>249</sup> Lienert 1969, a.a.O., S.114

Nun ergeben sich bei der Trennschärfenanalyse neben durchaus befriedigenden Koeffizienten mehrmals auch Werte um Null, etwa bei dem in deutscher Sprache geführten Interview in den Bereichen Hörverständnis, Morphosyntax und Wortschatz. Während die fehlende Trennschärfe beim Hörverständnis und der Morphosyntax eindeutig auf die nur sehr geringe Schwierigkeit zurückzuführen ist, die schon bei der Schwierigkeitsanalyse hinsichtlich der deutschen Sprache auftrat, ist beim Bereich Wortschatz eher ein besonderes Phänomen für die mangelnde Trennschärfe verantwortlich. Hinsichtlich der lexikalischen Fähigkeiten in der deutschen Sprache waren nämlich die Schüler der italienischen Sprachgruppe den deutschen Schülern überlegen, obwohl sie insgesamt in der Partnersprache niedrigere Ergebnisse erzielten als die Schüler der deutschen Sprachgruppe. Dieses Phänomen ist vermutlich durch ein beim zweisprachigen Aufwachsen entwickeltes Sprachbewusstsein zu erklären, das sich anscheinend besonders auf den Wortschatz auswirkt. Da es für die Schüler der italienischen Sprachgruppe leichter ( $P = 83\%$ ) war, den Kriterien im Bereich Wortschatz zu genügen, als den Schülern deutscher Herkunft ( $P = 61\%$ ), widerspricht das Ergebnis im Bereich Wortschatz dem Gesamtergebnis, woraus algebraisch eine mangelhafte Trennschärfe resultiert.

Auch beim Bereich der Lesefertigkeit ergeben sich in beiden Sprachen unabhängig vom Status der Sprache nur Trennschärfekoeffizienten um Null oder sogar solche mit negativen Kennzeichen. Wieder ist die mangelnde Trennschärfe nicht auf eine zu geringe Schwierigkeit zurückzuführen, da die Schwierigkeitsindizes in diesem Bereich von 0,83 bis 0,56 variieren, sondern auf das Verletzen der Homogenitätsregel, die der Trennschärfenberechnung zugrunde liegt. Da das Lesen einen spezifischen Aspekt der Sprachfähigkeit darstellt, stimmen die Bewertungen in diesem Bereich oft nicht mit der Tendenz des Gesamtergebnisses überein.

Die wenig zufriedenstellenden Trennschärfen sind zum Teil überdies auf das äußerst unzulängliche Rechenverfahren zurückzuführen, das zur Ermittlung der Trennschärfe jedoch benutzt werden musste, weil bei qualitativen Analysekrterien kaum andere Verfahren als Vierfelder-Berechnungen zur Verfügung stehen. Dabei werden aber die Beobachtungen definitorisch dichotomisiert, so dass viele Differenzierungen bei dieser Prozedur verloren gehen und mit dieser Methode ermittelte Koeffizienten besonders bei kleinen Stichproben nur eine geringe Präzision haben. Um diesen Mangel zu umgehen, wurde nach Möglichkeit die Berechnung des Neunfelder-Phi-Koeffizienten<sup>250</sup> bevorzugt, bei dem wenigstens 3 Abstufungen berücksichtigt werden können. Ansonsten wurde auf den Vierfelder-Phi-Koeffizienten zurückgegriffen.

---

<sup>250</sup> Nach Lienert / von Eye 1994, a.a.O., S.108.

Auf ein Beispiel zur Berechnung des Neunfelder-Phi-Koeffizienten folgt die Tabelle der Trennschärfekoeffizienten der Untersuchung.

| Schreibfertigkeit D<br>(Bezug 25) |                 |           | Soll (nach Gesamtwert) |                 |        | max. 3 |
|-----------------------------------|-----------------|-----------|------------------------|-----------------|--------|--------|
|                                   |                 |           | >21,3                  | 21,3 ≥ x < 17,6 | ≤ 17,6 |        |
|                                   |                 |           | > 2                    | 2 ≥ x < 1       | ≤ 1    |        |
| Ist                               | >21,3           | > 2       | a = 17                 | 1               | b = 0  | A = 18 |
|                                   | 21,3 ≥ x < 17,6 | 2 ≥ x < 1 | 1                      | 7               | 1      | 9      |
|                                   | ≤ 17,6          | ≤ 1       | c = 0                  | -               | d = 1  | B = 1  |
| Phi-9 = 0,72                      |                 |           | C = 18                 | 8               | D = 2  | N = 28 |

### Trennschärfekoeffizienten Phi bzw. Phi-9

|                                |       |       |                               |       |       |
|--------------------------------|-------|-------|-------------------------------|-------|-------|
| <u>Bild A (it.)</u>            |       |       | <u>Bild B (d)</u>             |       |       |
| Hörverständnis                 | 0,30  | Phi-9 | Hörverständnis                | 0,00  | Phi   |
| Phonematik/Prosodie            | 0,62  | Phi-9 | Phonematik/Prosodie           | 0,37  | Phi   |
| Wortschatz                     | 0,71  | Phi-9 | Wortschatz                    | 0,00  | Phi   |
| Morphosyntax                   | 0,58  | Phi-9 | Morphosyntax                  | 0,00  | Phi   |
| mündl. Sprachgebrauch          | 0,66  | Phi-9 | mündl. Sprachgebrauch         | 0,00  | Phi   |
|                                |       |       | <u>Bild A/B (it.+ d)</u>      |       |       |
|                                |       |       | Wortschatz                    | 0,62  | Phi-9 |
|                                |       |       | Morphosyntax                  | 0,59  | Phi-9 |
|                                |       |       | mündl. Sprachgebrauch         | 0,61  | Phi-9 |
| <u>I - 2 Textv.</u>            |       |       | <u>D - 2 Textv.</u>           |       |       |
| Textverständnis                | 0,32  | Phi   | Textverständnis               | 0,00  | Phi   |
| Wortschatz                     | 0,83  | Phi-9 | Wortschatz                    | 0,69  | Phi   |
| Morphosyntax                   | 0,82  | Phi-9 | Morphosyntax                  | 0,00  | Phi   |
| mündl. Sprachgebrauch          | 0,56  | Phi-9 | mündl. Sprachgebrauch         | 0,69  | Phi   |
| <u>Spiel 1 (it.)</u>           |       |       | <u>Spiel 1 (d)</u>            |       |       |
| Hörverständnis                 | 0,16  | Phi   | Hörverständnis                | 0,00  | Phi   |
| mündl. Sprachgebrauch          | 0,16  | Phi   | mündl. Sprachgebrauch         | 0,00  | Phi   |
| <u>Schreibfertigkeit (it.)</u> |       |       | <u>Schreibfertigkeit (d)</u>  |       |       |
| I - 3 Abschreiben*             | -0,12 | Phi   | D - 3 Abschreiben             | 0,32  | Phi   |
| I - 4 Diktat                   | 0,63  | Phi   | D - 4 Diktat                  | 0,91  | Phi   |
| Schreibfertigkeit G            | 0,30  | Phi   | Schreibfertigkeit G*          | 0,72  | Phi-9 |
| *mit monol. Kontrollgruppe     |       |       | *mit monol. Kontrollgruppe    |       |       |
| <u>concetti</u>                |       |       | <u>Konzepte</u>               |       |       |
| Punkte                         | 0,61  | Phi   | Punkte                        | 0,33  | Phi   |
| kom. Sprachverhalten I         | 0,76  | Phi-9 | concetti / Konzepte (it. + d) | 0,52  | Phi   |
|                                |       |       | kom. Sprachverhalten D        | 0,52  | Phi-9 |
| <u>I - 1 Lesen</u>             |       |       | <u>D - 1 Lesen</u>            |       |       |
| Phonematik/Prosodie            | 0,54  | Phi   | Phonematik/Prosodie           | 0,00  | Phi   |
| Lesefertigkeit                 | -0,09 | Phi   | Lesefertigkeit                | 0,00  | Phi   |
| <u>PS-I - Lesen</u>            |       |       | <u>PS-D - Lesen</u>           |       |       |
| Phonematik/Prosodie            | 0,44  | Phi   | Phonematik/Prosodie           | 0,35  | Phi   |
| Lesefertigkeit                 | 0,09  | Phi   | Lesefertigkeit                | -0,22 | Phi   |
| <u>I-1 + PS-I - Lesen</u>      |       |       | <u>D - 1 + Ps-D Lesen</u>     |       |       |
| Phonematik/Prosodie            | 0,54  | Phi-9 | Phonematik/Prosodie           | 0,20  | Phi   |
| Lesefertigkeit                 | 0,00  | Phi   | Lesefertigkeit                | -0,14 | Phi   |

Demnach scheinen die Untersuchungskriterien in Bezug auf die italienische Sprache im Schnitt trennschärfer zu sein als bei der deutschen Sprache, was sich wiederum rechnerisch daraus ergibt, dass die italienische Sprachkompetenz bei den SESB-Schülern eine größere Streuung aufweist als die deutsche. Im Italienischen waren schwächere von guten Fähigkeiten öfter eindeutig zu „trennen“. In der Regel fallen die Phi-9-Koeffizienten höher aus als die mit einer Vierfeldertafel berechneten Phi-Koeffizienten, was aber mehr auf die Mängel der wenig differenzierenden Methode hinweist als auf eine geringere Trennschärfe dieser Bereiche.

Nach den Maßstäben der Testtheorie und ohne Berücksichtigung der aufgezeigten Mängel der zur Verfügung stehenden Berechnungsmethode weisen bei der deutschen Sprache nur die Bereiche Wortschatz (Anordnung Textverständnis) und die Schreibfertigkeit aufgrund der Komponente Diktat befriedigende Trennschärfen auf. Allgemein reichen Koeffizienten ab  $r_{it} = 0.70$  bei informellen Tests oder zu Gruppenvergleichen aus, während zur individuellen Diagnose Trennschärfekoeffizienten von mindestens  $r_{it} = 0.90$  gefordert werden, besonders wenn das Verfahren zur Selektion eingesetzt werden soll, was beim vorliegenden Beobachtungsverfahren aber nicht der Fall ist. Dagegen erreicht ein Verfahren am ehesten sowohl eine hohe Validität wie Reliabilität, wenn die Trennschärfekoeffizienten etwa von 0,3 bis 0,8 variieren.<sup>251</sup>

Bei der italienischen Sprache erweisen sich die Bereiche Hörverständnis (Bild A und Spiel 1), Textverständnis, Lesefertigkeit, motorische Schreibfertigkeit sowie der mündliche Sprachgebrauch beim Spiel 1 als wenig geeignet, starke und schwache Sprachfähigkeiten zu unterscheiden. Die Bereiche Phonematik/Prosodie, Wortschatz, Morphosyntax, Rechtschreibfertigkeit (Diktat), Begriffsbildung (concetti) und das kommunikative Sprachverhalten erfüllen dagegen selbst bei Anwendung der groben Berechnungsmethode die an eine Mindesttrennschärfe gestellten Anforderungen mit Koeffizienten zwischen  $r_{it} = 0.44$  (Phonematik/Prosodie bei Ps-I-Lesen) und  $r_{it} = 0.83$  (Wortschatz beim Textverständnis).

Diese Werte sind wegen der genannten Vorbehalte gegenüber der äußerst ungenauen Berechnung der Trennschärfekoeffizienten keineswegs als absolut gültige Ergebnisse zu betrachten. Sie wurden hier dennoch dargestellt, weil Angaben zur Trennschärfe bei einer seriösen Aufgabenanalyse als verbindlich gelten und die Koeffizienten neben anderen Kennwerten außerdem zur statistischen Ermittlung der Reliabilität des Untersuchungsverfahrens erforderlich sind. Trotz der laut den Berechnungen nur geringen Trennschärfe einiger Bereiche erweist sich die Reliabilität des Verfahrens im Übrigen für alle Lerngruppen erstaunlicherweise als sehr gut.

---

<sup>251</sup> Lienert 1969, a.a.O., S.295