REGIONALE
Schulberatungsstelle
DES KREISES BORKEN

# Diagnostik

## IN DER (INKLUSIVEN) SCHULE

WEST MÜNSTERLAND
KREIS BORKEN

## Impressum

Herausgeber: Kreis Borken / 2016
Regionale Schulberatungsstelle
Burloer Str. 93
46325 Borken

Autor: Dr. Sascha Borchers

Redaktion: Michael Sylla

Druckerei: Hausdruckerei

1. Auflage: 500

ISBN-10: 3-937432-54-X

ISBN-13: 978-3-937432-54-0

Copyright: © Kreis Borken 2016

# INHALTSVERZEICHNIS


Diagnostik – Irrungen und Wirrungen 4

Diagnostik in der Schule 5

- Zum Begriff der „Diagnostik“ 5
- Tradition von Diagnostik in der Schule 6
- Leistungsfeststellung, Lernverlaufskontrolle, Förderplanung 9
- Das Etikettierungs-Ressourcen-Dilemma & “Wait to fail” 10

Diagnostik bei Inklusion 12

- Zur Rolle von Diagnostik in der Inklusion 12
- Lösung von „wait to fail“: Integration von Diagnostik in das Curriculum 16
- Standardisierte Wiederholungen im Rahmen des Unterrichtscurriculums 17
- Was curriculumbasierte Messungen nicht sein sollten 20
- Beispiele für curriculumbasierte Messungen und angedockte Unterrichtskonzepte in der Grundschule 21
- „Lernen sichtbar machen“ durch Rückmeldung von diagnostischen Ergebnissen („Feedback“) 26

Diagnostik mit psychometrischen Verfahren 29

- Standardisierte psychometrische Diagnostik hilft bei der Beantwortung bestimmter Fragestellungen 29
- Konstruktionsprinzipien psychologischer Test 32
- Grenzen standardisierter Diagnostik 37
- Standardisierte Diagnostik in der Schule 38
- Standardisierte Diagnostik in der Schule ist sinnvoll 40

Fazit & Ausblick 42

Literatur 43

# DIAGNOSTIK – IRRUNGEN UND WIRRUNGEN

Nähert man sich dem Thema „Diagnostik“ im Bereich der Schule, dann wird schnell deutlich, wie viele Missverständnisse, Trugschlüsse und überzogene Erwartungen mit diesem Begriff verknüpft sind. Viele „Arten“ von Diagnostik (pädagogische-, didaktische-, Förder-, inklusive, selektive ...) werden genannt. Diese betonen die unterschiedlichen Aspekte bzw. Funktionen. Zudem gibt es Äußerungen wie „unnötig“, „überholt“, „gefährlich“ und viele andere mehr. Dabei stellt Schlee (2012) fest:

*„Diagnostik kann nur etwas Seiendes feststellen [...] Wer der Diagnostik spezifische Wirkungen und Folgen unterstellt, trägt zum Entstehen von Mystifikationen bei.“*

Gemeint ist, dass Diagnostik als solche wertfrei ist. Die verschiedenen, eingangs erwähnten „Arten“ von Diagnostik betonen immer schon ein Ziel des diagnostischen Prozesses.

Dies ist verständlicherweise verwirrend. Da sich das Schulsystem aktuell im Umbruch befindet, gibt es derzeit Verunsicherungen in Bezug auf Diagnostik in der Schule.

Das aktuell vorherrschende Thema im Bereich Schule lautet „Inklusion“. Und gerade in diesem Kontext bedürfen Ziele und Methoden von „Diagnostik“ einer Klärung.

Durch das Erfordernis der individuellen Bewertung von Schülerinnen und Schüler müssen Lehrerinnen und Lehrer tagtäglich diagnostizieren. Um die Leistung eines Schülers zu werten, muss sie erst einmal festgestellt sein. Somit ist jede Leistungsfeststellung immer ein diagnostischer Vorgang.

Es ist also eine Standortbestimmung nötig, um wieder für Sicherheit zu sorgen. Diese Standortbestimmung soll Aufgabe der vorliegenden Handreichung sein. Ihr Fazit ist: Diagnostik ist in vielfacher Hinsicht eine Hilfe

- als standardisierte Testdiagnostik, zur Bestimmung der Lernausgangslage
- zur Evaluation von Förderprozessen
- zur Begründung von Entscheidungen bei AO-SF-Verfahren.

# DIAGNOSTIK IN DER SCHULE

## Zum Begriff der „Diagnostik“

„Diagnostik“ bedeutet von der griechischen Wortabstammung her in etwa „gründlich kennenlernen“ und dann „beschließen“. Um etwas beschließen zu können, ist es nötig, ein Ziel zu haben (denn nur damit kann ein Beschluss sinnvoll sein). Daher hat auch Diagnostik immer ein **Ziel**: Die Überlegung des Diagnostikers, was er selbst wissen möchte. Oder anders ausgedrückt: Was ist die genaue **Fragestellung**, zu deren Beantwortung die Ergebnisse des diagnostischen Prozesses beitragen sollen. Solche Fragen können im Schulkontext sehr vielfältig sein:

- Ist ein Kind schulfähig?
- Worin sind die Disziplinschwierigkeiten einer Klasse begründet?
- „Hat“ Andreas „LRS“?
- Wie kann ich als Lehrer mein Zeitmanagement verbessern?
- Ist Lena hochbegabt?
- Welche Note soll Tobias im Fach Englisch bekommen?
- Was sind Ursachen meines „Stresses“?
- Welche Stärken hat Marion?
- Gestalte ich guten Unterricht?
- Wie kann ich Tobias besser fördern?
- ...

Je nach zugrundeliegender Fragestellung wird die durchgeführte Diagnostik einen anderen Schwerpunkt haben bzw. mit anderen Methoden durchgeführt werden. Solche Methoden sind:

- Klassenarbeiten / Vergleichsarbeiten/ Lernstandserhebungen
- Befragungen, Interviews
- Gespräche
- (Unterrichts-)Beobachtungen
- Standardisierte Testverfahren
- Familienanamnesen
- Umfelddiagnostik
- ...

Im schulischen Kontext begegnen uns im Zusammenhang mit Diagnostik viele Begrifflichkeiten:

- Psychologische Diagnostik
- Pädagogische Diagnostik
- Sonderpädagogische Diagnostik
- Psychometrische Diagnostik
- Didaktische Diagnostik
- Förderdiagnostik
- Standardisierte Diagnostik
- Inklusive Diagnostik
- ....

Es würde den Rahmen dieser Broschüre sprengen, alle Konnotationen dieser Begriffe zu klären. Diese Begriffe dienen dazu, Ziele des diagnostischen Prozesses zu beschreiben. So hat beispielsweise pädagogische Diagnostik das Ziel, aufgrund der Ergebnisse, Aussagen über mögliche erzieherische Interventionen treffen zu können. Didaktische Diagnostik möchte Hinweise zum weiteren Lernverlauf ableiten usw. Für das Anliegen dieser Handreichung sind genauere Begriffsdefinitionen auf dieser Ebene zunächst zweitrangig, da alle für sich in Anspruch nehmen, in einem regelgeleiteten Prozess Fragestellungen zu beantworten: Der Diagnostiker baut sich aus den ihm zu Verfügung stehenden Methoden seinen „Diagnostikbaukasten" zusammen, um damit eine Fragestellung möglichst objektiv, zuverlässig, gültig, aber auch fair und ökonomisch[1], beantworten zu können. Am Ende des Prozesses steht ein Auswertungsgespräch mit den Auftraggebern der Fragestellung oder ein Bericht bzw. Gutachten, in dem die Fragestellung nachvollziehbar beantwortet wird.

*Diagnostik ist also „...ein theoretisch begründetes System von Regeln und Verfahren zur Gewinnung und Analyse von Kennwerten für inter- und intraindividuelle Merkmalsunterschiede" (Tent & Stelzl, 1993).*

Sie hat im Wesentlichen drei Funktionen[2]. Sie ist

- **beschreibend:** es wird aufgrund einer durch die Diagnostik erfasstes Merkmal eine Beschreibung abgegeben.
- **erklärend:** Sie soll ergründen, warum sich bestimmte Merkmale oder Entwicklungen bei einer Person zeigen.
- **vorhersagend:** Aus der Diagnostik sollen sich Konsequenzen / Handlungsanweisungen für jetzt oder die Zukunft ergeben, d.h aus der Beschreibung folgt ein Urteil z.B. durch einen Vergleich mit anderen Personen oder einer Theorie. Daraus erfolgt die Planung einer angemessenen Intervention.

# *Tradition von Diagnostik in der Schule*

Schule hat unter anderem die Funktionen von Qualifikation, Integration und Selektion. Speziell die Selektivität des deutschen Schulsystems ist in internationalen Studien wiederholt kritisiert worden.

*„Bildlich gesprochen ist die Schule einem großen Rüttelsieb vergleichbar, das zwischen den Generationen angeordnet ist und den Zugang zu beruflichen Positionen, sozialem Prestige und materiellem Erfolg steuert. Deshalb ist in diesem Zusammenhang auch gelegentlich von der Allokationsfunktion die Rede."*[3]

Die Zuweisungsfunktion der Schule greift weitreichend in die Entwicklung junger Menschen ein, weshalb Entscheidungen fachlich und rechtlich Bestand haben müssen.

Begründungen für Laufbahnentscheidungen, Gewährung von Ressourcen oder Nachteilsausgleichen erfolgen über Benotungen und traditionsgemäß über pädagogische, sonderpädagogische, ärztliche oder psychologische Diagnostik. Vorbehalte gegen „Diagnostik" resultieren in erster Linie

aus dieser Funktion: Dabei geht es selten um „Aufstieg“, viel häufiger aber um „Sitzenbleiben“, Abschulungen oder Zuweisung zu Förderschulen.

Das Unbehagen gegenüber Etikettierungen und Platzierungssentscheidungen hat daher schon Anfang der 70er Jahre zu der Debatte um „Förderdiagnostik“ geführt:

*„Mit Diagnostik wurde gesellschaftliche Ausgrenzung (anstelle von Förderung) verbunden. Der bösen Selektionsdiagnostik, die Kinder in Förderschulen und damit aus der Gesellschaft aussondert, wurde die gute Förderdiagnostik gegenübergestellt“ (Schöler, 2008).*

Fachleute betrachten die Notwendigkeit einer Abgrenzung von „Förderdiagnostik“ zur „normalen“ Diagnostik aber durchaus kontrovers:

*„Wenn Förderdiagnostik darin bestehen soll, dass sie im Dienst der pädagogischen Bemühungen steht, dann gilt dies ebenfalls für jede pädagogisch-psychologische Diagnostik“ (Schöler, 2008).*

Gerade im Kontext von Inklusion kann auf fachgerecht durchgeführte Diagnostik nicht verzichtet werden. **Diagnostik** – von der inklusiven Seite (und nicht der ausgrenzenden) betrachtet – **ist der Schlüssel zur Förderung aller Schülerinnen und Schüler**. Durch den Fokus der individuellen Förderung im Schulgesetz des Landes NRW einerseits und durch den gesellschaftlichen Prozess der Inklusion andererseits, bekommt der Förderaspekt von Diagnostik zunehmend wieder höhere Bedeutung.

Bei schulischer Diagnostik muss auf eine weitere grundlegende Problematik hingewiesen werden: Lernleistungen / Schulprobleme von Schülerinnen und Schülern werden gewöhnlich monokausal auf individuelle Veranlagungen / Ursachen zurückgeführt. Allenfalls werden familiäre Faktoren als Verursachung angesehen. Das ist im Sinne von Laufbahnentscheidungen durchaus verständlich, entspricht aber nicht einem modernen Verständnis von Entstehung und Aufrechterhaltung von Verhalten. Eine solche Sichtweise lässt Umweltfaktoren weitgehend außer Acht und Diagnostik und Interventionen fokussieren dabei ausschließlich den einzelnen Schüler. Diese „medizinische“ Sichtweise ist in Schulen außerordentlich stabil und entspricht der Logik des mehrgliederigen Schulsystems, in dem Schüler entsprechend ihren Anlagen platziert werden. Hattie (2013) zitiert eine Studie über 5000 Kinder, die auf Grund schlechter Schulleistung an Psychologen überwiesen wurden:

*„Nicht eine einzige Lehrperson schreibt das Problem einem schlechten Unterrichtsprogramm, einer schlechten Schulpraxis, einer schlechten Lehrperson oder etwas anderem zu, was mit der Schule zusammenhängt. Die Probleme, so die Lehrpersonen, seien auf Ursachen im Elternhaus oder im Lernenden selbst zurückzuführen.“*

Fazit dieses Kapitel ist, schulische Diagnostik nicht als etwas Selektives zu betrachten, sondern mit der "alten Tradition" zu brechen und sie als Hilfsmittel zur Förderung aller Schülerinnen und Schüler anzusehen.

# *Systemische Sicht auf Lernleistungen / Verhalten*

Die systemische Betrachtung geht davon aus, dass Lernleistungen bzw. Schülerverhalten in komplexer Interaktion zwischen individuellen Anlagen und Umwelteinflüssen entstehen und aufrechterhalten werden. Um sich der Beantwortung diagnostischer Fragestellungen anzunähern, ist es daher wichtig zu überlegen, aus welchen Bereichen Informationen gesammelt werden sollen:

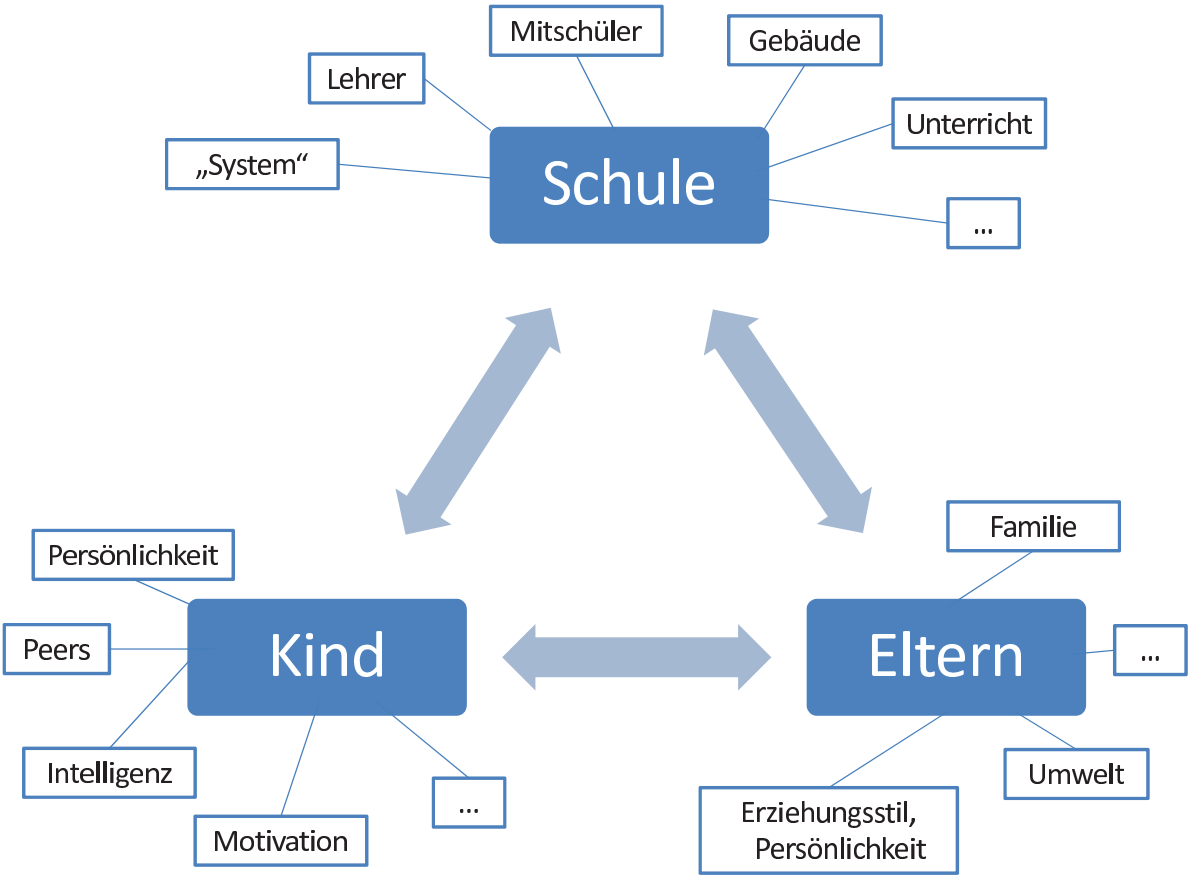


*Abbildung 1: Diagnostische Informationsquellen*

Die Abbildung 1 verdeutlicht, dass sich die Informationsgewinnung nicht auf das Kind konzentrieren darf, sondern Informationen aus dem Lebensumfeld wichtige Beiträge zur Aufklärung einer Fragestellung liefern können. In der Regel sind Ursachen für Probleme nicht singulär, sondern beruhen auf einer (ungünstigen) Verflechtung verschiedener Faktoren. Ein entscheidender Vorteil der systemischen Analyse ist, dass sich daraus auch vielfältigere Interventionsmöglichkeiten ergeben.

# *Leistungsfeststellung, Lernverlaufskontrolle, Förderplanung*

Durch das Erfordernis der individuellen Bewertung der Schülerinnen und Schülern müssen Lehrerinnen und Lehrer tagtäglich diagnostizieren. Um die Leistung zu werten, muss sie erst einmal festgestellt sein. Somit ist jede Leistungsfeststellung immer ein diagnostischer Vorgang. Die vorliegende Broschüre vermittelt grundlegendes Wissen zum Messen und zur „Testtheorie". Hier werden zunächst die wichtigsten diagnostischen Anwendungen des schulischen Alltages kurz vorgestellt:

- **Leistungsfeststellung:** Der „älteste" und wichtigste Aspekt schulischer Diagnostik ist die Leistungsfeststellung. Schülerinnen und Schüler werden aufgrund einer Leistung (z.B. in einer Klassenarbeit, einer mündlichen Mitarbeit etc.) bewertet. Dies geschieht in der Regel auf der bekannten Notenskala von „sehr gut" bis „ungenügend". Diese Leistungsfeststellung dient dazu, am Ende des Schuljahres eine Selektionsentscheidung zu treffen: Wer kommt weiter, wer „bleibt sitzen", wer kommt auf welche Schule, wer bekommt welchen Studienplatz...? Die Kriterien für eine Leistungsfeststellung beruhen auf den von den Schulministerien herausgegebenen Curricula. Allerdings fließen hier oft subjektive Beurteilungskriterien mit ein und die Beurteiler neigen dazu, eine Rangfolge innerhalb der Lerngruppe zu bilden.

- **Lernverlaufskontrolle:** Durch verschiedene Methoden (z.B. kleine "Tests" kann kontrolliert werden, welchen Wissensstand ein Lernender hat. Wiederholt man diese Messung, findet eine Prozessbetrachtung statt. Einige Unterrichtsmaterialien bieten Hilfen, den Lernverlauf für die Lehrperson zu kontrollieren und ggf. sogar für die Schülerinnen und Schüler abzubilden. Dabei werden kurze, unterrichtsbegleitende Wissenstests durchgeführt, die Rückmeldung über das erfolgreiche Bearbeiten einer Lerneinheit geben. Auf diese Weise wissen Lehrende und Lernende immer, wo sie gerade stehen.

- **Förderplanung** ist die Planung von Fördermaßnahmen und den Einsatz von geeigneten Methoden und Materialien z.B. zur Schließung einer Lernlücke. Im Kontext von Teilleistungsschwierigkeiten (z.B. in Mathematik oder Rechtschreibung) sollte vor der Förderung durch entsprechende Diagnostik abgeklärt werden, welcher Stoff schon beherrscht wird und wo die Förderung sinnvoll ansetzen kann.

Gerade der letzte Aspekt „Förderplanung" ist durch den Fokus der individuellen Förderung im Schulgesetz des Landes NRW im schulischen Alltag verankert. Daher gilt eine gezielte Förderplanung nicht nur im Bereich von Teilleistungsschwierigkeiten, sondern sollte Teil „normalen" Unterrichts sein. Da individuelle Förderung nur dann effektiv sein kann, wenn die Lernausgangslage der Schülerinnen und Schüler bekannt ist, ist vor der Förderung diese Diagnostik notwendig. Je heterogener Lerngruppen sind, desto zentraler werden also diagnostische Prozesse im Unterricht, um Schülerinnen und Schüler einzeln im Blick zu behalten. Konkrete Konzepte, die eine Einbindung von Diagnostik und Förderung in das Curriculum vorsehen, werden auf Seite 21 vorgestellt.

Solange Diagnostik jedoch das Ziel hat, eine Leistung zu bewerten oder ein Merkmal zu bestimmen, die eine „Konsequenz" zur Folge haben, kommt noch ein weiterer Gesichtspunkt hinzu: Das Etikettierungs-Ressourcen-Dilemma.

# *Das Etikettierungs-Ressourcen-Dilemma & "Wait to fail"*

In vielen Bereichen unseres täglichen Lebens treffen wir auf ein ähnliches Problem: Eine Unterstützung gibt es nur, wenn es ein „klassifiziertes" Problem gibt. Die Krankenkasse rechnet den Arztbesuch nur bei einer Diagnose ab, eine Unterstützung im Rahmen des Bildungs- und Teilhabepaketes gibt es nur bei der Bescheinigung der Bedürftigkeit. Die finanzielle Unterstützung im Rahmen der Jugendhilfe wird bei Lernschwierigkeiten nur bei Vorlage eines psychiatrischen Gutachtens gewährt und so weiter. Solche Regelungen treten immer dann auf, wenn eine Ressource (also z.B. die finanzielle Unterstützung) knapp ist oder der Ressourcengeber (z.B. die öffentliche Hand) den Eindruck hat, dass er ausgenutzt werden könnte. Daher wird die Ressource an eine Bedingung gebunden. Dagegen ist grundsätzlich wenig einzuwenden – problematisch wird es aber, wenn mit der Feststellung der Bedürftigkeit eine Ettikettierung einhergeht, wie z.B. die Diagnose „lernbehindert", „psychische Störung" oder „erziehungsschwierig". Das „Dilemma" besteht also darin, dass man Kindern und Jugendlichen zunächst eine „Störung", „Krankheit" oder „Behinderung" attestieren muss, um an Mittel (Geld, „Nachteilsausgleiche") zu gelangen. Dies fördert eher Separierung und weniger Inklusion.

Solche Regelungen sind Ursachen für die nach wie vor vorgenommenen Klassifikation und Kategorisierung von beeinträchtigten Kindern. Sie sind oft sowohl von Schul- als auch von Elternseite gewollt, da beide Seiten einen Nutzen davon haben. Die langfristige Folge in Deutschland war jedoch auch eine deutliche Separation. Das Dilemma ist also ein Relikt aus dem Zeitalter der Separation, hat es aber gut durch das Zeitalter der „Integration" (Separation innerhalb der Gruppe) geschafft und ist nun ein Hindernis auf dem Weg zur Inklusion.

Hartke (2012) bringt dies auf den Punkt:

*„Jegliche Form von Feststellungsdiagnostik geht von der Annahme aus, es gäbe zwei eindeutig abgrenzbare Kategorien von Schülern: Schüler mit und ohne sonderpädagogischem Förderbedarf und diese seien über Ist-Standsmessungen von Parametern wie Schulleistungen, Intelligenz, Sprachentwicklung trennscharf zu unterscheiden. Diese Annahmen sind nicht zutreffend und wenig hilfreich für Inklusion."*

Eine Erkenntnis aus dem Dilemma müsste daher sein, Ressourcen losgelöst von einer Diagnose zu vergeben bzw. auf einer solchen Ebene zur Verfügung zu stellen, dass eine Förderung für alle in Frage kommt und nicht nur für Schülerinnen und Schüler mit „Etikett".

Die Folge bei Betreuungspersonen (Eltern, Lehrer, Erzieher, Jugendhilfe...) kann sein, dass man so lange mit der Problemlösung wartet, bis eine Etikettierung sicher stattfinden kann und es Zugang zur Ressource gibt: **„Wait to fail"** ( vgl. z. B. Huber & Grosche, 2012).

Das bedeutet, dass eine Intervention oder die Zuweisung von Ressourcen (z.B. Förderunterricht oder eine außerschulische Unterstützung) erst dann stattfindet, wenn die Lerndefizite der Schülerinnen und Schüler chronifiziert sind. Dieser Warte-Effekt wird oft noch verstärkt, wenn zunächst außerschulische Diagnostik (zumeist medizinische) durchlaufen wird, um abzuklären, was jetzt „mit dem Kind ist." Der Fokus der außerschulischen Diagnostik richtet sich oft nur auf das Kind, ist in der Regel defizitorientiert und liefert wenig Anhaltspunkte für schulische Interventionen. Wichtige Umwelteinflüsse wie Familie, Freunde oder das Schulsystem werden selten berücksichtigt. Als Ergebnis gibt es einen Bericht, der in

weiten Teilen – meistens – das Lehrerurteil bestätigt. Das ist einerseits vielleicht hilfreich, andererseits wurde wieder viel Zeit (und Ressource) verschenkt, die man gut hätte in Fördermaßnahmen investieren können. Hinzu kommt, dass oft ein „Expertenurteil“ abgewartet wird. Es fühlen sich unterschiedliche Experten zuständig bzw. bekommen diese Zuständigkeit zugeschrieben (vgl. z.B. „Ärztehopping“). Nicht selten entsteht so eine „Hilflosigkeit im Expertendschungel“, die eine Intervention verzögert.

Bei der sich anschließenden Förderung ist jedoch der Lernrückstand des Schülers möglicherweise so groß, dass er im verbreiteten Jahrgangsunterricht mit geringen Individualisierungsanteilen kaum noch eine realistische Chance hat, diesen Lernrückstand aufzuholen. Lyon et al., (2001, zitiert nach Hartmann & Müller, 2009) beschreiben genau diesen Effekt:

*„Die Forschung der letzten Jahrzehnte relativiert die Effektivität von remedialen Maßnahmen (Therapien, Klassenwiederholung, Sonderschulplatzierung); obgleich Kinder mit Lernproblemen durch kompetente Interventionen durchaus individuelle Fortschritte verzeichnen können, ist bekannt, dass sie in ihren Schulleistungen oft nicht zu unauffälligen Lernenden aufzuholen vermögen.“*

Da es ökonomisch, bildungspolitisch und persönlich unglücklich ist, zu lange auf eine Intervention zu warten, stellt sich die Frage, wie ein schnelles Vorgehen aussehen kann. Damit sind vor allem Lösungen gemeint, die aus dem Schulsystem selbst heraus generierbar sein müssen. Der Prozess der Inklusion „zwingt“ das System Schule geradezu, hier über Lösungswege nachzudenken. Eine schulpolitische Lösung ist die Ressourcenbereitstellung im LES-Budget (sonderpädagogische Förderung im Bereich der Lern- und Entwicklungsstörungen), die direkt den Schulen zur Verfügung gestellt wird. Dadurch wird sonderpädagogische Diagnostik sichtbar neu definiert - nicht um zu selektieren oder Ressourcen zuzuweisen, sondern um Schüler individuell unterrichten und fördern zu können. Und dies geht nur, wenn man weiß, was ihre Stärken bzw. Schwächen sind und wie mit diesen umzugehen ist. Im Kapitel "Beispiele für curriculumsbasierte Messungen und angedockte Unterrichtskonzepte in der Grundschule" werden drei Beispiele für Diagnostik und Förderung im Rahmen des "normalen Unterrichts" vorgestellt.

# DIAGNOSTIK BEI INKLUSION

Infolge der Anerkennung der UNO-Behindertenrechtskonvention 2009 wurde in Deutschland der Prozess zur Inklusion in Schulen angestoßen. Daher ist es bei der Betrachtung von diagnostischen Prozessen in Schule unabdingbar, auch die Sichtweise „inklusiver Diagnostik“ einzunehmen. Von besonderer Bedeutung muss hier die Grundhaltung sein – weg vom Streben nach Homogenität und Separation hin zur Anerkennung von Heterogenität und Inklusion. Dies hat Auswirkung auf die Förderung aller Schülerinnen und Schüler, nicht nur auf die mit ausgewiesener „Behinderung“. Dabei ist eine wichtige Unterscheidung zu treffen (vgl. Prengel, 2015):

- **A o L** (assessment **of** learning): summative Erfassung eines “Status Quo”
- **A f L** (assessment **for** learning): formative Erfassung, um einen Lernverlauf zu begleiten und zu fördern. Dies ist der Kerngedanke didaktischer Diagnostik.

## *Zur Rolle von Diagnostik in der Inklusion*

Möchte man individuell fördern, muss man wissen, welches die Lernausgangslagen des Schülers sind, was seine Lerngeschichte ist und wie er den nächsten Entwicklungsschritt gehen kann. Dazu muss eine individuelle – und der Inklusion dienliche – Diagnostik durchgeführt werden. Neben der üblichen pädagogisch-psychologischen Diagnostik (Leistungstests, Fragebögen usw.) kommen diagnostische Methoden zum Einsatz, die eng mit der (Förder-) Didaktik verknüpft sind:

*„Inklusive Diagnostik ist eine in den pädagogischen Alltag eingelassene, stets mit den Lernprozessen einhergehende didaktische Prozessdiagnostik. Grundlage inklusiver Diagnostik bilden didaktische Stufenmodelle, die auch dem Kerncurriculum und den Lernmaterialien zugrunde liegen. Die didaktisch qualifizierten Lehrerinnen und Lehrer kennen die aufeinander aufbauenden Kompetenzstufen in den Lernbereichen, so dass sie im schulischen Alltag problemlos erkennen, auf welcher Stufe sich jedes Kind gerade befindet, was die Zone der jeweils nächsten Entwicklung ist und welches pädagogische Angebot zu diesem Zeitpunkt individuell passend ist“ (Prengel, 2013).*

Im Idealfall stellen die genutzten Unterrichts-Materialien auch Methoden der Diagnostik zur Verfügung. Dabei dient Diagnostik dem Ziel, die individuellen Lernangebote im inklusiven Unterricht vorzubereiten. Inklusive Diagnostik durchläuft einen Kreislauf:

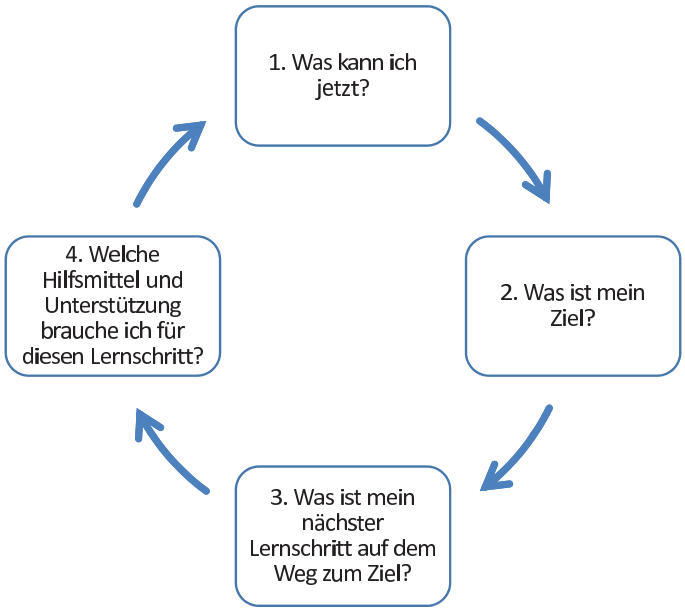


*Abbildung 2: Visualisieren des inklusiven diagnostischen Kreislaufs[4].*

Dieser Kreislauf verdeutlicht, dass Diagnostik nicht dann einsetzt, wenn der „Normalfall" gestört wird, sondern sie bildet einen permanenten Prozess (vgl. auch Eberwein, 1997).

Folgende Gelingensbedingungen für inklusive Diagnostik können formuliert werden:

- **Geringer Aufwand:** Inklusive Diagnostik darf für den alltäglichen Einsatz nicht aufwändig sein und muss sich in das Curriculum integrieren.
- Die inklusive Diagnostik muss im Alltag handhabbar sein.
- Grundlage sind **didaktische Stufenmodelle:** Die Entwicklung der wichtigen Lernbereiche (Rechtschreibung, Lesen, Rechnen, Geometrie) erfolgt bei den Schülerinnen und Schüler (meistens) gestuft. Zur Vermittlung des Lernstoffes gibt es didaktische Stufenmodelle. Lernmaterialien folgen diesen Stufen. Stufenmodelle gibt es in ähnlich einsetzbarer Funktion auch für andere Entwicklungsbereiche, z.B. im Bereich des emotionalen und sozialen Lernens.
- **Kenntnisse dieser Modelle:** Je klarer diese Lernstufen den Lehrern im Alltag präsent sind, umso problemloser erkennen sie im Unterricht anhand der mündlichen, schriftlichen und kreativen Aktivitäten und Produkte der Kinder, auf welcher Stufe sich jedes Kind gerade befindet.
- Kenntnisse der **Bezugsnorm.** Zur Bewertung der Ergebnisse einer Diagnostik ist das Heranziehen einer kriterial vergleichenden Bezugsnorm[5] maßgeblich. Grund: Diese ermöglicht, das erreichte Kompetenzniveau im Vergleich zu den oben genannten fachlichen Stufenmodellen zu setzen. So kann herausgefunden werden, welches die „Zone der nächsten Entwicklung" ist und was getan werden muss, damit der Schüler diese erreicht[6].

Diese Aufzählung zeigt, dass es an vielen Stellen auf dem Weg zu einer inklusiven Diagnostik gar nicht mehr so weit ist – eben weil sie zentraler Bestandteil des Lernprozesses und der verwendeten Materialien ist. Allerdings setzt die Beschreitung dieses Weges eine strukturierte und lerntheoretisch fundierte Vorgehensweise voraus. Die Planung sollte damit frühzeitig und am besten in Anlehnung an bereits etablierte und funktionierende Konzepte erfolgen. Dabei muss aber „das Rad" nicht neu erfunden werden: „Vor allem kleine Verlage bieten

gegenständliche Lernmaterialien an…“ schreibt Prengel (2013, S. 49), mit der Idee, ein inklusives Klassenzimmer zu gestalten. Es gebe schon eine Menge guter und funktionierender Materialien. Mit dieser Sichtweise wird auch deutlich, dass der oft diskutierte Begriff der „Förderdiagnostik“ in der Inklusion gar nicht nötig ist. Pädagogische Diagnostik hat immer das Ziel, ein Kind zu fördern. In der inklusiven Diagnostik ist sie Teil des Lernalltages. Dennoch muss dieser Begriff gewürdigt werden, denn er war der erste Schritt **weg** von einer etikettierenden Diagnostik.

Da Diagnostik zunächst wertneutral ist, gibt der Vorsatz „didaktische“ das Ziel des Prozesses vor: Informationen über lehr- und lernförderliche Bedingungen im Mikrosystem Lerner-Lehrer zu erlangen. In Anlehnung an Prengel (2015) folgt daraus, dass **didaktische Diagnostik** eine (zentrale) Tätigkeit der Lehrerarbeit ist. Sie hat ihren Ort im alltäglichen Unterricht der heterogenen Lerngruppe.

Die Elemente der **didaktischen Diagnostik** sind:

- **Ziel:** angemessene Gestaltung pädagogisch-didaktischen Handelns und kindlichen Lernens
- **Zeitraum:** während der gesamten Unterrichtsphase
- **Erkenntnis- und Handlungssubjekte:** Lehrende und Lerner
- **Erkenntnisgegenstand:** Produkte der Schüler in den jeweiligen Schulfächern, Lernbereichen und fächerübergreifenden Vorhaben
- **Erkenntnisprozess:** spiralförmig (siehe Abbildung 3)
- **Instrumente** sind (u.a.)
  - Kompetenzraster
  - Lernentwicklungsmodelle
  - Benennen relevanter geistiger Operationen
  - Lerntagebücher
  - Unterschiedliche Dokumente zu Lernaktivitäten

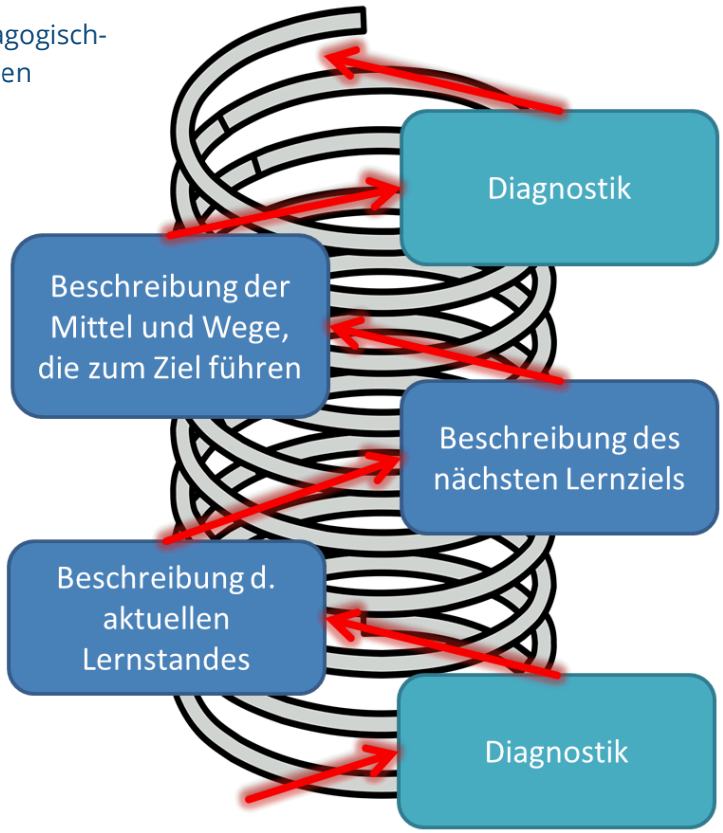


*Abbildung 3: Spiralförmiger Erkenntnisprozess in der „didaktischen Diagnostik“*

# Exkurs: Inklusion und Noten

Viele Leserinnen und Leser werden jetzt sagen: Inklusive Diagnostik – gut und schön. Aber wir müssen doch Noten geben, und daher eine Leistung messen und bewerten. Wie verträgt sich das mit Inklusion?

Der Widerspruch lässt sich nicht ohne Weiteres aufheben. Das kann nur gehen, wenn das Schulsystem auf Notengebung und Selektion verzichtet. Dies ist aber aufgrund politischer und gesellschaftlicher Überzeugungen aktuell (noch) nicht möglich. Allerdings bieten Schulversuche wie „Primus“[7] eine gute Diskussionsgrundlage hierfür. Vielleicht sind wir also doch nicht allzu weit davon entfernt!

Dennoch ist eine Leistungsbewertung im Kontext Inklusion denkbar. Auch hier bietet die Arbeiten von Prengel (2013, 2015) interessante Aspekte. Inklusive Leistungsbewertung sollte aus vier Perspektiven stattfinden:

1. **Perspektive der grundlegenden Wertschätzung:** Jedes Kind ist willkommen. Es gibt keine ausgrenzenden Sanktionen.

2. **Perspektive der Schulleistung auf Basis der kriterialen Bezugsnorm:** Werden bestimmte Lernstufen erreicht?

3. **Perspektive der Schulleistung auf Basis der individuelle Bezugsnorm:** Wie sieht der individuelle Lernfortschritt der Schülerinnen und Schüler aus?

4. **Perspektive des Leistungsvergleichs mit anderen:** Dies ist die bestehende, selektierende Perspektive und gilt vor allem bei Übergangsentscheidungen.

Natürlich können im bestehenden System solche Perspektiven nur unzureichend in Noten einfließen. Sie sollten es aber.

Im Gutachten von Prof. Werning liest sich das so:

*„Konsequent wäre es, hier vollständig auf das AO-SF-Verfahren zu verzichten. Dies setzt jedoch die Umsetzung einer inklusionspädagogischen Beurteilungskultur (z.B. der Verzicht auf Noten bis Klasse 8) und die Einführung einer intensiven Rückmeldungskultur (Lernentwicklungsberichte; Portfolios) für die Schülerinnen und Schüler sowie für die Erziehungsberechtigten über die Lern- und Leistungsentwicklung voraus.“ (Werning, 2011, S. 7)*

Eine individuelle Rückmeldung auf den 4 nebenstehenden Ebenen ist schon jetzt jederzeit möglich, zumal innerhalb einer Schulstufe. Kritisch wird dies beim Übertritt in weiterführende Systeme. Es gilt, dort wo es sinnvoll ist, die bestehenden Spielräume auszuschöpfen: Ein teilweiser Verzicht auf Noten ist (in Grenzen und ohne AO-SF-Verfahren) schon jetzt möglich (vgl. Verzicht auf Noten in Klassen 1 bis 3, im Kontext individueller Förderung oder bei Nachteilsausgleichen etc.)

# *Lösung von „wait to fail“: Integration von Diagnostik in das Curriculum*

Im Verlauf dieser Darstellung wurde schon darauf verwiesen, dass ein Umdenken im Menschenbild (weg von dem Begriff der "Behinderung", hin zur Heterogenität) und damit verbunden in der Haltung zu den Themen „Förderung“ und „Unterrichten“ nötig ist. Eine in den Unterricht nahtlos integrierte (Mikro-)Diagnostik kann hier einen ertragreichen Beitrag leisten. Man nennt diese Vorgehensweise „Curriculums-basierte-Messung“, kurz CbM. Das Ziel der CbM muss sein, Probleme frühzeitig und schnell zu erkennen, um so die Beteiligten in die Lage zu versetzen, schnell (und angemessen) zu handeln und damit Förderbedarfe zu vermeiden. Damit verlieren sowohl das Thema „Behinderung“ als auch die mögliche Ursache für diese Behinderungen an Bedeutung. In den Mittelpunkt rückt der individuelle Lernforschritt der einzelnen Schülerinnen und Schüler. In diesem Kapitel wird eine strukturierte Vorgehensweise vorgestellt, die aus dem amerikanischen Bildungssystem stammt und helfen soll, Probleme frühzeitig zu erkennen und ein Handwerkszeug zur Lösung dieser Probleme bereitzustellen.

Wie kann  Diagnostik im Rahmen des alltäglichen Unterrichts funktionieren?

Die Lehrperson führt ca. 1x wöchentlich[8] in seinem Fach Diagnostik durch, die nicht dazu dient, die Schülerleistung zu bewerten, sondern zu schauen, ob der Schüler den Lerninhalt der zurückliegenden Einheit verstanden hat. Im Idealfall ist diese Messung Teil des verwendeten Unterrichtsmaterials, und gehört zum Unterrichtsalltag. Damit wird es Teil der Feedbackkultur innerhalb der Klasse. Der Vorteil ist, dass Diagnostik Teil des Unterrichts ist und so als Normalfall von den Schülerinnen und Schülern wahrgenommen wird – eben als nichts Besonderes.

Das Ziel dieser Messungen ist es festzustellen, ob der Lernstoff beherrscht wird. Ist dies der Fall, dann verläuft der Unterricht wie gewohnt weiter. Fallen bei Schülerinnen und Schülern wiederholt Lücken auf, so kann er an einer strukturierten Wiederholung des Stoffes teilnehmen. Auch diese Wiederholungen sind im Unterrichtsalltag eingeplant, um zeitliche Räume dafür zu haben. Dies kann z.B. in Freiarbeitsphasen der anderen Schülerinnen und Schüler geschehen.

# *Standardisierte Wiederholungen im Rahmen des Unterrichtscurriculums*

*Nur vom Wiegen wird die Sau nicht fett!*

(Bauernweisheit)

Nach der Durchführung einer Diagnostik ist derjenige, der sie ausgewertet hat „schlauer“: Er weiß z.B., wo der diagnostizierte Schüler Kenntnisse und Lernlücken hat, wo es noch offene Fragen gibt und wie er im Vergleich zum Lernziel bzw. zu anderen Mitgliedern der Lerngruppe vom Wissensstand her einzuordnen ist.

Nicht schlauer ist zunächst derjenige, der diagnostiziert wurde. Denn aus dem Vorgang der Diagnostik ist kein Lerneffekt zu erwarten. Diese auf den ersten Blick triviale Erkenntnis wird leider bei vielen Formen der schulischen Diagnostik vernachlässigt. Im Kontext der didaktischen Diagnostik ist es jedoch die zentrale Frage: Welche Erkenntnisse und welche Maßnahmen leite ich aus meinen Messungen ab? Das bedeutet, dass  Diagnostik im schulischen Kontext immer eine didaktische (!) Intervention zur Folge haben soll („assessment **for** learning“). Eine Intervention soll den Schülerinnen und Schülern helfen, nicht verstandenes Wissen zu verstehen, eine Lernlücke zu schließen oder sie an der Stelle „abzuholen“, an der sie stehen.

Diagnostik hat dann keinen selektiven, wertenden oder gar etikettierenden Zweck, sondern gehört zur Alltagsselbstverständlichkeit von Schülerinnen und Schülern (und Lehrerinnen und Lehrern). Dann ist eine maximale Wirkung erreicht.

Eine vergleichsweise einfache Möglichkeit der ersten Intervention ist die Wiederholung. Beim ersten Feststellen von Schwierigkeiten wird der Lernstoff der zurückliegenden Unterrichtseinheit wiederholt. Diese Wiederholung kann in ohnehin vorgesehenen Differenzierungsphasen stattfinden, aber auch in das bestehende Förderkonzept der Schule integriert werden. Eine häufige, am besten tägliche Implementation dieser Vorgehensweise ist von Vorteil.

Diese Vorgehensweise wirkt auf mehreren Ebenen:

- Ebene der **Lernstoffüberprüfung** und des **Classroom-Managements**: Lehrerinnen und Lehrer erhalten Kenntnis darüber, was ihre Schülerinnen und Schüler können und was nicht (Identifikation von Lernlücken). Werden bei mehreren Schülerinnen und Schülern ähnliche Lücken festgestellt, weisen diese möglicherweise auch auf Unzulänglichkeiten in der Didaktik hin und können dann behoben werden.
- Ebene der **Wiederholung**: Anhand von bestehendem und bekanntem Lernmaterial wird erneut gelernt. Die Einarbeitungszeit ist kurz, und da das Material schon einmal gelernt wurde, sollte die dafür aufgewendete Zeit auch kurz sein. So ist in der Regel mit relativ wenig Aufwand ein hoher Lernertrag möglich.
- Ebene des **Feedbacks**: Schülerinnen und Schüler bekommen ein deutliches Feedback darüber, welcher Stoff schon beherrscht wird, und welcher nicht. Dieses Feedback hilft ihm, sein eigenes Lernen einzuschätzen und zu reflektieren, dass der Lernprozess im nicht gekonnten Stoff ungenügend war.
- Ebene der einfachen **Organisation**: Das Lernmaterial ist bereits vorhanden und bekannt. So haben die Lehrerinnen und Lehrer keine und die Schülerinnen und Schüler nur eine geringe Einarbeitungszeit. Zudem wird

ein hohes Maß an selbständigem Lernen ermöglicht. Die Förderung kann sofort und im Klassenverband beginnen. Die Schüler haben nicht das Gefühl, dass sie etwas Besonderes sind oder ein Problem „haben". Diese Vorgehensweise soll eine Alltäglichkeit sein.

Nach dem erneuten Durcharbeiten einer Lektion ist es nötig, den Lernerfolg kurzfristig zu überprüfen:

- Gibt es (erneut/immer noch) Lernlücken?
- Welcher Art sind diese Lücken?
- Den Schülerinnen und Schülern soll dann zurückgemeldet werden, was sie inzwischen können und was sie schon geschafft haben. Dies ist ein guter Motivator und bringt alle Beteiligten weg von einer Defizitorientierung.

Für diese Phasen muss auch die Betreuung der betroffenen Schüler durch die Lehrperson enger stattfinden. Dafür muss Zeit veranschlagt werden. Diese Zeit kann am besten durch Freiarbeitsphasen der anderen erfolgen bzw. durch eine zeitweilige Doppelbesetzung der Klasse. Ein Interessantes Konzept hier ist „I choose" (vgl. Lohri & Steiner, 2015), bei dem ritualisiert im Tagesablauf Zeiträume von ca. 30min geschaffen werden, in denen Schülerinnen und Schüler, die die Lernziele erreicht haben, einer selbstgewählten (schulischen!) Aktivität nachgehen dürfen, während andere mit Verständnisschwierigkeiten im Stoff in einer Kleingruppe einer intensiven Wiederholung nachgehen.

Wurde der **Rückstand** zur Klasse **aufgeholt**, dann fällt die Teilnahme an den Wiederholungen weg und der Schüler nimmt in den bisher dafür vorgesehenen Zeiten wieder an den standardmäßig durchgeführten Differenzierungsangeboten teil.

Wenn **noch nicht alle Lernlücken geschlossen** sind, dann kann entweder die „Wiederholung wiederholt" werden (mit einer neuen Zusammenstellung von Materialien, die den bisherigen Lernfortschritt berücksichtigen). Erkennt der Lehrer jedoch, dass der Lernfortschritt geringer ausfällt als erwartet, kann es sein, dass das Defizit des Schülers grundlegender ist als angenommen. In diesem Fall sollte der Schüler in den Differenzierungszeiträumen an einem **„Standard-Treatment-Training"** teilnehmen. Das ist ein standardisiertes und evaluiertes Trainingsprogramm zum Üben im jeweiligen Förderbereich[9]. Vor der Anwendung eines solchen Trainings ist zu überlegen, welche Fähigkeit dem Schüler fehlt und aus welchem Funktionsbereich das mögliche Defizit stammt. Dieser Bereich wird von der für den jeweiligen Fachunterricht zuständigen Lehrkraft betreut. Auch hier verbleibt der Schüler in der Lerngruppe und arbeitet im Rahmen der alltäglichen Individualisierung mit dem „standard Treatment". Es ist keine umfangreiche Diagnostik notwendig, die Lehrkraft entscheidet den Einsatz aus ihrer Erfahrung im Umgang mit dem Lernstoff.

Erst wenn der Test des Lernstandes nach der Lerneinheit zeigt, dass der Lernzuwachs geringer ausgefallen ist als erwartet, ist sonderpädagogische Unterstützung nötig. Abbildung 4 verdeutlicht diese Vorgehensweise.

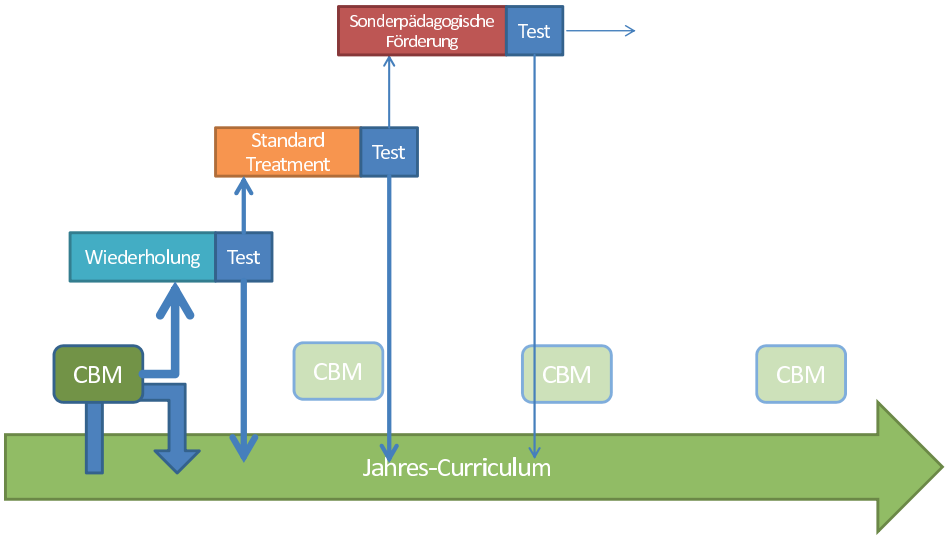


*Abbildung 4: Organisatorischer Ablauf für CbM und hieraus resultierenden Interventionen*

Die **sonderpädagogischer Unterstützung** wird erst dann eingeschaltet, wenn zuvor durchgeführte Interventionen ohne Erfolg bleiben. Sie ist somit die nächste Maßnahme, die innerschulisch durchgeführt werden kann. Die Bereitstellung der sonderpädagogischen Ressource erfolgt im LES-Budget (sonderpädagogische Förderung im Bereich der Lern- und Entwicklungsstörungen), die direkt den Schulen zur Verfügung gestellt wird.

Durch die vorherigen Maßnahmen liegen schon viele Informationen über das Lernverhalten und die Lernfähigkeit des jeweiligen Schülers vor, so dass der Sonderpädagoge diese Informationen für seine Diagnostik nutzen kann. Darüber hinaus können (und müssen manchmal!) eingesetzte standardisierte Testverfahren Teil dieser Diagnostik sein.

Nach sonderpädagogischer Diagnostik und Intervention schließt sich in regelmäßigen Abständen eine Überprüfung der in der Förderung aufgestellten Ziele an, immer mit der Option der Aufhebung der sonderpädagogischen Förderung.

Im Idealfall findet diese Überprüfung der Förderziele so „dicht“ es geht in und um die Lerngruppe des Schülers statt, um eine hohe Durchlässigkeit sicherzustellen.

In der sonderpädagogischen Literatur werden Teile dieser Vorgehensweise auch als „**RTI-Modell**“ (Response-to-Intervention) beschrieben. Das ist ein aus dem nordamerikanischen Kontext kommender Ansatz mit dem Ziel, Lern- und Verhaltensschwierigkeiten möglichst früh zu identifizieren und darauf mit standardisierten Interventionen zu reagieren. Die Vorgehensweise durchläuft – analog des hier Vorgestellten - drei Stufen (oder Ebenen, im englischen „Tires“). Auf der basalen Ebene finden regelmäßige Screenings mit allen Schülerinnen und Schülern statt. Zusätzlich werden von den Lehrerinnen und Lehrern vielfältige, für das Lernen und Verhalten relevante Informationen gesammelt. In der nächsten Stufe findet zunächst eine Analyse des individuellen Lernfortschritts der Schüler statt. Es wird dann eine darauf aufbauende intensivere Lernbegleitung parallel zum Unterricht durchgeführt. In dieser engen Begleitung finden sehr regelmäßig (ein- bis zweimal pro Woche) Lernerfolgskontrollen statt. Tritt auf dieser Ebene keine Leistungsverbesserung ein, wird der Schüler in Stufe drei individuell diagnostiziert, um ein differenziertes Bild vom Lernpotential und den jeweiligen Wirkmechanismen zu erhalten.

RTI bedeutet damit eine Kopplung von Unterricht, Diagnostik und individueller Förderung mit weitgehend standardisierten

und evaluierten Methoden (vgl. Abbildung 5). Auf der Ebene der Klasse gedacht kann das RTI-Modell auch einen Einfluss auf den Unterricht des jeweiligen Lehrers haben. Denn: Sind zu viele Kinder „Nonresponder" (also Lerner mit einer unzureichenden Leistungsentwicklung), dann muss sich der jeweilige Lehrer fragen, ob nicht sein Unterricht insgesamt einer Veränderung bedarf (Gold, 2011, S. 220). Leider wird diese Sichtweise zu selten eingenommen.

Eine sehr ausführliche Beschäftigung mit dem Konzept des RTI und praktischen Anwendungen findet man u.a. bei Liebers & Seifert (2012).

**Diagnostik in der Inklusion: Eher die Regel als die Ausnahme**

Inklusion bedeutet die Abkehr von der Vorstellung, man könne Schülerinnen und Schüler einer Klasse im Gleichschritt unterrichten. In inklusiven Klassen können Schülerinnen und Schüler nicht zielgleich unterrichtet werden. Um trotzdem möglichst individualisierten Unterricht anbieten zu können, müssen Lehrerinnen und Lehrer darüber informiert sein, wo ihre Schülerinnen und Schüler im Lernstoff stehen und wie ihr Lernfortschritt aussieht. Daher ist die Methode der curriculumbasierten Messungen in inklusiven Schulen eher Standard als Ausnahme und sollte zunehmend ein Modell für alle Schülerinnen, Schüler und Schulformen werden.

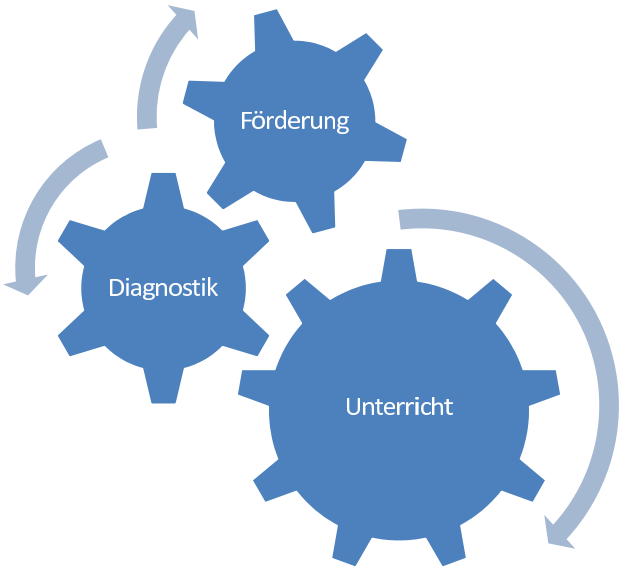


*Abbildung 5: RTI als Kopplung von Unterricht, Diagnostik und individueller Förderung*

# *Was curriculumbasierte Messungen nicht sein sollten*

Die hier vorgestellte Verfahrensweise dient primär dazu, festzustellen

- ob eine Leistungsveränderung eines Schülers (oder einer Klasse) zwischen den Messungen stattgefunden hat und
- ob diese Veränderung auf Lernlücken hinweist

Im letzteren Fall macht das Verfahren darauf aufmerksam, dass pädagogisch umgesteuert werden muss (Klauer 2006) – es hat

einen formativen Charakter. Das bedeutet, dass aufgrund der Messergebnisse und des festgestellten Lernverlaufs Interventionen abgeleitet werden sollten. Solche Interventionen sollten nicht nur für (lern)schwache Schülerinnen und Schüler gedacht werden, sondern für die gesamte Lerngruppe, um der Vorgabe der individuellen Förderung möglichst nahe zu kommen. Ohne eine Kenntnis über den Lernverlauf ist dies nicht möglich.

Aus diesem Grund sind curriculumbasierte Messungen und damit der gesamte RTI-Ansatz **nicht**

- ein Mittel zur Einteilung nach Niveaugruppen
- Klassifikation von Schülerinnen und Schülern im Sinne der AO-SF
- die Voraussetzung für spezielle Förderung.

CbMs dienen **nicht** der Platzierung und **nicht** der Selektion. Sie dienen auch **nicht** dazu festzustellen, ob ein bestimmtes Lernziel erreicht wurde (Klauer 2006).

Der formative Charakter des Ansatzes bedeutet zudem, dass eine mögliche Ursachensuche „systemweit" geschehen muss. Oft wird jedoch das Problem beim Subjekt – in diesem Fall der Schülerinnen oder dem Schüler – gesucht[10]. Die Folge davon ist, dass beim mehrfachen „Nicht-Ansprechen" auf eine Intervention der Verbleib in der Lerngruppe in Frage gestellt wird. Im Kontext von Inklusion stellt sich diese Frage jedoch nicht. „Diagnostische Mittel sind hier Bestandteil des Lehrens und Lernens von Lehrkräften und Kindern" schreibt Prengel (2013).

# *Beispiele für curriculumbasierte Messungen und angedockte Unterrichtskonzepte in der Grundschule*

Die Methode der curriculumbasierten Messungen ist in Deutschland bisher vorwiegend mit Materialien aus der Primarstufe erforscht. Grund ist, dass bisher vergleichsweise wenig Unterrichtsmaterialien hierzu erstellt worden sind. Aus dem angloamerikanischen Raum – wo das Konzept schon länger Einzug in den schulischen Alltag gefunden hat – gibt es jedoch eine Menge Ansätze, die zeigen, wie RTI auf weiterführenden Schulen funktionieren kann. Aber:

*„RTI looks different at the secondary level than it does at the elementary school level." (Quinn, 2012).*

## Lesen

Lesen ist die wichtigste Kulturtechnik. Darum sollte es Aufgabe der Schule sein, den Schülerinnen und Schülern eine gute Lesekompetenz zu vermitteln. Dies ist auch der Grund, warum es eine Reihe von Methoden und Materialien im Kontext von CbM gibt, die Lesefähigkeit zu fördern[11]. Der zentrale Aspekt ist eine diagnostische Komponente mit dem Ziel, den Schülerinnen und Schülern zurückzumelden, welche Lernfortschritt sie gemacht haben und den Lehrerinnen und Lehrern zu zeigen, welche Lernaufgaben die Schüler als nächstes vor sich haben.

Ein aktueller Ansatz wird gerade an der Universität Münster untersucht und soll hier kurz vorgestellt werden: **Der Lese-Sportler** (Schulte & Souvignier, 2013)

Das Programm „Lese-Sportler“ kombiniert die wissenschaftlich gut fundierte Methode des wiederholenden Lesens (zur Steigerung der Leseflüssigkeit) mit Elementen des „Reziproken Lehrens"[12]. Im 3-wöchigen Abstand füllen Schüler einen ca. 10-minütigen Test am PC aus und erhalten danach ein Feedback über ihre Entwicklung hinsichtlich ihrer Leseflüssigkeit und ihres Leseverständnisses. Über ein onlinebasiertes Lehrermenü (www.quop.de) erhalten Lehrerinnen und Lehrer einen detaillierten Einblick. Sie können so über das Schuljahr hinweg die Leseentwicklung jedes Kindes sowie die Entwicklung der ganzen Klasse beobachten und daraus Hinweise für individuelle Förderung und Unterrichtsplanung ableiten (vgl. Abbildung 6).

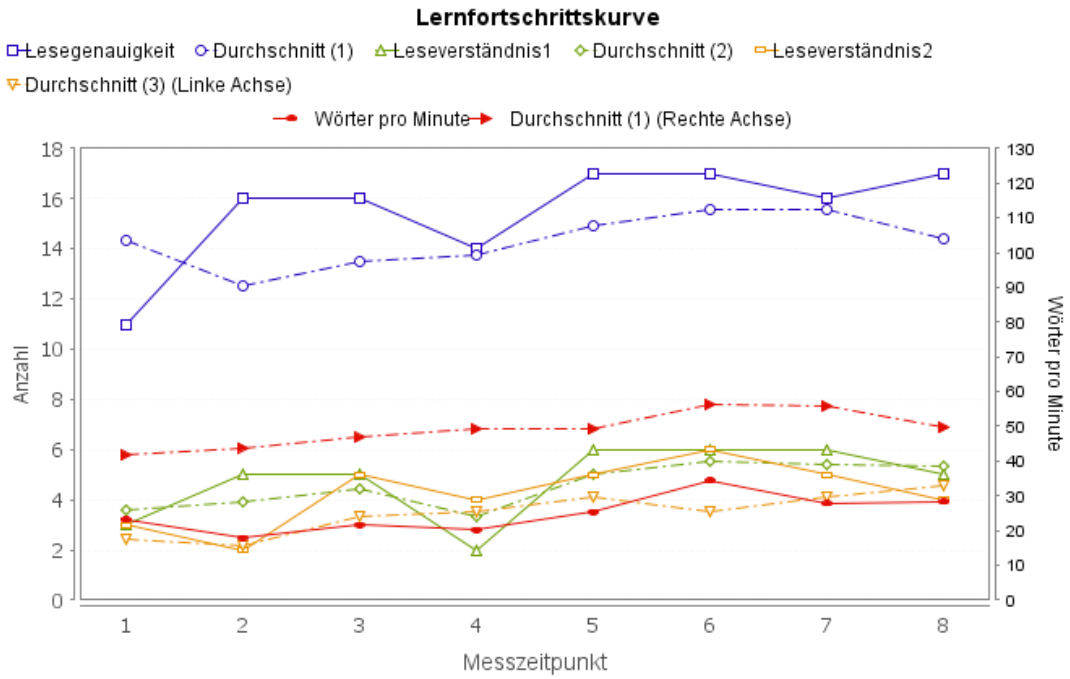


*Abbildung 6: Entwicklungsverlauf eines Schülers im Vergleich zum Klassendurchschnitt (aus Schulte & Souvignier, 2013. Abbildung mit freundlicher Genehmigung der Autoren.)*

## Rechtschreiben

Um beim Rechtschreiben eine sinnvolle curriculumbasierte Messung durchzuführen, müssen dem Unterricht zuvor entweder ein Kompetenzraster mit zugehörigen Unterrichtseinheiten oder Entwicklungsstufen der Rechtschreibung zugrunde gelegt werden. Bei beiden Vorgehensweisen sind die Zwischenziele klar zu definieren, um diese messbar zu machen und sie dann im Sinne eines lernförderlichen Feedbacks an Schüler und Lehrer zurückzumelden. Zudem ist eine Kombination aus beiden Vorgehensweisen möglich (das Kompetenzraster orientiert sich an einer Entwicklungsreihenfolge).

Das Konzept der **Rechtschreibwerkstatt** (z.B. Sommer-Stumpenhorst & Hötzel, 2007) verwendet beide Vorgehensweisen. Über das Schuljahr verteilt sollen regelmäßig Messungen des aktuellen Kenntnisstandes der Schüler erhoben werden. Diese werden

- quantitativ (also Häufigkeiten von Fehlschreibungen) und
- qualitativ (auf Kriterien bezogen)

ausgewertet und anschließend in ein Kompetenzraster („Lernentwicklungsbericht“, vgl. Abbildung 7) eingetragen. Der Lehrer kann so schnell die jeweilige Entwicklung, den nächsten Lernbereich erkennen und dies auch mit Schülerinnen, Schülern und Eltern besprechen.

Gerade die Schülerkommunikation ist ein wichtiger Aspekt des Konzeptes. Die Schülerinnen und Schüler sollen lernen, dass sie für ihr Lernen Verantwortung übernehmen. Es komme darauf an, das „richtige“ zu tun (und danach etwas können) und nicht nur „irgendetwas“ tun (und beschäftigt zu sein)[13].

Mehr Information hierzu gibt es z.B. unter www.rechtschreibwerkstatt.de

| | | 1/1 | 1/2 | 2/1 | 2/2 | 3/1 | 3/2 | 4/1 | 4/2 |
|---|---|---|---|---|---|---|---|---|---|
| **1.** | **Lautebene** | | | | | | | | |
| UW | Buchstaben bewegungsrichtig schreiben | | | | | | | | |
| LB | Laute heraushören / Buchstaben zuordnen | | | | | | | | |
| LB | Buchstaben unterscheiden | | | | | | | | |
| LB | Besonderheiten kennen und anwenden | | | | | | | | |
| LB | au, ei, eu, sch, ch, ng, nk, qu, x | | | | | | | | |
| LB | Endungen –el, -en, -er | | | | | | | | |
| LB | Wortanfang: sp, st | | | | | | | | |
| **2.** | **Wortebene** | | | | | | | | |
| UW | Wortgrenzen einhalten | | | | | | | | |
| UW | Wörter lesbar, formklar, flüssig schreiben | | | | | | | | |
| LB | Wörter lauttreu schreiben | | | | | | | | |
| LD | Laut-Buchstaben-Folgen einhalten | | | | | | | | |
| LD | Silben trennendes *h* | | | | | | | | |
| LD | vokalisches *r* | | | | | | | | |
| LV | Kennzeichnung langer und kurzer Vokale | | | | | | | | |
| LV | langer Vokal, ie, ä, ß | | | | | | | | |
| LV | Doppelkonsonant | | | | | | | | |
| LV | ck, tz | | | | | | | | |
| WA | Wortarten kennen und bestimmen können | | | | | | | | |
| WA | Adjektive, Verben, sonst. Wörter | | | | | | | | |
| WA | Namen, Konkreta kennen und groß | | | | | | | | |
| WA | sonst. Nomen (Abstrakta) groß schreiben | | | | | | | | |
| WU | Wortstammprinzip kennen und anwenden | | | | | | | | |
| WU | Auslautverhärtung | | | | | | | | |
| WU | Umlautschreibung | | | | | | | | |
| WU | Erhalt des Wortstammes in Ableitungen | | | | | | | | |
| WZ | Wortbildungen erkennen, Rechtschreibmuster | | | | | | | | |
| WZ | Vorsilben, Verben | | | | | | | | |
| WZ | Adjektivendungen | | | | | | | | |
| WZ | Nomenendungen | | | | | | | | |
| WZ | zusammengesetzte Wörter | | | | | | | | |
| AF | Sensibilität für Ausnahmeschreibungen | | | | | | | | |
| AF | schreibwichtige Wörter mit *pf, chs, v, ai* | | | | | | | | |
| AF | wichtige W. mit *i*, aa/ee/oo, Dehnungs *h* | | | | | | | | |
| AF | häufige und schreibwichtige Fremdwörter | | | | | | | | |
| **3.** | **Satzebene** | | | | | | | | |
| SZ | Satzanfang groß, Punkt am Satzende | | | | | | | | |
| UW | Texte lesbar und klar gegliedert schreiben | | | | | | | | |
| SZ | besondere Satzschlusszeichen (?, !) | | | | | | | | |
| SZ | einfache wörtliche Rede kennzeichnen | | | | | | | | |

*Abbildung 7: Lernentwicklungsbericht der Rechtschreibwerkstatt. Quelle: http://www.rechtschreib-werkstatt.de/rsl/bew/lentwb/html/bericht.html (Abbildung mit freundlicher Genehmigung des Autors.)*

## Rechnen

Für Messungen im Lernbereich „Rechnen“ gilt analog dem Rechtschreiben die Grundlage eines Kompetenzstufenrasters und einer Entwicklungsstufenfolge.

Das Lehrwerk **„Matinko“** (Jansen, 2010) basiert ebenso wie die „Rechtschreibwerkstatt“ auf einer Kombination von Entwicklungsstufen des mathematischen Verständnisses und einem daraus abgeleiteten Kompetenzraster (vgl. Abbildung 8). Der

*Abbildung 8: Drei Säulen der mathematischen Verständnisgrundlagen nach Jansen. www.matinko.de (Abbildung mit freundlicher Genehmigung des Autors.)*

Lernfortschritt der Kinder wird einerseits mit standardisierten Verfahren für jedes Kind in halbjährlichen Abständen erfasst. Der Lernzuwachs in einzelnen Unterrichtseinheiten wird zudem in einem Vergleich der Leistungen in einem jeweiligen (also zur Unterrichtseinheit gehörenden) Prä- und im Posttest deutlich. Ein Vergleich mit einer Referenzgruppe ermöglicht Aussagen über die Wirksamkeit des Unterrichts allgemein. Jansen (2010) sieht gerade in der Rückgabe der Posttests eine gute Gelegenheit für Motivations- und Lernersteuerung. Der Unterricht wird dadurch maximal individualisiert. Mehr Information hierzu gibt es z.B. unter www.matinko.de.

## (Curriculumbasierte) Diagnostik und Förderprogramme der Schulbuchverlage

Auch die Schulbuchverlage haben erkannt, dass es eine sinnvolle (und gewinnbringende) Idee ist, für Schulen diagnostische Verfahren und Förderprogramme aufzulegen und mit den Lehrwerken zu verzahnen. Die Idee ist in der Regel: Es gibt eine PC-basierte oder gestützte Testung, eine Vollautomatische Auswertung und eine Rückmeldung über die Leistung von Schülerinnen und Schülern. Zu dieser Rückmeldung wird eine Förderempfehlung gegeben. Diese sind entweder spezielle Fördermaterialien aus dem Verlagsprogramm oder Materialien, die zu einem bestimmten Lehrwerk passen.

*Tabelle 1: Beispiele für Diagnostische Verfahren und Förderprogramme von Schulbuchverlagen (Liste nicht vollständig)*

| Name | Verlag und URL | Beschreibung |
|---|---|---|
| Grundschuldiagnose | Schrödel-Verlag<br>http://www.grundschuldiagnose.de/ | Klasse 2 bis 4.<br>Das Programm unterstützt Lehrerinnen und Lehrer darin, den Lernstand der Schüler in den Fächern Deutsch und Mathematik zu testen. Fördermaterialien werden auf Basis der Testergebnisse automatisch und individuell zusammengestellt. |
| Fördern@cornelsen | Cornelsen-Verlag<br>http://www.cornelsen.de/foerdern | Grundschule und Sek. 1.<br>Ermittlung des Leistungsstands und darauf abgestimmte Möglichkeiten zur spezifischen Förderung von Schülerinnen und Schülern.<br>Durchführung von Kompetenztests mit differenzierten Auswertungen. Es werden Förderpläne mit Fördermaßnahmen und Fördermaterialien zusammengestellt. |
| Testen und Fördern | Klett-Verlag<br>http://www.klett.de/testen-und-foerdern | Grundschule und Sek. 1<br>Lehrwerksbezogenes Testprogramm mit Tipps zur Förderung und Materialien aus dem Lehrwerkangebot. |

Gerade diese letzte Variante passt zu der oben vorgestellten Vorgehensweise bei curriculumbasierten Messungen. Die Schüler und Lehrer bekommen sofortige Rückmeldung über das Erreichen eines Lernziels und Hinweise darauf, mit welchen Materialien die identifizierten Lernlücken geschlossen werden können. Beispiele für solche Verfahren sind in Tabelle 1 abgebildet.

Die Einführung solcher Unterstützungshilfen kann – je nach verwendeten Lehrwerken an der Schule – sinnvoll sein. Gefahr hierbei besteht jedoch auf zwei Ebenen:

- allzu sorgloser Umgang mit den Programmen. Alles sieht „so einfach aus", so dass Schülerinnen und Schüler „schön alleine" damit arbeiten können. Aber auch hier sollten die jeweiligen Fachlehrer sich einbringen und die Auswertungen für eine Lernberatung und Feedback zu Grunde legen. Dann können diese Programme sinnvoll und effektiv sein!
- Durch die Zentrierung auf den Lernstoff und Schülerinnen und Schülern alleine kann schnell eine verkürzte Sicht auf individuelle Variablen entstehen. Bei der Entwicklung von Schwierigkeiten ist jedoch immer ein „Bedingungsgefüge" zu betrachten (vgl. Abbildung 1: Diagnostische Informationsquellen).

# *„Lernen sichtbar machen" durch Rückmeldung von diagnostischen Ergebnissen („Feedback")*

In seinem Buch „Visible Learning" geht Hattie (2013) im Kapitel „Feedback" (S. 206ff) auf die Effekte von Ergebnissen unterschiedlicher Arten von Diagnostik ein. Sein zentraler Ansatz hier ist die Rückmeldung von Erkenntnissen zwischen Lernenden und Lehrenden. „Feedback ist eine Information, die von einem Akteur über Aspekte (...) des eigenen Verstehens gegeben wird". (Hattie 2013, S. 206). Also ist „Feedback eine ‚Folge' der Leistung" (ebenda, S. 207). Die effektivste Form von Feedback ist diejenige, die dem Lernenden einen Hinweis in Bezug auf die jeweiligen Lernziele gibt – und nicht z.B. Belohnungen in Form von Noten etc. Dieses Feedback muss aber bei den Lernenden auch ankommen (was man nur im persönlichen Gespräch mit ihnen herausfindet). Im Idealfall verschmelzen Feedback und Unterricht miteinander. Dies stellt sich Hattie wie in Abbildung 9 vor.

Nach dieser „Verschmelzung" gibt es mehrere Stadien, die Feedback durchlaufen muss, um effektiv zu wirken. Zunächst muss es eine bei Lehrenden oder Lernenden wahrgenommene Diskrepanz zwischen Zielerreichung und Leistung geben. Diese Diskrepanz wird aufgedeckt durch die in den Unterricht integrierte Diagnostik. Lehrer und Lernender haben nun die Möglichkeit, diese Diskrepanz zu reduzieren, auf Seite des Lehrenden geht dies über das Anbieten geeigneter Ziele oder über effektives Feedback. Auf Seiten des Lerners über eine erhöhte Anstrengung, effektive Strategien oder das eigene Modifizieren der Ziele. Das Feedback selbst erfolgt auf drei Ebenen. Diese Ebenen sind

- **Feed Up:** Was ist das/dein Ziel? Wohin geht es?
- **Feed Back:** Wie kommst du voran? Wo stehst du gerade?

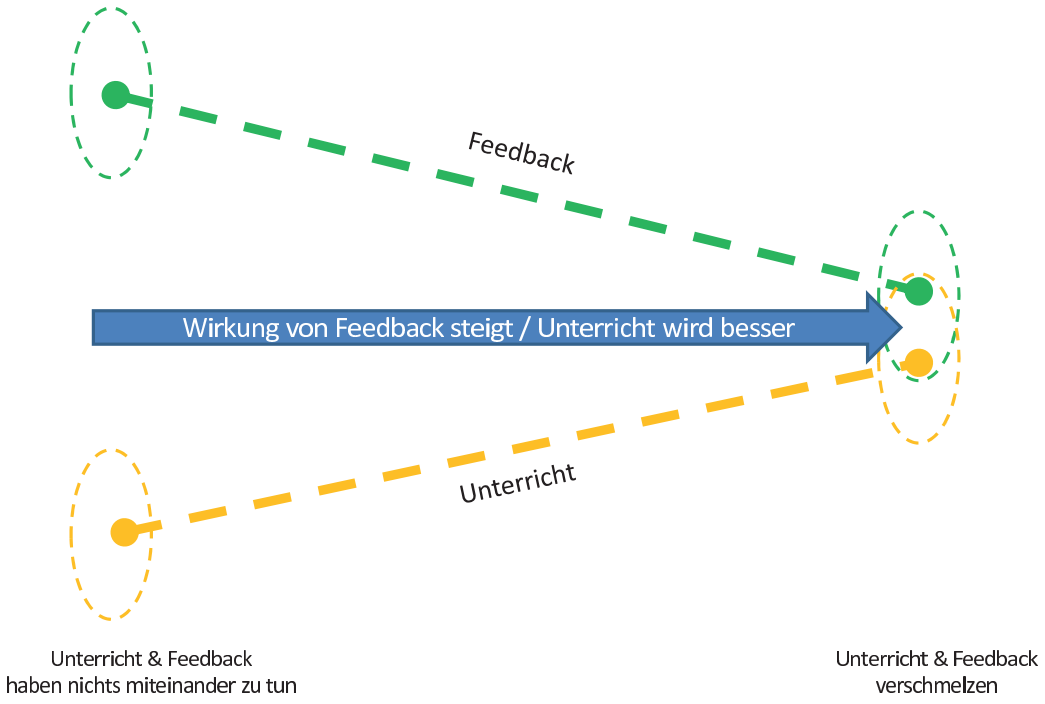


*Abbildung 9: Feedbackmodell von Hattie (2013) zur Verschmelzung von Unterricht und Feedback (eigene Darstellung)*

- **Feed Forward:** Blick auf das nächste Ziel, also „wohin geht es nach der Erreichung des jetzigen Ziels?"

  Um möglichst konkret zu werden, sollen auf jeder Ebene vier gleiche Stufen betrachtet werden:

- Bezug zu **Aufgabe** oder Produkt: Ist etwas falsch oder richtig? Wo kann sich der Lernende weitere Informationen beschaffen? Beispiel: *„Du musst dir zur Lösung der Aufgabe mehr Hintergrundinformationen zum 30-jährigen Krieg beschaffen."*
- Bezug zum **Prozess**, der zur Lösung der Aufgabe beschritten wird. Beispiel: *„Du musst die Charakterisierung der Hauptperson so durchführen, dass deine Leser die innere Zerrissenheit der Person nachvollziehen können. Überlege, was für die eine Seite steht und was für die andere!"*
- Bezug zur **Selbstregulation** des Lernenden. Beispiel: *„Du hast mir im letzten Gespräch gezeigt, dass du die Laute den Buchstaben korrekt zuordnen kannst. Prüfe die Wortanfänge, ob diese alle richtig sind."*
- Bezug auf das **Selbst**: Dieses Feedback wendet sich an die Person ohne Bezug zur Aufgabenleistung. Beispiel: *„Das hast du gut gemacht!".* Dieses **Lob** ist jedoch selten effektiv und nützt der Lernverbesserung wenig. Diese Art von Feedback sollte selten gegeben werden.

  Die folgende Abbildung 10 fasst das oben gesagte noch einmal zusammen:

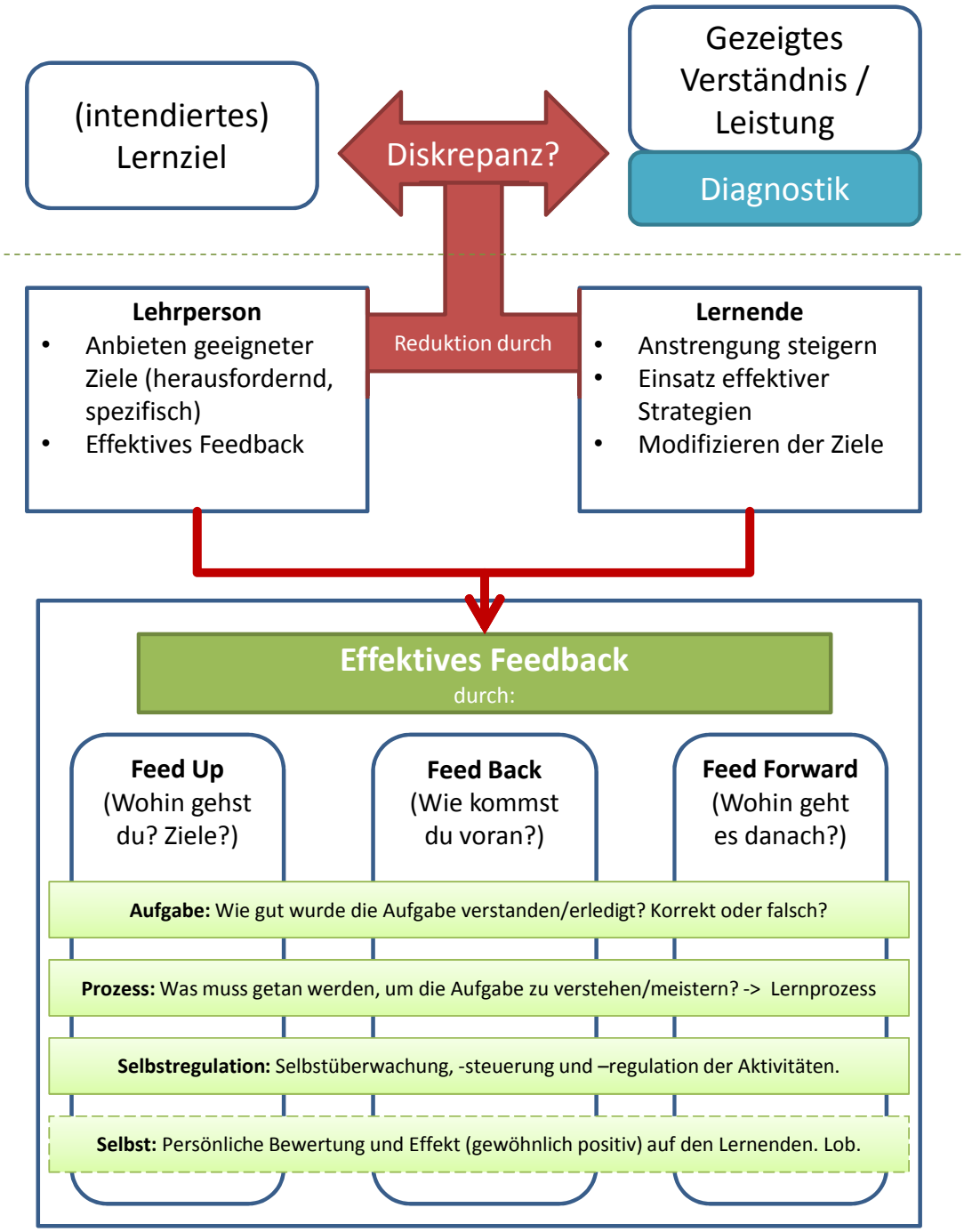


*Abbildung 10: Modifiziertes Feedbackmodell nach Hattie (2013, S. 209). Die Rolle der Diagnostik wurde ebenfalls durch den Autor ergänzt.*

Aus diesem Feedbackmodell wird deutlich, wie sehr Diagnostik (hier als Feststellung einer Lernzielerreichung) und Feedback zusammenhängen – und dass dies ein rekursiver Prozess ist, der weiter fortgeschrieben wird.

# DIAGNOSTIK MIT PSYCHOMETRISCHEN VERFAHREN

## *Standardisierte psychometrische Diagnostik hilft bei der Beantwortung bestimmter Fragestellungen*

Im Gegensatz zur didaktischen Diagnostik, die individuelle Lernfortschritte fokussiert, gibt standardisierte psychometrische Diagnostik Auskunft über bestimmte individuelle Merkmalsausprägungen anhand des Vergleichs mit Normstichproben. Dies ist nach wie vor erforderlich, wenn z.B. Platzierungsentscheidungen („behindert" vs. „nicht behindert") nachvollziehbar begründet werden müssen.

Um zu möglichst gültigen und zuverlässigen Aussagen zu kommen, sollte der diagnostische Prozess, d.h. die eingesetzten Methoden und die verwendete Vorgehensweise, bestimmten Gütekriterien genügen. Diese Kriterien sind der Grund, weshalb es sich auch im schulischen Kontext lohnt, sich mit standardisierter Diagnostik zu beschäftigen. Standardisierte Diagnostik umfasst Methoden, die diesen Kriterien genügen. Sie sind von Wissenschaftlern anerkannt und erforscht. Praktiker kennen sich damit aus und können sie einsetzen. Standardisierung meint, dass im diagnostischen Prozess

- Diagnostiker wissen, wie sie ein Verfahren einzusetzen haben und wie sie die Ergebnisse interpretieren sollen
- ein Verfahren von verschiedenen Personen eingesetzt werden kann und zu gleichen Ergebnissen kommt
- ein Verfahren möglichst genau misst
- ein Verfahren das misst, was es messen soll
- die Ergebnisse mit einer Normgruppe verglichen werden können, um Aussagen zu treffen, die die Testergebnisse des gemessenen Subjektes mit denen einer Bezugsnorm vergleichen.

Folgende Vorteile sind es, die den Einsatz im schulischen Kontext rechtfertigen oder gar notwendig machen:

- Standardisierte Diagnostik ermöglicht es, zuverlässige Aussagen über einen gemessenen Sachverhalt zu treffen. Die Messinstrumente sind so konstruiert, dass sie das messen, was sie messen sollen.
- Durch diese Genauigkeit bekommen Aussagen, die mit Hilfe standardisierter Diagnostik getroffen werden, ein hohes Maß an Sicherheit. Die häufige Unsicherheit, die bei Fragestellern und deren Auftraggebern auftritt, wird damit deutlich reduziert. Das, was man weiß, weiß man relativ sicher.
- Standardisierte Diagnostik hilft, Entscheidungen über die Beibehaltung oder das Verwerfen von Hypothesen zu treffen, die vorher zur Klärung einer Fragestellung generiert worden sind.

- Standardisierte Diagnostik ermöglicht einen Vergleich des Testergebnisses mit den Ergebnissen einer Normstichprobe. Da diese Stichproben repräsentativen Charakter haben, kann man den individuellen Testwert vergleichen mit z.B. „allen deutschen Kindern der Klasse 5“. Dieser Vergleich ermöglicht einen anderen Blickwinkel als der Vergleich mit einer kleineren Bezugsgruppe (also z.B. den Kindern der Klasse 5 der eigenen Schule).
- Aus diesen Gründen dient standardisierte Diagnostik auch immer der Qualitätssicherung und erleichtert die Kommunikation mit allen am Erkenntnisgewinn Beteiligten.
- Die Verfahren werden auch in außerschulischen Kontexten (Praxen, Kliniken) verwendet, was den interdisziplinären Austausch erleichtert.

Durch diese Vorteile kann standardisierte Diagnostik beispielsweise zur Klärung folgender Fragen herangezogen werden:

- Welche Potentiale hat der Schüler?
- Wo kommen Lern-/Leistungsprobleme von Schülerinnen und Schülern her? Beruhen sie auf mangelnder Fähigkeit oder mangelnder Anstrengungsbereitschaft?
- Sind Verhaltensprobleme von Schülerinnen und Schülern möglicherweise auf eine Unterforderung zurückzuführen?
- Wie sind die Fähigkeiten von Schülerinnen und Schülern in einem bestimmten Bereich (bspw. Rechtschreibung) im Vergleich zu seiner Altersgruppe ausgeprägt?
- Wie hoch oder niedrig ausgeprägt ist die Lernmotivation von Schülerinnen und Schülern im Vergleich zu anderen aus ihrer Klassenstufe?

Mit der Beantwortung dieser Fragen kommt man im Klärungsprozess bestimmter Sachverhalte schon recht weit. Oft ist es das Ausschließen einzelner Hypothesen, die dazu führen, dass neue Hypothesen generiert und überprüft werden können, um sich so der Lösung einer Fragestellung anzunähern.

**Beispiele für Fragestellungen, in denen standardisierte Diagnostik hilfreich ist:**

**1. Matthias**

Matthias geht in die 6. Klasse eines Gymnasiums. Seine Leistungen liegen im 4er-Bereich. Ein Hauptgrund sei seine chaotische Organisation und seine teilweise „schnippischen“ Bemerkungen in Situationen, in denen Leistungen erbracht werden sollen. Die Schule hat die Frage, ob Matthias für diese Schulform geeignet sei. Besonders im Matheunterricht zeigt er ab und zu überraschend gute Leistungen.

In einem standardisierten Intelligenztest erreichte er einen Gesamtwert von IQ = 124. Sein Profil ist dabei weitgehend homogen. In einem Fragebogen zur Leistungsmotivation beschreibt er sich selbst als sehr gering motiviert, seine eigenen Leistungen zu zeigen, hat aber hohe Lernziele.

Durch die Testdiagnostik wurde deutlich, dass Matthias überdurchschnittlich intelligent ist. Er möchte jedoch nicht, dass andere davon etwas mitbekommen. Daher verbirgt er dieses Potential geschickt im Unterricht - auch vor den Lehrern. Hinzu kommt, dass er zweisprachig aufgewachsen ist und daher Schwierigkeiten im Erlernen von Französisch (als dann dritte Sprache) hat.

In einer Intervention wurde mit ihm an der Klärung seiner eigenen Ziele gearbeitet. Daran anschließend wurden die Schritte besprochen, die zum Erreichen dieser Ziele nötig sein würden. Schlussendlich wurde mit ihm der dysfunktionale Mechanismus aus guter Begabung und dem Verstecken von dieser erläutert. Die „dritte“ Fremdsprache wird als „Lernlücke“ betrachtet und der Fokus auf die Stärken gelegt. Langsam verbesserten sich die allgemeine Situation und auch die schulischen Leistungen.

**2. Julia**

Julia ist Schülerin der 3. Klasse. Sie wirkt oft bedrückt. Ihre schulischen Leistungen sind seit der zweiten Klasse unterdurchschnittlich, besonders im Lesen und Rechtschreiben. Ab und zu hat sie „helle Momente“ und kann eine gute Leistung abrufen. Hypothese der Beratungslehrerin ist eine schulische Überforderung. Mittels einer standardisierten Intelligenztestung wird Julias kognitives Potential überprüft. Es zeigt sich eine gut durchschnittliche Leistung mit einem IQ von 110. Somit kann eine generelle Überforderung nicht der Grund für Ihre Leistungsprobleme sein. Im Elterngespräch stellen sich schwerwiegende häusliche Probleme heraus. Diese führen zu einer hohen Ängstlichkeit bei Julia. Die Beschäftigung mit familiären Dingen ist für sie wichtiger als schulisches Lernen. Dadurch erwarb sie einen Lernrückstand in den beiden Hauptfächern, der durch ihr Potential aber lange verschleiert geblieben war. Mittels standardisierter Rechtschreib- und Rechtentests werden ihre Lerndefizite identifiziert, um in diesem Bereich eine intensive Förderung zu planen. Julias Sorgen durch die Trennungsproblematik der Eltern werden mit Unterstützung durch einen niedergelassenen Kinder- und Jugendlichen-Psychotherapeuten bearbeitet.

Beide Beispiele zeigen, dass die jeweiligen Fragestellungen gut mit Hilfe eines standardisierten Intelligenztests beantwortet werden konnten. Damit wird deutlich, wie sinnvoll es sein kann, solche Verfahren zur Überprüfung von Hypothesen einzusetzen. Wichtig hierbei ist jedoch, dass dem Einsatz eine reflektierte Fragestellung zu Grunde liegt. Neben der Betrachtung von Einzelfällen kann es durchaus sinnvoll sein, standardisierte Diagnostik „flächendeckend“ anzuwenden. Auch hierzu ein Beispiel:

**3. Basisförderung an einem Gymnasium**

Der Fachgruppe Deutsch an einem Gymnasium fiel die teilweise sehr schlechten Rechtschreibleistungen der Schüler der 5. Klasse auf. Mit Hilfe einer standardisierten Testung aller neuen Fünftklässler sollten Schüler mit Defiziten in der Rechtschreib- oder Leseleistung identifiziert werden. Neben der Identifikation solcher Schüler war ein weiteres Ziel, über eine qualitative Auswertung Fördermaßnahmen möglichst gut individualisieren zu können. Mit den verantwortlichen Lehrern wurde folgendes Testkonzept besprochen:

- standardisierter Rechtschreibtest „Hamburger Schreibprobe“
- standardisierter Lesetest „ELFE“

Das letztgenannte Beispiel zeigt, dass einem Testeinsatz immer eine konkrete Fragestellung zugrunde liegen sollte, um gut planen und mit den vorhandenen Ressourcen sparsam umgehen zu können. Es stellt sich immer die Frage: „Und was machen wir mit den Ergebnissen?“ Im Falle des Gymnasiums hier ging es um die Identifikation und rechtzeitige Intervention bei Problemen im Rahmen von Schwierigkeiten in den Basisfähigkeiten Lesen und Schreiben. Sie bieten aber auch noch weitere interessante Information: Man hat einen Vergleich der Leistungen der Schüler einer Jahrgangsstufe. Zudem ist eine Verlaufsmessung möglich. Es kann also kontrolliert werden, ob die Förderung wirksam war. So kann auf Ebene einzelner Schüler nachgesteuert werden.

Gerade bei lernschwachen Schülern kann es interessant sein, deren Leistungen mit einem standardisierten Test in Bezug zu einer größeren Norm zu setzen, als nur der Vergleich mit der Klasse. Dies wird beispielsweise gestärkt durch die Forschungsergebnisse von Hoffmann & Böhme (2014), die zeigen konnten, dass das Lehrerurteil hinsichtlich der Schwierigkeitseinschätzung von Mathematik- bzw. Deutschaufgaben teilweise deutlich von der „echten“

Schwierigkeit abweichen kann, und es so zu Urteilsverzerrungen auf Klassenebene kommen kann (worunter schwächere Schüler möglicherweise stärker beeinträchtigt werden als stärkere). Verdeutlich wird dies durch folgendes Beispiel:

**4. Leni**

In einer sehr leistungsstarken Klasse fallen die schwachen Rechenleistungen der Schülerin Leni auf. Das Ergebnis eines standardisierten Rechentest ergab einen Prozentrang von PR = 27 (T-Wert = 44). Das heißt, das Ergebnis ist so gut wie das von mindestens 27% aller Schülerinnen und Schüler ihrer Jahrgangsstufe. Damit erreicht Leni ein noch durchschnittliches[14] Ergebnis. In einer leistungsschwächeren Klasse wäre sie damit (vermutlich) nicht aufgefallen. Die damit verbundene Frage nach Dyskalkulie konnte also verneint werden.

# *Konstruktionsprinzipien psychologischer Test*

Damit psychologische Tests überhaupt „funktionieren", müssen sie verschiedenen Qualitätsmerkmalen genügen. Diese Merkmale dienen dazu, den diagnostischen Prozess „sicherer" zu machen: Der Untersucher soll sich darauf verlassen können, dass das Testergebnis auch dem entspricht, was er davon erwartet. Um die Qualität der einzelnen Tests einschätzen zu können, ist es wichtig zu wissen (oder nachschlagen zu können), was die Begriffe bedeuten und wie die damit einhergehenden Messgrößen für jeden Nutzer zu interpretieren sind (vgl. z.B. Lienert, 1989).

## Hauptgütekriterien

Die wichtigsten Qualitätsnormen heißen „**Hauptgütekriterien**". Es sind:

- **Objektivität**: Ist ein Test nach objektiven Kriterien gestaltet, dann sind die Ergebnisse nach einer korrekten Testdurchführung unabhängig von Einflüssen der Untersucher, der Untersuchungssituation oder der Auswertung zustande gekommen.
- **Reliabilität (= Zuverlässigkeit)**: Die Reliabilität eines Tests gibt Auskunft darüber, wie zuverlässig das zu messende Merkmal gemessen und wie sehr diese Messung mit Messfehlern behaftet sein wird.
- **Validität (= Gültigkeit)**: Die Validität eines Tests macht Aussagen darüber, wie sehr das Verfahren tatsächlich das gewünschte Merkmal misst. Um in der Beantwortung der Fragestellung möglichst sicher zu gehen, sollte im Testhandbuch genau nachgelesen werden, welches Merkmal mit dem verwendeten Test gemessen wird. Möchte man beispielsweise die Rechenleistung eines Schülers der 3. Klasse messen, sollte man genau nachlesen, welche Fähigkeiten mit dem Test genau überprüft werden (Zahlenverständnis? Mengenverständnis? Addition, Subtraktion, Multiplikation? ...) und ob dies zum eigenen Erkenntnisinteresse passt.

Was hat man sich unter den „Testgütekriterien" genau vorzustellen? In Abbildung 11 findet sich am Beispiel „Schulnoten" ein Vergleich, der helfen soll, die Begriffe nachzuvollziehen.

Beispiel „Gütekriterien von Schulnoten“:

- **Objektivität:** Vergeben verschiedene Lehrer bei gleicher Leistung die selbe Note? Oder fließen subjektive Verzerrungen ein?
- **Reliabilität/Zuverlässigkeit:** Werden gleiche Leistungen verschiedener Schülerinnen und Schüler in einer Arbeit gleich bewertet? Sind Leistungsunterschiede von Schülerinnen und Schüler auch in unterschiedlichen Noten repräsentiert?
- **Validität/Gültigkeit:** Ist die Note auch aussagekräftig für einen entsprechenden Wissensstand in einem Fach (oder misst die Note z.B. eher die Verhaltensauffälligkeiten im Unterricht)?
- **Fairness:** Gibt es bei der Benotung Benachteiligungen für bestimmte Personengruppen (z.B. Geschlechtsunterschiede oder Schüler mit Migrationshintergrund, die Sprachprobleme haben)?

*Abbildung 11: Beispiel für die Anwendung der Testgütekriterien auf das Beispiel „Noten“*

Die Hauptgütekriterien kann man auch statistisch ausdrücken. Die Kennwerte geben an, wie sehr der jeweilige Test dem Gütekriterium entspricht. Dabei bedeutet „0“ entspricht nicht dem Kriterium, „1“ = entspricht vollständig dem Kriterium. Die Werte können zwischen diesen beiden Extremen abgebildet werden. Je näher der Wert an „1“ ist, desto besser. In der Regel sind Kennwerte ab 0,75 als gut und brauchbar zu bezeichnen. Tabelle 2 gibt eine kurze Übersicht über die Hauptgütekriterien mit ihren jeweiligen Kennwerten:

*Tabelle 2: Hauptgütekriterien und wichtige statistische Kenngrößen dazu (vgl. Lienert, 1989)*

| Gütekriterium | Kurzbeschreibung | Statistischer Kennwert |
|---|---|---|
| Objektivität | Grad, in dem die Ergebnisse des Tests unabhängig vom Untersucher sind | Kontingenz- oder Korrelationskoeffizient (als Maß für die Untersucherübereinstimmung). Variiert zw. 0 und 1. Sollte über 0,7 liegen. |
| Reliabilität | Grad der Zuverlässigkeit, mit dem der Test ein bestimmtes Merkmal misst. | Korrelationskoeffizient. Variiert zw. 0 und 1. Sollte über 0,8 liegen. |
| Validität | Grad der Genauigkeit, mit dem der Test das misst, was er messen soll. | Korrelationskoeffizient. Variiert zw. 0 und 1. Sollte über 0,5 liegen. |

## Nebengütekriterien

Weiterhin gibt es sogenannte **„Nebengütekriterien“**. Diese lassen sich nicht so gut in statistischen Kennwerten abbilden. Dennoch ist es sinnvoll, sich bei der Entscheidung für einen Test anzuschauen, welche Aussagen die Testautoren zu diesen Werten treffen:

- **Normen:** In der Regel vergleicht ein Test den Wert, den ein einzelner Proband erreicht hat, mit Werten von Probanden aus einer ähnlichen Stichprobe (z.B. alle Schülerinnen der 8. Klasse). So kann eine Aussage darüber gemacht werden, wie ein Proband im Verhältnis zu anderen einzuschätzen ist. Die Norm dient als Bezugssystem. Die Norm ist die wichtigste diagnostische Alltagsgröße im diagnostischen Prozess. Um die Güte eines Testes einschätzen zu können, ist es auch hilfreich, sich einmal anzusehen, an welcher Stichprobe und in welchem Zeitraum die Normierung stattgefunden hat. Aus welchen Bundesländern kamen beispielsweise die Schüler? Gab es besondere Regionen oder wurde flächendeckend getestet etc.?
- **Vergleichbarkeit:** Ein Test ist vergleichbar, wenn es ähnliche Tests gibt, die das gleiche oder ein sehr ähnliches Konstrukt messen. Aus der Vergleichbarkeit kann man ableiten, wie gut die Validität eines Tests ist.
- **Testökonomie:** Ein Test ist dann ökonomisch, wenn er schnell und einfach durchzuführen ist, wenig Material verbraucht und schnell und bequem auszuwerten ist. Je nach diagnostischem Kontext ist diese Größe wichtig für die Testauswahl (z.B. bei der Frage ob der Test als Gruppen- oder Einzeltests durchgeführt werden soll).
- **Nützlichkeit:** Ein Test ist nützlich, wenn er etwas misst, für das ein praktisches Bedürfnis besteht (und was von keinem anderen Test gemessen wird).
- **Testfairness:** Die Testfairness gibt an, wie sehr das Testergebnis systematisch beeinflusst wird von Faktoren der ethnischen, soziokulturellen oder geschlechtsspezifischen Gruppenzugehörigkeit der Testperson. Dieses Kriterium ist besonders dort von Bedeutung, wo Auswahlentscheidungen aufgrund des Testergebnisses getroffen werden. Besonders häufig wird Testfairness untersucht in Bezug auf Geschlecht, sozioökomischen Status, Bildungsnähe, Migrations- und sprachlichem Hintergrund.

**Schlechtes Beispiel für Testfairness:**

In einem bundesweit an allen Schulen durchgeführten Test zum Allgemeinwissen wird im Bereich „Erdkunde“ folgende Frage gestellt: „Wie heißt die Hauptstadt von Rheinland-Pfalz?“ Schülerinnen und Schülern aus diesem Bundesland haben hier einen Vorteil: Sie kennen mit größerer Wahrscheinlichkeit die Hauptstadt ihres eigenen Bundeslandes. Die Schwierigkeit der Frage ist somit abhängig von der Gruppe. Damit könnte der (falsche!) Eindruck entstehen, dass Schülerinnen und Schülern aus Rheinland-Pfalz besser in Erdkunde seien als solche aus anderen Bundesländern.

## Normvergleich

Die Normierung eines Tests ist die wichtigste Größe, um diagnostische Aussagen zu treffen. Der Normwert dient dazu, die gemessenen Eigenschaften (z.B. Intelligenz, Leistung in einem Rechentest...) einer Person mit einer bestimmten Gruppe in Beziehung zu setzen. Diese „Gruppe“ bilden bei Schultests in der Regel die Klassenstufen (also „alle Schülerinnen und Schüler der Klasse 7“). Es können auch Subgruppen gebildet werden („alle Mädchen der Klasse 7“). Durch den Vergleich einer Person mit einer Norm kann man eine Aussage darüber machen, wie sich ein Merkmal der Person im Vergleich zur Gruppe verhält („X hat eine schlechtere Rechenleistung als der Durchschnitt der Schüler ihrer Klassenstufe.“)

Normwerte gewinnt man, in dem nach der Testkonstruktion eine ausreichend große Stichprobe (in der Regel über 500 Personen) mit diesem Test untersucht wird. Mittels

statistischer Verfahren (Berechnung von Mittelwert, Standardabweichung, Prozenträngen) werden diese Normen erstellt und in Tabellen abgebildet. Dabei wird häufig eine quasi kontinuierliche Verteilung („Normalverteilung“, siehe Abbildung 12) der Ausprägung der Eigenschaft angenommen.

Es existieren verschiedene Normwertskalen, die direkt ineinander umgerechnet werden können. Die bekannteste dieser Skalen ist die „IQ-Skala“. Ein IQ-Wert von 100 bedeutet ein durchschnittliches Ergebnis. Weitere Skalen werden in Tabelle 3 und Abbildung 12 vorgestellt und verglichen. Von diesen Skalen unterscheidet sich der oft benutzte Prozentrang (PR). Dieser gibt an, welcher Prozentsatz der Normstichprobe einen gleich hohen oder kleineren Wert aufweist. Ein PR von 50 bedeutet ein durchschnittliches Ergebnis, ein PR von 10 dagegen, dass nur 10% der Normstichprobe ein gleiches oder kleineres Ergebnis hatten. Die Prozentrangskala wird oft benutzt. Jedoch führt sie leider oft zu Fehlurteilen , da sie sehr hohe oder niedrige Werte suggeriert. So ist der PR = 20 noch ein durchschnittliches Ergebnis, wird aber oft schon als deutlich unterdurchschnittlich eingeschätzt. Dieser Denkfehler kommt zustande, weil sich der PR auf die Fläche unter der Normalverteilung bezieht, die anderen Normwerte jedoch den kontinuierlichen Verlauf der X-Achse abbilden. Dieser Zusammenhang wird in Abbildung 12 veranschaulicht.

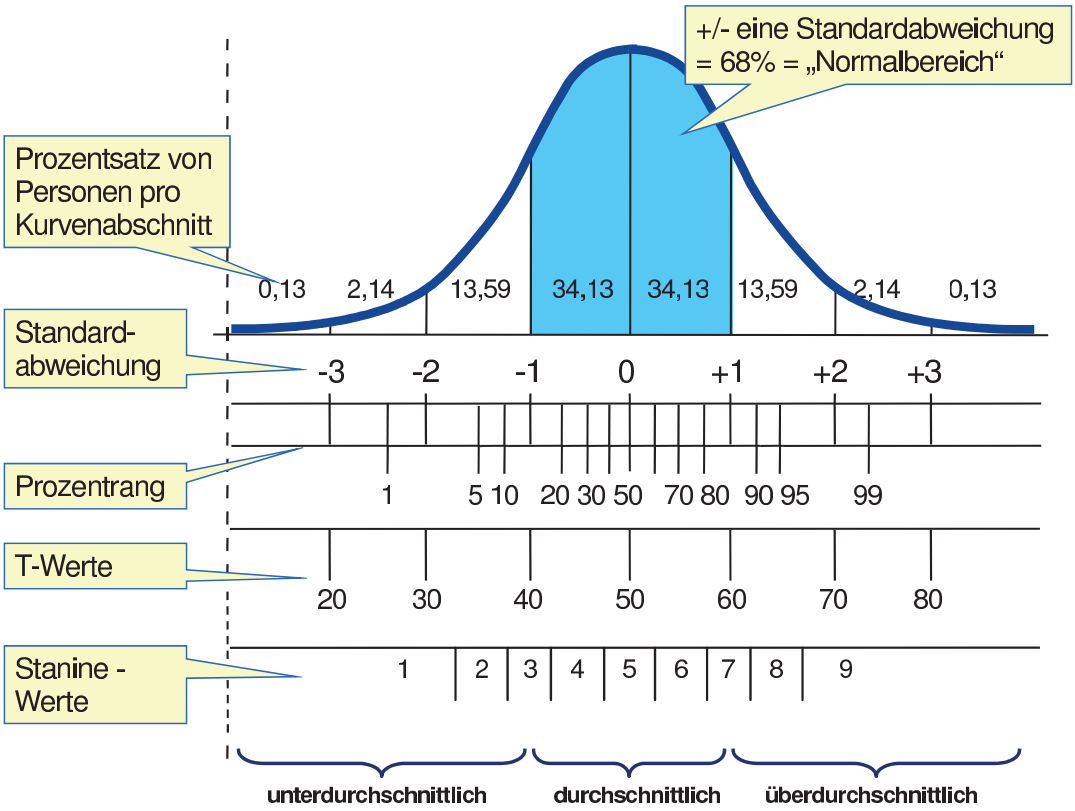


*Abbildung 12: Vergleich von Normmaßen anhand der Normalverteilung*

Welche Normskala letztlich verwendet wird ist beliebig. Wichtig ist allerdings, dass verschiedene Werte zum Vergleich in derselben Norm vorliegen. Für Intelligenztests hat sich beispielsweise die IQ-Norm weitgehend durchgesetzt. Manche Tests verwenden allerdings dennoch andere Normen, so greift z. B. das AID auf T-Werte zurück, weil das Konzept dieses Tests den Gesamt-IQ als Mittelwert der Untertests ablehnt. In Tabelle 3 findet sich eine kurze Übersicht über die gebräuchlichsten Normwerte und deren statistischen Kennwerte.

*Tabelle 3: Wichtige Norm-Maße im Vergleich*

| **Norm-Maß** | **Kurzbeschreibung** | **Statistische Kennwerte** | | | |
|---|---|---|---|---|---|
| | | Mittelwert | Standard-abweichung | Skala (von – bis) | „Normal-bereich“ |
| **Prozentrang (PR)** | Globales Maß der Einordnung auf einer Rangreihe mit einer Skala von 0 bis 100% | 50 | +/- 34 | 0 - 100 | 16-84 |
| **z** | Skala auf Basis der Kennwerte der Standardnormalverteilung | 0 | +/- 1 | -3 - +3 | -1-+1 |
| **T** | Häufig verwendete Skala | 50 | +/- 10 | 20 - 80 | 40-60 |
| **Stanine** | „Standard-Nine“, Skala aus dem angloamerikanischen Raum mit einer 9-teiligen Skala | 5 | +/- 2 | 1 - 9 | 3-7 |
| **IQ** | Wert, der nur für das Konstrukt Intelligenz gilt (!) | 100 | +/- 15 | 40 - 160 | 85-115 |

Bei der Interpretation werden Bewertungen wie „Normal“, „Durchschnitt“, „Über Durchschnitt“ oder „Unter Durchschnitt“ usw. aus der Normalverteilung abgeleitet: Eine Standardabweichung über und unter Mittelwert umfasst den „Normalbereich“ (z.B. IQ-Skala 85-115).

# *Grenzen standardisierter Diagnostik*

Die Gütekriterien verleiten dazu, standardisierten Tests einen „Absolutheitsanspruch“ zuzuschreiben, den sie aber nicht beanspruchen. Das ist immer der Fall, wenn psychologische Variablen gemessen werden. Daher ist es sinnvoll, sich der Grenzen bewusst zu sein. Dies gilt vor allem für die Interpretation der Testergebnisse. In Anlehnung an Rackwitz (2011) kann man Begrenzungen standardisierter Diagnostik grob in vier Punkte einteilen:

1. Einschränkung in den Aufgabentypen

   Aufgrund der komplizierten statistischen Berechnungen, die bei der Konstruktion standardisierter Tests durchgeführt werden müssen, ist die Nutzung bestimmter Aufgabentypen eingeschränkt. Es bieten sich richtig-falsch-Zuordnungen, Auswahlzuordnungen, Fehleranzahlen, Lösungszeiten etc. als Aufgabentypen an. Diese bilden jedoch oft nicht das Fähigkeitsspektrum der Schülerinnen und Schüler ab. Es kann dann sinnvoll sein, die standardisierte Testung um offene Aufgabenformen zu ergänzen und beim diagnostischen Urteil beide Informationsquellen miteinfließen zu lassen.

2. Prozessbetrachtung

   Die Probleme des eingeschränkten Aufgabentyps führen auch dazu, dass der Lösungsprozess der Schülerinnen und Schüler nur unzureichend abgebildet werden kann. Man kennt dies aus dem Mathematikunterricht: Oft können beim Lösen einer Klausuraufgabe noch Punkte für den Lösungsweg erreicht werden. Bei der standardisierten Testung zählt nur das Ergebnis. Wir wissen aber nicht, was die Schülerinnen und Schüler verstanden haben vom Lösungsweg und was nicht. Da die Art und Weise, wie Schülerinnen und Schüler zu Lösungen kommen, aber ein wichtiger Faktor in der Förderplanung ist, sollten standardisierte Tests nicht (alleine) zur Planung einer Förderung herangezogenen werden. Eine sogenannte qualitative Auswertung ist unabdingbar. Nur so können die Erkenntnisse über die Denkweise der Schülerinnen und Schüler, die Zone des aktuellen Entwicklungsstandes und die Zone der nächsten Entwicklung identifiziert und geplant werden.

   Im Kapitel "Beispiele für curriculumsbasierte Messungen" wurden solche Verfahren vorgestellt. Im Idealfall bieten Testverfahren auch immer die Möglichkeit zu einer solchen qualitativen Auswertung.

3. Blick auf die Lernbiographie

   Wie der Blick auf den Lösungsweg ist es notwendig, die gesamte Lernbiographie des Schülers heranzuziehen. Gibt es Bereiche, die ihm immer schon Schwierigkeiten bereitet haben? Hat man bei der Förderplanung bewusst auf bestimmte Inhalte verzichtet, da diese noch zu schwer gewesen wären? Werden solche Inhalte bei kommenden Aufgaben vorausgesetzt? Diese und weitere Fragen sind bei der Bewertung von Testergebnissen sehr hilfreich und dienen der korrekten Einschätzung der Lernleistung.

4. Stärken nicht aus dem Blick verlieren

   In unserer Gesellschaft haben wir uns angewöhnt, eher auf die Schwächen und Lücken zu achten, als auf das, was wir gut können. Das ist in der Diagnostik nicht anders. Schnell wird darauf geachtet, was „links“ vom Normalen ist. Dabei werden so sehr schnell wichtige Ressourcen in der Intervention übersehen. Daher soll der Diagnostiker auch **immer die Stärken und Ressourcen der Schülerinnen und Schüler im Blick haben**, da sich eine mögliche Förderung gut an diese anschließen lässt. Dazu dienen Profilauswertungen von Tests.

Die oben genannten Punkte machen deutlich, dass es trotz aller Vorteile Grenzen von

standardisierten Verfahren gibt. **Die Stärke liegt sicherlich in der Anwendung** – es ist sofort klar, was man wie machen muss, wie ein Test auszuwerten ist und wie der Testwert des Schülers einzuordnen ist. Dadurch ist es möglich, einen „Fremdblick" auf den Schüler zu bekommen, der vielleicht eigene Vorurteile und Vorannahmen korrigiert – oder bestätigt.

Als Fazit kann man formulieren: **Standardisierte Diagnostik ist sinnvoll, wenn sie reflektiert und mit einer konkreten Fragestellung verbunden eingesetzt wird.** Zur Förderplanung sollte sie ergänzt werden durch qualitative Verfahren. Umgekehrt sollte Förderplanung nicht auf ungeprüften Mutmaßungen beruhen.

# *Standardisierte Diagnostik in der Schule*

In der Schule wird an sich „ständig" Diagnostik betrieben. Jede Klassenarbeit, jeder Test dient der Beantwortung der Frage: Welchen Lernstoff beherrschen Schülerinnen und Schüler? Wie gut schneiden sie (im Vergleich zur Klasse) ab? Natürlich können Lehrerinnen und Lehrer hier keine standardisierten und normierten Verfahren entwickeln, die den Qualitätsmerkmalen der Gütekriterien genügen. Und doch lohnen sich hier ein Vergleich und eine detailliertere Beschäftigung.

Der in Tabelle 4 dargestellte Vergleich soll helfen, die verschiedenen Testarten in der Schule einordnen zu können.

*Tabelle 4: Vergleich von Testformen in der Schule*

| Kategorie | Beschreibung | Norm? | Beispiel |
|---|---|---|---|
| **Leistungstest** | Standardisierte Testverfahren für unterschiedliche Leistungsbereiche in der Schule (i.d.R. Lesen, Schreiben, Rechnen) | Ja (in der Regel). Vergleich mit jeweiliger Klassenstufe | • Stolperwörterlesetest<br>• Hamburger Schreibprobe<br>• DEMAT<br>• ... |
| **Klassenarbeit** | Vom Lehrer selbst entwickelte Aufgabensammlung in gestufter Schwierigkeit zur Überprüfung des unterrichteten Lernstoffes | Nein | • Überprüfung von Bruchrechnen<br>• Analyse von Gedichten<br>• ... |
| **„Test"** | Kürzere Aufgabensammlung zu einem isolierten Unterrichtsthema | Nein | • Vokabeltest<br>• Diktat<br>• ... |
| **CBM** | Curriculum-basierte-Messung | Nein | • Prä- und Posttest bei Matinko<br>• Diagnosediktate der Rechtschreibwerkstatt<br>• ... |
| **Förderdiagnostik** | Intention des jeweiligen diagnostischen Werkzeugs ist es, eine Förderplanung zu ermöglichen | teilweise | Qualitative Auswertung standardisierter Verfahren, differenzierte Analyse von Teilbereichen (z.B. BADYS, BAKO...) |

Wie können die dargestellten Testformen nun mit Blick auf die Einsatztauglichkeit in der Schule bewertet werden?

- **Leistungstest:** Durch den Vergleich z.B. mit einer Jahrgangsnorm bekommt man ein Bild über die Leistung eines Schülers im Vergleich zur Norm. Dies kann helfen festzustellen, wie der generelle Leistungsstand des Schülers bei diesem Merkmal (z.B. Rechtschreibleistung) ist. Das Testergebnis ist hierbei objektiv. Es können jedoch nur sehr eingeschränkt Aussagen zum Wissensstand im Unterrichtsstoff bzw. der individuellen Förderung abgegeben werden.

- **Klassenarbeit:** Die Schwierigkeitsstufe können Lehrerinnen und Lehrer steuern, Schwierigkeit und Bewertung sind stark von dessen subjektivem Eindruck geprägt. Objektivität ist nur in der Auswertung gegeben, jedoch auch hier eingeschränkt und abhängig von den zuvor festgelegten Kriterien. Das Ergebnis sagt etwas über das Lernen des abgetesteten Stoffes aus, kann aber keine umfassenderen Aussagen zulassen.

- **„Test“:** Hier gilt die gleiche Einschränkung wie bei der Klassenarbeit. Da aber kleinere Stoffmengen überprüft werden (die sich „dichter“ an den vorangegangenen Unterrichtsstunden befinden), sind gute Aussagen über entstehende Lernlücken und noch aufzuarbeitenden Stoff möglich. Hier gilt die Regel: Besser mehrere kleine Tests als wenige größere.

- **CbM:** Vergleichbar mit einem „Test“. Da dem CbM jedoch ein Messkonzept über eine Unterrichtsreihe zugrunde liegt, haben diese eine hohe Aussagekraft bei der Erreichung des jeweiligen Lernziels und dem Gesamtcurriculum. Zudem ist für Lehrende und Lernende immer deutlich, wo der Lernende im Bezug auf den Lernstoff steht.

- **Förderdiagnostik:** Häufig benutzter Begriff. Im Begriff steckt die Intention: die diagnostischen Ergebnisse sollen eine Förderplanung ermöglichen. Daher ist die Diagnostik differenzierter als z.B. bei den CbM. Im Kontext von Inklusion ist dieser Begriff hinfällig (vgl. Begriffsklärung von Diagnostik).

Bei der Testinterpretation standardisierter Tests sollte auf folgende Aspekte geachtet werden:

- **Testzeitraum:** Die standardisierten Leistungstests sind z.B. auf Klassenstufen normiert. Diese Norm bezieht sich jedoch auf einen Zeitraum im Schuljahr. So ist beispielsweise der Rechtschreibtest DRT2+ auf das Schuljahresende der zweiten Klasse und den Beginn der dritten Klasse normiert. Eine Testdurchführung und Auswertung zu anderen Zeiträumen um Vergleichsaussagen zu treffen ist nicht zulässig. Testet man z.B. eine Schülerin zu Beginn der zweiten Klasse, wird sie mit großer Wahrscheinlichkeit einen deutlich schlechteren Wert erhalten als am Ende der zweiten Klasse. Das ist natürlich zu erwarten, da nahezu der Stoff der gesamten zweiten Klasse noch nicht vermittelt worden ist, aber abgefragt wird.

- **Vergleichbarkeit:** Die Vergleichswerte, die ein standardisiertes Verfahren liefert, beziehen sich auf eine „Population“ – also theoretisch auf „alle Schüler der Klasse X in Deutschland“. Regional (also z.B. in der Astrid-Lindgren Schule in Bullerbü) können davon aber deutliche Abweichungen auftreten (vgl. Beispiel von Leni oben). Dies kann mit der Einwohnerstruktur des Einzugsgebietes der Schule zusammenhängen oder auch mit den verwendeten Lehrmaterialien. Auch daher sind die Werte, die man erhält, immer in einen „Kontext“ einzuordnen.

- **Standard-Schulnoten:** Manche Tests geben sog. „Standard-Schulnoten“ aus. Hierzu wird eine 5-stufige Notenskala zugrunde gelegt. Diese hat den Wert „3“ als Mittelwert und eine Standardabweichung von „1“. Die Testautoren haben hierzu lediglich Normwerte (z.B. T- oder z-Werte) „linear transformiert“. Diese Standard-Noten können daher von den von Ihnen verwendeten (möglicherweise kriteriumsbasierten) Noten abweichen.

Falls Sie also einen solchen Test zur Leistungsbewertung heranziehen, sollten Sie sich diesen Unterschied verdeutlichen.

- **Standardschätzfehler:** standardisierte Tests geben einen „Standardschätzfehler" an, der sagt, mit welcher Wahrscheinlichkeit ein Testwert im Bereich eines bestimmten „Vertrauensintervalls" (sog. Konfidenzintervall) liegt. Das könnte dann z.B. wie folgt lauten: „Mit 95% Wahrscheinlichkeit liegt das wahre Testergebnis im Bereich von 96 – 104." Dieser Wert hilft, Testergebnisse mit möglicherweise abweichenden Alternativergebnissen zu vergleichen und zu einer Einschätzung zu kommen, wo der „wahre Wert" eines Testergebnisses angesiedelt ist.

# *Standardisierte Diagnostik in der Schule ist sinnvoll*

Im bisherigen Verlauf dieser Broschüre wurde schon mehrfach auf Vor- und Nachteile von standardisierter Diagnostik eingegangen. Grundsätzlich sollte die Schule diesem Thema positiv gegenüberstehen, da diagnostische Aspekte in der Schule alltäglich sind (herkömmliche Leistungsfeststellung) und sein sollten (Lernfortschrittsdiagnostik) und zudem im Kontext von Inklusion ein wichtiges Handwerkszeug darstellen. So schreibt Hartke (2012, S. 62): „Diagnostik und Inklusion passen ausgesprochen gut zusammen." Als Gründe nennt er unter anderem:

- „Effektive individuelle Förderung (...) verlangt eine gültige – valide – Beschreibung der aktuellen Fähigkeiten des Kindes oder Jugendlichen als Grundlage für ein passendes Lernangebot.
- Bleiben Lernerfolge oder Entwicklungsfortschritte in Bereichen wie Sprachverständnis oder soziale Kompetenz aus, gilt es hierfür Erklärungen zu finden, um Ansatzpunkte zur Neubestimmung der Art der Förderung zu ermitteln."

Demgegenüber steht eine historische gewachsene Antipathie, die möglicherweise in der Darstellung von Eggert (1997) fußt:

*„Die mit den klassischen psychometrischen Tests quantitativ gemessenen Fähigkeiten lassen in der Regel keine verlässlichen Hinweise auf Ursachen von Lernproblemen oder die zukünftige Lernentwicklung eines Kindes noch auf detaillierte Möglichkeiten der Förderung zu."*

Damit begründet er die Abkehr u.a. in der Sonderpädagogik von standardisierter Diagnostik hin zur „Förderdiagnostik". Nicht nur im Kontext der Inklusionsdebatte wurde gezeigt, dass es Situationen geben kann, in denen standardisiert gemessene Werte, auch mit einem Normvergleich, ihre Berechtigung haben. Dann nämlich, wenn eine „Standortbestimmung" erforderlich ist oder wenn grundlegende Entscheidungen darüber getroffen werden müssen, welche Förderung sinnvoll ist (und welche nicht). In einem Bericht über die Qualität von sonderpädagogischen Gutachten in Mecklenburg-Vorpommern kommen deren Autoren zu folgenden Schlüssen[15]:

*„Die Einschätzung der Schulleistung erfolgt häufig ohne adäquate Messungen. Damit wird ein Mindestkriterium gutachterlicher Dignität nicht erfüllt, welches die Voraussetzung für das Erfüllen von anderen Qualitätsstandards darstellt."*

Die Autoren stufen sonderpädagogische Gutachten als „entwicklungsbedürftig“ ein und verweisen hier auf eine „unzureichende Nachvollziehbarkeit der getroffenen Entscheidungen“ (S. 167). An anderer Stelle (S. 164) stellen sie fest,

*„… dass Schulleistungsprüfungen relativ selten auf standardisiertem Niveau stattfinden. Dies ist insofern erwähnenswert, weil Schulleistungen im Überprüfungsanlass oder in den Fragestellungen in 40% aller Gutachten aufgeführt werden.“*

Damit wird schon eine mögliche Lösung der Mängel aufgezeigt: Der Einsatz standardisierter Diagnostik. Die Verfahren eignen sich deshalb, weil sie eine gute Basis für den Prozess der Urteilsfindung und damit einiges an Sicherheit und möglicherweise juristischer Belastbarkeit bieten.

Stellt man beide Standpunkte gegenüber, dann wird deutlich, dass es bei der Frage, nach dem Einsatz standardisierter Diagnostik in der Schule **nicht um ein ja oder nein geht, sondern es darauf ankommt, was das Ziel des Diagnostikers** sein soll. Die eingesetzten Methoden sollen dem Ziel dienen zu helfen, die Fragestellung des Diagnostikers zu beantworten. Dem Einsatz der ausgewählten Verfahren soll eine angemessene Kosten-Nutzen-Relation zugrunde liegen. Hier können qualitative Verfahren, quantitative, oder auch „anekdotische“ genutzt werden.

# FAZIT & AUSBLICK

Diese Broschüre „rund um“ das Thema Diagnostik hat die Aufgabe, Hintergrundinformationen, Basiswissen und eine Implementierung von diagnostischem Arbeiten in der Schule vorzustellen. Es sollte verdeutlicht werden, dass Diagnostik ein hilfreiches Handwerkszeug ist, mit deren Hilfe man den Lehr-Lern-Herausforderungen des schulischen Alltags besser begegnen kann.

Gerade im Kontext Inklusion ist es hilfreich, sich die diagnostischen Prozesse an der eigenen Schule zu vergegenwärtigen. Dabei sollte hinterfragt werden, was noch zeitgemäß ist, was neu gedacht werden sollte und was gut funktioniert. Das heißt, Diagnostik ist ein guter Baustein für Inklusion an der eigenen Schule. Es können hier viele Probleme und Herausforderungen sichtbar gemacht werden, Diagnostik ist dabei aber konkret genug, um Veränderungsprozesse deutlich zu machen und anzustoßen. Daher wurde der Inklusion ein großer Schwerpunkt in der vorliegenden Broschüre eingeräumt.

*Nochmal sei hier deutlich darauf verweisen, was Diagnostik in der Schule ist (oder sein sollte):*

| Inklusive Diagnostik ist/sollte sein | Diagnostik ist nicht/sollte nicht sein |
|---|---|
| • **eine hilfreiche Methode**<br>• **unterstützend für Lehrende und Lernende**<br>• **inklusiv**<br>• **essentiell zur Steuerung von Lernprozessen**<br>• **motivierend**<br>• **informativ (Feedback)** | • **diskriminierend**<br>• **wait to fail**<br>• **selektierend**<br>• **kategorisierend**<br>• **klassifizierend**<br>• **etikettierend**<br>• **Bedingung für Ressourcenzuweisung** |

Wer Inklusion im Schulsystem wirklich effektiv umsetzen will, der kommt um diagnostische Methoden nicht herum, weil sie erst individuelle Förderplanung ermöglicht. Durch die größere Heterogenität, die stärkere individuelle Sichtweise und Förderung mussen allen Beteiligten Informationen darüber vorliegen, wo sie im Lernprozess stehen und was die nächsten lernförderlichen Schritte sind. Diese Informationen können nur eine inklusive – und damit didaktische – Diagnostik liefern (Schumann, 2013).

# LITERATUR


Dummer-Smoch, L. & Hackethal, R. (2007). Kieler Leseaufbau. Handbuch und Übungsmaterialien. Kiel: Veris.

Eberwein, H. (1997): Systemisch-ganzheitliche Diagnostik in der Schule. In: R. Voß (Hrsg). Die Schule neu erfinden. Luchterhand. S. 223-229.

Eggert, D. (1997): Von den Stärken ausgehen...: Individuelle Entwicklungspläne in der Lernförderungsdiagnostik. Dortmund: Borgmann.

Ferri, B. (2011). Undermining inclusion? A critical reading of response to intervention (RTI). In: International Journal of Inclusive Education. 1-18.

Gold, A. (2011). Lernschwierigkeiten. Ursachen, Diagnostik, Intervention. Stuttgart: Kohlhammer.

Hartke, B. (2012): Besteht ein Widerspruch zwischen Inklusion und Diagnostik? In: M. Brodkorb & K. Koch (Hrsg.): Das Menschenbild der Inklusion. Erster Inklusionskongress M-V – Dokumentation. Rostock.

Hattie, J. (2013). Lernen sichtbar machen. Überarbeitete deutschsprachige Ausgabe von Visible Learning. Besorgt von W. Beywel und K. Zierer. Schneider-Verlag Hohengeren.

Hartmann, E. & Müller, C.M. (2009). Schulweite Prävention von Lernproblemen im RTI-Modell. Schweizerische Zeitschrift für Heilpädagogik. Jg. 15, 9 /09.

Hinz, A. (2013). Inklusion – von der Unkenntnis zur Unkenntlichkeit!? - Kritische Anmerkungen zu einem Jahrzehnt Diskurs über schulische Inklusion in Deutschland. Zeitschrift Für Inklusion, 0(1). http://www.inklusion-online.net/index.php/inklusion/article/view/201/182.

Hoffmann, L. & Böhme, K. (2014). Wie gut können Grundschullehrkräfte die Schwierigkeit von Deutsch- Mathematikaufgaben beurteilen? Psychologie in Erziehung und Unterricht, 61, 42-55.

Huber, C. & Grosche, M. (2012). Das response-to-intervention Modell als Grundlage für einen inklusiven Paradigmenwechsel in der Sonderpädagogik. Zeitschrift für Heilpädagogik 63, 312–322.

Jansen, P. (2010): Matinko - Mathematik individualisiert und kompetenzorientiert. Ein Lehrgang zur Individualisierung des Klassenunterrichts in der Schuleingangsphase. Coesfeld: Fachberatung für mathematisches Lernen.

Klauer, K.-J. (2006). Erfassung des Lernfortschritts durch curriculumsbasierte Messung. Heilpädagogische Forschung (32; 1), 16-26.

Lenhard, W. & Lenhard, A. (2010). Rechenspiele mit Elfe und Mathis 1. Göttingen: Hogrefe.

Liebers, K. & Seifert, C. (2012). Assessmentkonzepte für die inklusive Schule – eine Bestandsaufnahme. http://www.inklusion-online.net/index.php/inklusion/article/view/166/156

Lienert, G. A. (1989). Testaufbau und Analyse. Weinheim: PVU.

Lohri, S. & Steiner. F. (2015). Leistungstief adieu! – Eine Schule auf Erfolgskurs. http://

lsm2014.lernensichtbarmachen.ch/wp-content/uploads/2015/06/02_SN-Lohri-Steiner.pdf

Lyon, G.R. et al. (2001). Rethinking learning disabilities. In C.E. Finn, A.J. Rotherman & C.R. Hokanson (eds.), Rethinking special education for a new century

(S. 259–287). Washington: Fordham Foundation and the Progressive Policy Institute.

Prengel, A. (2015). Didaktische Diagnostik als Element alltäglicher Lehrerarbeit – „Formatives Assessment" im inklusiven Unterricht. In: B. Amrhein, K. Ziemen (Hrsg.). Diagnostik im Kontext inklusiver Bildung – Theorien, Ambivalenzen, Akteure, Konzepte. Bad Heilbrunn: Klinkhardt.

Prengel, A. (2013). Inklusive Bildung in der Primarstufe. Grundschulverband. http://www.grundschulverband.de/fileadmin/bilder/Publikationen/Mitgliederbaende/Prengel_Kurzfassung.pdf

Quinn, P. (2012). Response to Intervention for High Schools. www.TotalRTI.com

Rackwitz, R.-Ph. (2011): Dialogische Lernbeobachtung statt standardisierter Tests. gekürzter und leicht geänderter Nachdruck aus Pädagogik, Heft 5/2011, in: BLZ, Heft 06/07-2011, S. 10 – 11.

Reuter-Liehr, C. (2008). Lautgetreue Lese-Rechtschreibförderung. Bochum: Winkler.

Rheinberg, F. (2001). Bezugsnormen und schulische Leistungsbeurteilung. In: F. E. Weinert (Hrsg.): Leistungsmessung in Schulen. Weinheim: Beltz.

Schlee, J. (2012). Was kann und sollte Diagnostik in einer „inklusiven Pädagogik leisten? In: M. Brodkorb & K. Koch (Hrsg): Das Menschenbild der Inklusion. Erster Inklusionskongress M-V – Dokumentation. Rostock.

Schnotz, W. (2006): Textverständnis. In: D. Rost (Hrsg): Handwörterbuch Pädagogische Psychologie. Weinheim: Beltz PVU.

Schumann, B. (2013). Inklusive Bildung braucht inklusive Diagnostik. Bildungsklick: http://bildungsklick.de/a/88459/inklusive-bildung-braucht-inklusive-diagnostik/

Schöler, H. (2008). Vom Unsinn der Förderdiagnostik. Vortrag auf Kongress der Deutschen Gesellschaft für Sprachheilpädagogik, Cottbus. http://www01.ph-heidelberg.de/wp/schoeler/Datein/Cottbus-2008_Vom%20Unsinn%20der%20Foerderdiagnostik_27-09-2008.pdf

Schulte, E. & Souvignier, E. (2013). Der Lese-Sportler. Die Grundschulzeitschrift. 7/2013. S. 58-61.

Schöning, A.; Fahrenwaldt, A.-M.; Rössig, A-M. G. & Hartke, B. (2013). Zur Qualität sonderpädagogischer Diagnostik in Mecklenburg-Vorpommern im Schuljahr 2008/09. http://www.sopaed.uni-rostock.de/fileadmin/Isoheilp/Diagnostik_MV/Dia-Bericht_13.6.13.pdf

Sommer-Stumpenhorst, N. (2011). Lernentwicklung und Beratung. Beckum: Graf Orthos Rechtschreibwerkstatt.

Sommer-Stumpenhorst, N. & Hötzel, M. (2007). Richtig schreiben lernen von Anfang an. Berlin: Cornelsen Scriptor.

Tent, L. & Stelzl, I. (1993). Pädagogisch-psychologische Diagnostik (Bd. 1: Theoretische und methodische Grundlagen). Göttingen: Hogrefe.

Werning, R. (2011). Gutachten: Kompetenzzentren für sonderpädagogische Förderung im Bereich der Lern- und Entwicklungsstörungen in Nordrhein-Westfalen. https://www.schulministerium.nrw.de/docs/Schulsystem/Inklusion/Rechtliches/Kompetenzzentren/index.html

**Online-Ressourcen:**
http://de.wikipedia.org/wiki/Psychologische_Diagnostik
http://de.wikipedia.org/wiki/G%C3%BCtekriterien_psychodiagnostischer_Verfahren
http://de.wikipedia.org/wiki/Normwertskala
https://de.wikipedia.org/wiki/Validit%C3%A4t_von_Intelligenzmessungen
(Alle Online-Ressourcen zuletzt abgerufen im September 2016)

## Endnoten

1. Objektivität, Reliabilität (Zuverlässigkeit), Validität (Gültigkeit) Fairness und Ökonomität sind sog. „Testgütekriterien, nach denen standardisierte psychometrische Testverfahren konstruiert werden.
2. modifiziert nach Schöler (2008)
3. Neues schulpädagogisches Wörterbuch. Weinheim, München 1993, S. 118. Zitiert nach: http://www.uni-potsdam.de/fileadmin/projects/erziehungswissenschaft/documents/studium/Textboerse/pdf-Dateien/hintz_funktionen.pdf
4. in Anlehnung an Prengel, 2013, S. 50
5. Hier ist der Bezug zu einem vordefinierten Kriterium gemeint. Dies könnte z.B. sein „Zehnerübergang sicher beherrschen". Die Beschäftigung mit dem Konzept der Bezugsnormorientierung ist bei der Interpretation diag-nostischer Daten – ob standardisiert, inklusiv oder qualitativ ein gewinnbringendes Feld. Der oben zitierte Artikel von Rheinberg (2001) sei hier empfohlen. Für eine erste Übersicht hilft auch http://de.wikipedia.org/wiki/Bezugsnorm
6. Diagnostik stellt fest, auf welcher Kompetenzstufe sich ein Schüler befindet. Durch das vorgegebene „Kompetenzraster" im jeweiligen Aufgabenfeld wird nach dieser Einordnung deutlich, welches die nächsten Kompetenzen sein können, die der Schüler erreichen kann.
7. Schulversuch PRIMUS. Näheres siehe hier: http://www.schulministerium.nrw.de/docs/Schulsystem/Versuche/Primus/index.html
8. Dies gilt für Fächer, die mehrmals pro Woche unterrichtet werden, das sind i.d.R. die Hauptfächer bzw. in der Grundschule Lesen, Schreiben, Rechnen.
9. Beispiele hierzu sind:
   - Lesen: Kieler Leseaufbau (Dummer-Smoch, L. & Hackethal, R., 2007)
   - Rechtschreiben: Lautgetreue Lese-Rechtschreibförderung (Reuter-Liehr, C.; 2008)
   - Rechnen: Rechenspiele mit Elfe und Mathis (Lenhard, W. & Lenhard A., 2010).
10. Die wichtigste Kritik am RTI-Ansatz ist der Vorwurf, RTI schaffe durch seine strikten diagnostischen Vorgehensweisen ein schnelles „Nicht-Passen" zum System und fördere durch die schon frühe Erkenntnis der Schwächen von Schülern ein ebenso frühes Aussondern. Siehe auch Schumann (2013) oder Ferri (2011).
11. Ein weiterer Aspekt ist, dass sich die Lesefähigkeit gut in Einzelschritte zergliedern lässt und vglw. einfach trainiert werden kann. Lernerfolge werden für die Schüler schnell sichtbar und können auch gut zur Motivation dienen.
12. Reziprokes Lehren ist eine Methode zur Steigerung des Leseverständnisses. Hierbei liest eine Gruppe von Schülern abschnittsweise einen Text und bespricht dessen Bedeutung. Diese „Besprechung" wird durch den Lehrer strukturiert (Zusammenfassung, Fragen, Klärung, Vorhersagen). Nach einer Einführungsphase zieht sich der Lehrer zurück. Die Lehrerrolle wird dann wechselnd von den Schülern übernommen (vgl. Schnotz, W. (2006)
13. Sommer-Stumpenhorst, 2011, S. 8ff und S. 20
14. Die Prozentrangskala geht von 0 bis 100. Der Durchschnitt liegt im Bereich von PR=18 bis PR=82.
15. Schöning, A.; Fahrenwaldt, A.-M.; Rössig, A-M. G. & Hartke, B. (2013, S. 188.) Die Studie wurde mit Gutachten im Bundesland Mecklenburg-Vorpommern in den Jahren 2008/2009 durchgeführt. Natürlich kann die Situation an anderer Stelle zu anderen Zeitpunkten von dem hier Berichteten unterschiedlich sein. Dennoch sind die zitierten Aspekte hilfreich und an der ein- oder anderen Stelle übertragbar.

# ZUR PERSON

## Schulpsychologe
## Dr. Sascha Borchers

Dipl.-Psych. Dr. Sascha Borchers arbeitet seit 2011 in der Regionalen Schulberatungsstelle des Kreises Borken und ist daneben als Fachbeauftragter für Schulpsychologie in der Bezirksregierung Münster tätig. Er hat zuvor an der Universität Gießen unter anderem in der Lehrerausbildung gearbeitet und zum Thema „E-Learning“ promoviert. Seine Arbeitsschwerpunkte sind neben Diagnostik und individueller Förderung die Themen Lese-Rechtschreib-Schwierigkeiten, Umgang mit besonderer Begabung in der Schule und Motivation. Dr. Borchers ist Geschäftsführer des Lenkungskreises Krisenprävention und –intervention an Schulen des Kreises Borken und beschäftigt sich in diesem Kontext mit schulpsychologischer Krisenintervention und der Ausbildung schulischer Teams für Gewaltprävention und Krisenintervention.

# NOTIZEN

REGIONALE
Schulberatungsstelle
DES KREISES BORKEN

Burloer Str. 93 | 46325 Borken
Tel. 02861 / 82-2527 | Fax 02861 / 822712527
schulberatung@kreis-borken.de | www.rsb-borken.de