

Das Problem der Selektion im Lichte von Forschungsbefunden

Autor(en): **Hedinger, Urs K.**

Objekttyp: **Article**

Zeitschrift: **Schweizer Schule**

Band (Jahr): **70 (1983)**

Heft 7: **Notengebung und Selektion : ein notwendiges Übel?**

PDF erstellt am: **26.09.2024**

Persistenter Link: <https://doi.org/10.5169/seals-532104>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern. Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden. Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Das Problem der Selektion im Lichte von Forschungsbefunden

Urs K. Hedinger

Einschlägige Forschungen

Kritik an der Auslese, wie sie die Schule betreibt, hat es zweifellos immer schon gegeben. In jüngerer Zeit ist diese Kritik aber grundsätzlicher geworden. Dies aus verschiedenen Gründen, unter anderem auch deshalb, weil die Kritik zunehmend aus Forschungsergebnissen Gesichtspunkte und Argumente gewinnen konnte. Im folgenden versuche ich, einige Forschungsbefunde teils zusammenzufassen, teils exemplarisch aufzuzeigen, die im Zusammenhang mit der Selektionsproblematik wichtig sind und Elemente zu einer Beurteilung der Selektionspraxis beitragen.

Insgesamt sind es sehr viele Forschungen aus verschiedenen wissenschaftlichen Disziplinen (Psychologie, Erziehungswissenschaft, Bildungsforschung, Soziologie u.a.m.), die zum Thema Selektion von Bedeutung sind, auch wenn sie sich nicht ausdrücklich darauf beziehen. Es gehören dazu unter anderem folgende Arten von Forschungen:

- Über verschiedene Arten von Fähigkeiten oder Begabungen, über ihre Entwicklung und die Einflüsse auf deren Entwicklung (z. B. Studien zur Anlage – Umwelt – Problematik der Begabung)
- Über Schülermerkmale, die Bedingungen für schulisches Lernen und für Schulerfolg bzw. Schulversagen darstellen (dazu gehören neben der Intelligenz und anderen Fähigkeiten auch Motivationen, Einstellungen und weitere Persönlichkeitsmerkmale)
- Über den Verlauf von Schullaufbahnen und über individuelle und soziale Faktoren, die ihn bestimmen (Bedeutung von Fähigkeiten, Einfluss des Elternhauses, Chancenungleichheiten und ihre Gründe)
- Über die Schülerbeurteilung, über die Möglichkeiten und Grenzen verschiedener Methoden der Beurteilung, über die Funktionen und Auswirkungen der Beurteilung
- Über den Selektionsvorgang im engeren Sinne, die verwendeten Verfahren und In-

strumente, die Möglichkeiten ihrer Verbesserung, sowie über die Ergebnisse und Auswirkungen der Selektion

- Über die Bedeutung der Schulstrukturen für die Selektion, deren Verlauf und Ergebnisse, die Bedeutung der Strukturen für die Entwicklung von Fähigkeiten beim Schüler und für die Schullaufbahn.

Hier kann nur auf einige wenige Forschungsergebnisse hingewiesen werden, die als gut gesichert gelten dürfen, weil sie sich in konsistenter Weise aus einer Vielzahl von Untersuchungen ergeben haben, und die mir für die Beurteilung der Selektionspraxis besonders wichtig erscheinen.¹

Schülerbeurteilung als Grundlage der Auslese

Selektionsentscheide oder, allgemeiner gesprochen, Entscheide über Zuweisungen von Schülern zu unterschiedlichen Unterrichtsangeboten und Bildungswegen stützen sich immer in irgendeiner Form auf eine Beurteilung früherer und aktueller Leistungen, allenfalls auf die Beurteilung weiterer leistungsbezogener Merkmale wie Motivationen, Arbeitshaltung etc. Die Qualität oder Richtigkeit von Ausleseentscheidungen ist somit unter anderem von der Qualität entsprechender Schülerbeurteilungen abhängig. Die pädagogische Diagnostik als Spezialzweig der Erziehungswissenschaft hat sich intensiv mit der Frage beschäftigt, welchen Kriterien Schülerbeurteilungen, insbesondere Leistungsbeurteilungen in Form von Noten, zu genügen hätten und in welchem Masse sie dies tatsächlich auch tun. In Anlehnung an die psychologische Testtheorie werden folgende Gütekriterien postuliert:

- *Objektivität*. Urteile und Bewertungen sind dann objektiv, wenn die gleiche Leistung (das gleiche Arbeits- oder Lernergebnis) von verschiedenen Beurteilern gleich beurteilt wird. Objektivität heisst Unabhängigkeit von der Person des Beurteilers.

- *Zuverlässigkeit*. Urteile sind zuverlässig, wenn sie bei wiederholter Beurteilung immer wieder gleich ausfallen, also nicht augenblicks- und nicht situationsabhängig sind.
- *Gültigkeit*. Urteile sind gültig (valide), wenn sie tatsächlich das (und nur das) erfassen, was sie zu erfassen vorgeben. Dies setzt eine eindeutige inhaltliche Bestimmung dessen voraus, was beurteilt wird. (Die Beurteilung von Aufsätzen kann als besonders einleuchtendes Beispiel dafür erwähnt werden, dass Urteile über die gleiche Arbeit oft unterschiedlich ausfallen, weil Unterschiedliches beurteilt wird.)

Empirische Untersuchungen haben wiederholt nachgewiesen, dass Leistungsbeurteilungen aus verschiedenen Gründen lehrerabhängig sind. Werden Schülerarbeiten von verschiedenen Lehrern beurteilt, so streuen die erteilten Noten über mehrere Werte der Notenskala. Dies gilt nicht nur für Aufsätze, von denen gerne zugegeben wird, dass es schwierig ist, sie objektiv zu bewerten, sondern auch für Mathematikarbeiten (vgl. Weiss, in Ingenkamp 1971). Dieser Befund wird unter anderem dadurch erklärt, dass den Lehrern bei der Beurteilung von Schülerarbeiten jeweils ihre eigene Klasse als wichtigster Bezugsrahmen dient, was bedeutet, dass sie unterschiedliche Massstäbe benützen. Deshalb sind Noten verschiedener Lehrer nur bedingt vergleichbar.

Aber auch vom gleichen Lehrer werden die gleichen Schülerarbeiten nicht immer gleich beurteilt. So fand zum Beispiel Eells (in: Ingenkamp 1971), dass bei wiederholter Beurteilung von aufsatzähnlichen Geschichts- und Geographiearbeiten im Abstand von 11 Wochen durch den gleichen Lehrer nur eine mässige Übereinstimmung zwischen erster und zweiter Beurteilung erzielt wurde (Korrelationen im Bereich von 0.25 bis 0.51). Ähnliches gilt auch hier ebenso für Mathematikarbeiten. Dicker (in: Zielinski 1974) fand, dass bei Mathematikarbeiten, die im Abstand von drei Monaten zweimal beurteilt wurden, die Lehrer nur bei einem Drittel der Arbeiten beim zweitenmal die gleiche Note setzten wie beim erstenmal.

Andere Untersuchungen zeigen, dass Leistungsurteile noch durch andere Faktoren beeinflusst werden als die vom Schüler effektiv erbrachte Leistung. Hadley (in: Ingenkamp 1971) hat Zusammenhänge aufgezeigt, die zwischen den mit Tests erfassten Leistungen

von Schülern (die als relativ objektiv gelten dürfen), ihren Noten und ihrer Beliebtheit beim Lehrer bestehen. Diejenigen Schüler, die beim Lehrer als die beliebteren gelten, erhielten bessere Noten als aufgrund der Leistungsergebnisse in Tests zu erwarten gewesen wäre; aber nur 16 % dieser Schüler bekamen schlechtere Zensuren als erwartet. Von den bei den Lehrern als weniger beliebt geltenden Schülern erhielten 50 % schlechtere Noten als der Leistungstest erwarten liess, aber nur 19 % dieser Schüler wurden, gemessen am Testergebnis, überbewertet. Darüber hinaus ergaben sich für die Mädchen merklich bessere Zensuren als für die Knaben. Der Grund dürfte darin liegen, dass die Mädchen in der Schule angepasster und fleissiger sind, bei den Lehrern deswegen als beliebter gelten und dafür mit besseren Noten belohnt werden.

Noch andere Arten von «Vorinformationen», die der Lehrer über den Schüler hat, sowie Erwartungen und Einstellungen des Lehrers spielen bei der Beurteilung eine Rolle. Als Halo-Effekt wird die Tatsache bezeichnet, dass der Lehrer dazu neigt, verschiedene Merkmale eines Schülers nicht unabhängig voneinander zu beurteilen, dass vielmehr das eine Urteil das andere beeinflusst. Aufgrund der festgestellten bisherigen Leistungen entwickelt der Lehrer auch Erwartungen bezüglich der künftigen Leistungen eines Schülers. Diese können, zum Beispiel über selektive Wahrnehmung, die Beurteilung späterer Leistungen beeinflussen. Ebenfalls durch selektive Wahrnehmung kann teilweise erklärt werden, dass der Lehrer dazu neigt, bei einem einmal gefällten Urteil über den Schüler zu bleiben und es auch unter veränderten Umständen (z. B. bei verbesserten oder verschlechterten Leistungen) nicht zu ändern (Ullrich und Woebecke 1981). Unter der Bezeichnung «Pygmalion-Effekt» wird in der pädagogischen Literatur das Phänomen geschildert und diskutiert, dass Leistungserwartungen des Lehrers nicht nur seine Urteile über Schüler «verzerren», sondern sich auch auf die effektiven Leistungen der Schüler auswirken können (Rosenthal und Jacobson 1971).

Die meisten der hier erwähnten Untersuchungsbefunde können nicht als Einzelfälle und zufällige Ergebnisse betrachtet werden. Sie werden durch die Ergebnisse vieler Studien gestützt und bestätigt und müssen des-

halb, zumindest in der Tendenz, als typisch gelten. Alle diese Untersuchungen legen den Schluss nahe, dass Leistungsbeurteilungen, vor allem in Form von Noten, nur bedingt den Anforderungen der Objektivität, der Zuverlässigkeit und der Gültigkeit genügen. Sie sind aus verschiedenen Gründen mit mehr oder weniger grossen Fehlern behaftet. Da solche Beurteilungen im wesentlichen die Grundlage für Ausleseentscheide darstellen, sind auch letztere mit entsprechenden Unsicherheiten und Fehlern belastet. Die Qualität der Selektionsentscheide kann nicht besser sein als diejenige der Schülerbeurteilungen, auf die sie sich stützen.

Unsichere Entwicklungsprognosen

Bei der Selektion kommen aber zu den Ungenauigkeiten in der Beurteilung früherer und aktueller Leistungen weitere Schwierigkeiten hinzu. Es stellt sich hier nämlich nicht allein ein Problem der *Diagnose* des aktuellen Zustandes, sondern eines der *Prognose* von künftigen Entwicklungen bzw. Entwicklungsmöglichkeiten. Für die Zuteilung der Schüler zu unterschiedlichen Bildungswegen sollten nicht allein die aktuellen Leistungen ausschlaggebend sein, sondern auch die zu erwartenden künftigen Entwicklungen und Lern- und Leistungsmöglichkeiten des Schülers. Die Berechtigung einer Aufteilung in Real- und Sekundarschüler zum Beispiel, mit all den weitreichenden Konsequenzen, die damit verbunden sind, wäre offenbar in Zweifel zu ziehen, wenn sich ein oder zwei Jahre nach der Aufteilung oder auch noch später erweisen sollte, dass die vorgenommene Gruppierung nicht mehr den effektiven Leistungen und Lernfähigkeiten der Schüler entspricht.

Genauere Angaben darüber, wie gut sich zum Zeitpunkt der Selektion die späteren Leistungen der Schüler und ihr Erfolg in den verschiedenen Schultypen voraussagen lassen, kann man vor allem jenen Untersuchungen entnehmen, die die Bewährung konkreter Ausleseverfahren systematisch überprüft haben, was oft im Zusammenhang mit Versuchen mit neuen Ausleseverfahren und -instrumenten geschehen ist. Die Bewährung wird in der Weise festgestellt, dass die Übereinstimmung zwischen der früheren Beurteilung (zum Zeitpunkt der Selektion) und den späteren Schulleistungen

der Schüler empirisch bestimmt wird. Das Mass der Übereinstimmung oder die Enge des Zusammenhangs zwischen Voraussage und späterer Bewährung wird mit Hilfe des statistischen Kennwerts der Korrelation ausgedrückt. Eine Korrelation von 1.0 würde dabei eine vollständige Übereinstimmung bedeuten, eine solche von 0.0, dass zwischen dem früheren Urteil und der späteren Bewährung überhaupt kein Zusammenhang besteht. Je höher die Korrelation, desto besser kann mit dem früheren Urteil die spätere Bewährung vorausgesagt werden. Man spricht deshalb von der *prognostischen Gültigkeit* der früheren Urteile. In den meisten Studien wird sowohl die Bewährung einzelner Formen der Schülerbeurteilung (wie: durchschnittliche Erfahrungsnote, durchschnittliche Prüfungsnote, Eignungsurteil des abgebenden Lehrers etc.) als auch von Kombinationen solcher Urteile, wie sie ja in der Praxis in der Regel für Ausleseentscheide massgebend sind, überprüft. Als typisches Beispiel für entsprechende Befunde führe ich hier diejenigen aus einer Studie von HUERSCH und PREISIG in der Stadt Bern an (vgl. Preisig 1979). Eine Vielzahl von anderen Studien hat in der Tendenz übereinstimmende Ergebnisse gezeitigt. (Siehe Tabelle 1, S. 306).

In der Studie von HUERSCH und PREISIG wurden vier Formen von Beurteilungen der Schüler einbezogen:

- Ein Notendurchschnitt aus dem letzten Schuljahr der abgehenden Primarschule (PRIM)
- Das Ergebnis der Aufnahmeprüfung für die Sekundarschule in Form des erreichten Notendurchschnitts (AUF)
- Das in einem Intelligenztest (Berner Übertrittstest) erzielte Ergebnis (INT)
- Ein Schülerbeurteilungsbogen (BEB). Mit Hilfe dieses speziell für die Zwecke der Übertrittsauslese entwickelten Instruments beurteilten die Primarlehrer ihre Schüler bezüglich folgender Merkmale: Sprachverhalten, rechnerisches Denken, Initiative, Aufmerksamkeit, geistige Beweglichkeit, Selbstständigkeit, Ordentlichkeit.

Die Ergebnisse, wie sie die Tabelle ausweist, decken einige wichtige Sachverhalte auf.

- Die einzelnen Beurteilungsarten weisen alle ähnliche Korrelationen mit dem Schulerfolg nach einem Jahr Sekundarschule auf, wobei diejenigen für die Primarschulnoten und für

Tabelle 1

Multiple Korrelationen zwischen dem Mittel der Sekundarschulnoten nach 1/4 Jahr bzw. nach einem Jahr und den Prädiktoren							
Prädiktoren:	Beurteilungsbogen Primarschulnoten (4. Kl.)	Intelligenztest	Aufnahmeprüfung	BEB	PRI	INT	AUF
Merkmale	R nach 1/4 Jahr Sekundarschule	R ² (%)	R nach 1 Jahr Sekundarschule	R ² (%)			
1. BEB	.65	42	.42	18			
2. PRI	.68	46	.50	25			
3. INT	.58	34	.43	18			
4. AUF	.52	27	.52	27			
5. INT/BEB	.78	61	.59				
6. PRI/BEB	.75	56	.56				
7. AUF/BEB	.75	56	.67				
8. PRI/INT	.75	56	.59				
9. INT/AUF	.64	41	.60				
10. PRI/AUF	.72	52	.61				
11. BEB/INT/PRI	.81	66	.63				
12. BEB/AUF/INT	.82	67	.71				
13. BEB/AUF/PRI	.79	62	.69				
14. AUF/INT/PRI	.77	59	.68				
15. BEB/PRI/AUF/INT	.83	69	.73				

R: Korrelation zwischen Prädiktor und Kriterium

R²: Gibt an, welcher Anteil der Kriteriumsvarianz durch die Prädiktoren erklärt wird.

Quelle: Preisig 1979

- die Aufnahmeprüfung etwas höher liegen. Aber auch diese Korrelationen sind nur mäßig hoch. Die einzelnen Beurteilungen (Prädiktoren) erklären lediglich rund 25 % der Unterschiede (Varianz) in den Leistungen ein Jahr später (ausgedrückt durch das R²).
- Kombiniert man zwei Beurteilungen, wird allgemein eine Verbesserung der Prognose erreicht. Werden die Erfahrungsnoten und das Ergebnis der Aufnahmeprüfung zusammen berücksichtigt, so beträgt die Korrelation jetzt 0.61 statt 0.50 bzw. 0.52. Es werden somit 10 % mehr der Leistungsunterschiede nach einem Jahr erklärt (oder vorausgesagt). Die Prognosemöglichkeit wird nochmals verbessert, wenn drei oder gar alle vier verfügbaren Urteile kombiniert werden.
 - Die Kombination mehrerer Urteile erlaubt somit etwas bessere Prognosen als die Berücksichtigung nur eines einzigen Urteils. Die damit erreichte Verbesserung bleibt jedoch beschränkt. Ein wesentlicher Grund dafür ist, dass die einzelnen Urteile weitgehend austauschbar sind, weil sie im wesentlichen das gleiche erfassen. Wurden zum Beispiel bereits zwei Urteile berücksichtigt, so bringt der Einbezug eines weiteren Urteils nur noch eine geringe Verbesserung der Prognose – vor allem auch gemessen am zusätzlichen Aufwand –, weil es wenig neue Information einbringt.
 - In Übereinstimmung mit andern Studien zeigen die Ergebnisse, dass auch ein sorgfältig entwickeltes und durchgeführtes Auslese-

verfahren, das mehrere Beurteilungen der Schüler miteinbezieht, nur eine mässige prognostische Gültigkeit erreicht. Eine Korrelation von 0.70, wie sie in den meisten Fällen etwa festgestellt wird, wenn die Bewährung nach einem Jahr überprüft wird, bedeutet, dass durch die Beurteilung zum Zeitpunkt der Auslese nur die Hälfte der Unterschiede in den Leistungen ein Jahr später erklärt werden kann.

- Wichtig ist ferner die Feststellung, dass die Prognose deutlich schlechter ausfällt, je weiter vom Selektionszeitpunkt entfernt derjenige Zeitpunkt liegt, in welchem die Bewährung überprüft wird. Das haben vor allem auch andere Untersuchungen gezeigt, die längere Zeiträume einbeziehen. Das heisst, dass zu einem gegebenen Zeitpunkt zwar eine mässig gute Voraussage über die Leistungen der Schüler ein Jahr später möglich ist, dass Voraussagen über zwei, drei und mehr Jahre hinweg dagegen sehr ungenau sind. Je früher also im Verlauf der Volksschule die Selektion getroffen wird, desto weniger ist es zum betreffenden Zeitpunkt möglich, das relative Leistungsniveau der Schüler am Ende der obligatorischen Schulpflicht vorausszusehen.
- Es ergibt sich eine weitere bedeutsame Feststellung: Auch wenn in das Ausleseverfahren Instrumente einbezogen werden, die auf wissenschaftlicher Grundlage (Testtheorie) entwickelt sind (Intelligenz-, Schulleistungstests und andere Beurteilungsverfahren), so wird keine wesentliche Verbesserung der Ausleseentscheide erreicht. Die Hoffnungen auf eine Verbesserung der Auslese durch Tests haben sich als nicht begründet erwiesen.

Wie viele Fehlentscheide gibt es?

Dass bei der Selektion Fehlentscheide vorkommen, weiss man aufgrund vieler Einzelerfahrungen. Dass ihre Zahl erheblich sein muss, vermutet man aufgrund von Befunden der Art, wie sie oben aufgezeigt wurden. Empirisch gesicherte Daten über die Häufigkeit von Fehlzuweisungen sind zwar nicht direkt verfügbar, aber Schätzungen sind möglich.

Aufgrund der gefundenen Masse des Zusammenhangs zwischen früherer Beurteilung und

späterer Bewährung kann man bestimmen, welche Anteile von richtigen und falschen Zuordnungen zu erwarten sind, wenn man die Schüler aufgrund der Beurteilung in einem gegebenen Zeitpunkt in zwei gleich grosse Gruppen einteilt, nämlich «bessere» und «schwächere». Die Anteile sind von der Enge des Zusammenhangs (prognostische Gültigkeit) abhängig. Bei einer entsprechenden Korrelation von 0.70 – was, wie wir gesehen haben, offenbar etwa dem erreichbaren Optimum entspricht – ergibt sich ein Anteil von 25 % Fehlzuordnungen. Obwohl die Bewährung von Beurteilungen und Ausleseentscheidungen meist nur noch an ausgelesenen Gruppen überprüft wird, muss angenommen werden, dass an jener Stelle, wo bei der Selektion effektiv der Schnitt gezogen wird, ein ähnliches Ausmass an Ungenauigkeit und damit ähnlich viele Fehlzuweisungen vorkommen.

Diese Zahl mag zunächst überraschen, da sich in der Praxis offenbar lange nicht so viele Fälle als tatsächliche Fehlzuweisungen herausstellen. Diese Divergenz lässt sich auf verschiedene Art erklären:

- Man spricht in diesem Zusammenhang von der *Haltekraft der Schulen* (Fend 1982). Damit ist die Tatsache gemeint, dass Schüler, wenn sie einmal in eine bestimmte Schule eingetreten sind, mit grosser Wahrscheinlichkeit auch dort verbleiben, und zwar relativ unabhängig von ihren tatsächlichen Eignungs- und Leistungsvoraussetzungen.
- Selektionsentscheide wirken wahrscheinlich oft im Sinne von «Sich-selbst-erfüllenden-Prophezeiungen». Die Einstufung in einen wenig anspruchsvollen Schulzweig wird vom Schüler als Urteil über seine schulischen Fähigkeiten in sein Selbstbild übernommen; er orientiert sich an den bescheidenen Erwartungen, die an Schüler dieses Schultyps gerichtet werden, und erbringt auch tatsächlich entsprechend bescheidene Leistungen. Umgekehrt kann die Einstufung auf einem anspruchsvollen Schulniveau zusätzliche Kräfte mobilisieren.
- Schliesslich ist darauf hinzuweisen, dass unsere Schulstrukturen nur beschränkt Korrekturen an einmal gefällten Selektionsentscheiden erlauben. Die bei der Selektion getrennten Schülergruppen entwickeln sich, da sie unterschiedlich unterrichtet werden,

auseinander, so dass Wechsel bald einmal schwierig werden, vor allem «Aufstiege» von weniger anspruchsvollen in anspruchsvollere Schulzweige.

Im vorliegenden Zusammenhang sind auch Befunde über das Ausmass der *Mobilität in verschiedenen Schulstrukturformen* von Interesse. Im vertikal gegliederten Schulsystem, wie es bei uns die meisten Kantone kennen, ist die Mobilität, d. h. die Häufigkeit von Schullaufbahnwechseln in der Volksschuloberstufe, beschränkt, wobei die Abstiege meist überwiegen. Sie ist merklich grösser in integrierten, durchlässigeren Schulformen. Dort nehmen rund 30 % der Schüler im Verlauf der Volksschuloberstufe eine Entwicklung, die einem Schultypenwechsel im vertikal gegliederten Schulsystem entsprechen würde und die zu einem anderen Abschlussniveau am Ende der obligatorischen Schulpflicht führt, als aufgrund der Beurteilung dieser Schüler am Ende der Primarschule zu erwarten gewesen wäre (Fend 1982). Dabei sind Aufstiege mindestens so häufig wie Abstiege. Man kann daraus schliessen, dass sich bei manchen Schülern nach dem 4. Schuljahr und auch noch nach dem 6. Schuljahr Entwicklungen ergeben, die wesentliche Änderungen in ihren Lern- und Leistungsmöglichkeiten bedeuten. Integrierte und durchlässige Schulstrukturen erleichtern das Umsetzen solcher Veränderungen in Schullaufbahnwechsel. Das vertikale, wenig durchlässige Schulsystem dagegen erschwert entsprechende Wechsel und verdeckt damit die Häufigkeit solcher Entwicklungen und Veränderungen bei Schülern.

Die Zusammensetzung ausgelesener Schülergruppen

Ziel der Selektion ist es unter anderem, Gruppen von Schülern mit ähnlichen Lern- und Leistungsvoraussetzungen zu bilden, die dann entsprechend ihren je besonderen Voraussetzungen unterschiedlich unterrichtet werden (Unterschiede im Lerntempo, im Anspruchsniveau, im Stoffprogramm). Die erwähnten Befunde über die Ungenauigkeit von Schülerbeurteilungen und über die begrenzte prognostische Gültigkeit von Urteilen und Ausleseentscheidungen lassen vermuten, dass es nur unzureichend gelingt, homogene Schülergruppen zu bilden. Dass dies empirisch belegt ist,

soll am Beispiel der Studie von HALBHERR (1980, 1981) gezeigt werden.

Halbherr untersuchte in zwei Gemeinden des Kantons Baselland die Leistungen (Intelligenz und Schulleistungen) von rund 400 Schülern im 6. Schuljahr, drei Monate nach ihrer Einweisung in die Realschule, die Sekundarschule oder das Progymnasium. Abbildung 1 veranschaulicht die Ergebnisse, die in der Tendenz zwar zu erwarten sind, in ihrer Deutlichkeit aber überraschen. Kurz nach der am Ende der 5. Klasse erfolgten Selektion unterscheiden sich die Schüler der drei Schultypen in ihren Leistungen nur bedingt voneinander. Die mittleren Testwerte der drei Schülergruppen sind zwar mehr oder weniger deutlich voneinander verschieden, aber die Verteilungen der Leistungswerte überschneiden sich sehr stark. Je nach Leistungsart kann dies folgendes bedeuten: Ein schlechter Sekundarschüler ist nicht leistungsfähiger als ein durchschnittlicher Realschüler, und ein guter Sekundarschüler ist ebenso leistungsfähig wie ein durchschnittlicher Progymnasiast. Der Grossteil aller Schüler liegt also in einem weiten mittleren Leistungsbereich, der an sämtlichen drei Schultypen repräsentiert ist. Die Selektion hat also nicht zu einer klaren Trennung von Gruppen von Schülern mit deutlich unterschiedlichen Fähigkeiten und Leistungen geführt.

Das heisst, dass die Selektion nur sehr bedingt «begabungsgerecht» erfolgt ist. Es ist nicht so, dass die Zuordnung zu einem der drei Schultypen sich in eindeutiger Weise aus den Fähigkeiten und Leistungen des Schülers ergeben hat. Die Studie enthält auch Hinweise darauf, welche anderen Faktoren neben den Fähigkeiten die Verteilung der Schüler über die Schultypen beeinflussen.

Es sind zum einen *regionale Faktoren*. Es gibt beträchtliche regionale Unterschiede in der prozentualen Verteilung der Schüler über die drei Oberstufentypen. In der stadtnahen Gemeinde besuchen mehr Schüler die höheren Schultypen als in der ländlichen Gemeinde. In den einzelnen Leistungsarten unterscheiden sich die Schüler des gleichen Schultyps zwischen den Gemeinden deutlich voneinander. Schüler des gleichen Schultyps sind an verschiedenen Orten unterschiedlich leistungsfähig; ein «durchschnittlicher Sekundarschüler» ist je nach Ort durch ein anderes Leistungsniveau charakterisiert.

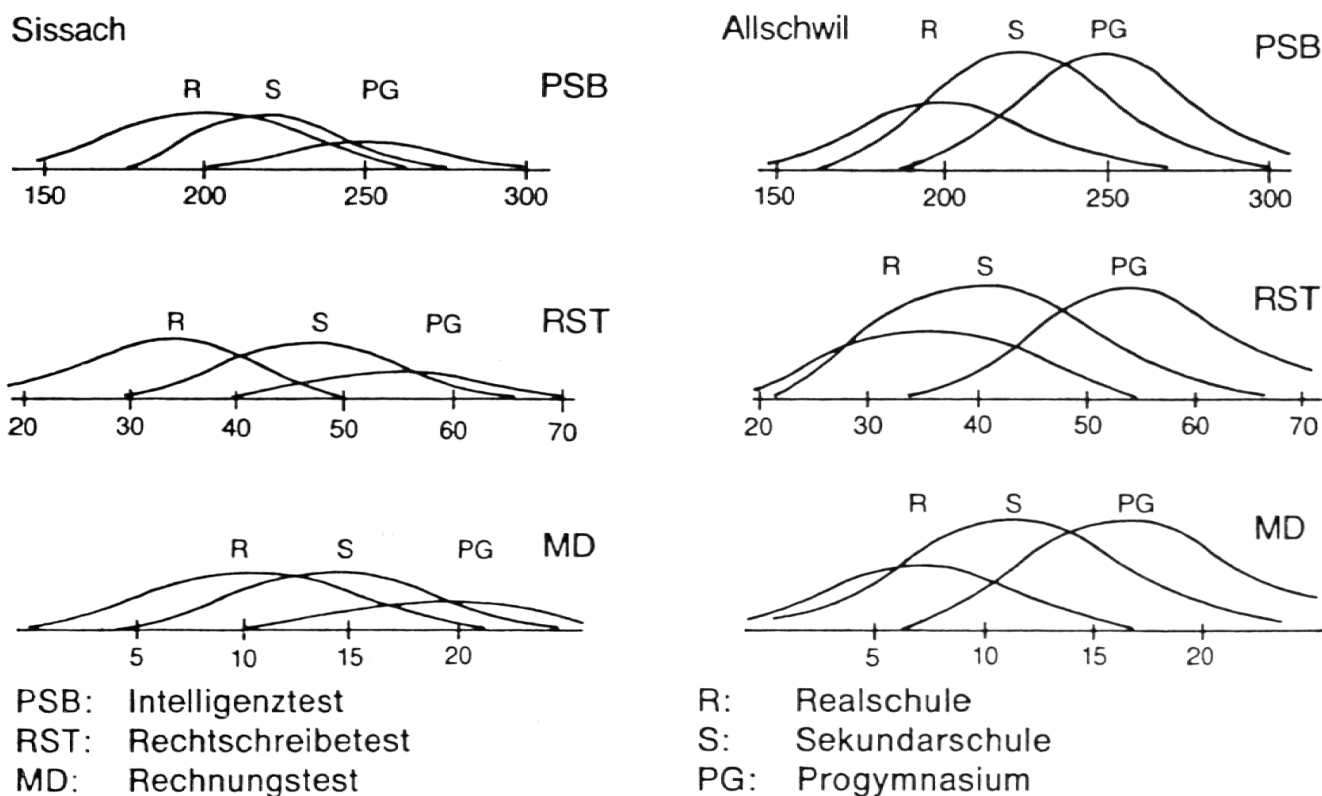


Abbildung 1

Leistungsspektrum der Real-, Sekundar- und Progymnasialschüler in Sissach und Allschwil (6. Schuljahr)

Quelle: Halbherr 1981

Von erheblichem Einfluss erweist sich ferner auch in dieser Studie das *soziale Herkunft* der Schüler. Es besteht ein deutlicher Zusammenhang zwischen dem sozioökonomischen Herkunft (Ausbildung und Beruf der Eltern) und dem von den Kindern besuchten Schultyp. Kinder aus den höheren Gesellschaftsschichten sind im obersten Schulniveau massiv «übervertreten» (gemessen an ihrem Anteil am gesamten Schülerjahrgang), im untersten Niveau «untervertreten». Das Umgekehrte gilt für Kinder aus den unteren Schichten. Diese ungleiche Verteilung kann zum Teil durch Unterschiede in den Fähigkeiten erklärt werden. Innerhalb der Gesamtgruppe der erfassten Schüler zeigt sich nämlich ein gewisser Zusammenhang zwischen sozialer Herkunft und Intelligenz: Je höher die Herkunftsschicht, desto höher im Durchschnitt die Intelligenz. Die soziale Herkunft bestimmt aber noch zusätzlich zur und unabhängig von der Intelligenz die Schullaufbahn der Kinder. Innerhalb der einzelnen Schultypen sind die Zusammenhänge zwischen Herkunft und Intelligenz nämlich gerade

umgekehrt. In den höheren Schulzweigen zeigen die Kinder aus unteren Schichten merklich höhere Intelligenzleistungen, diejenigen aus den höheren Schichten niedrigere Leistungen. Das bedeutet, dass in den höheren Schichten das vorhandene Fähigkeitspotential in stärkerem Masse im Hinblick auf Schulerfolg und den Besuch von anspruchsvollen Schulen «mobilisiert» wird. In den unteren Schichten dagegen wird im Gegenteil «überselektioniert», d. h. nur wenige Kinder und nur solche mit überdurchschnittlicher Intelligenz gelangen in die höheren Schulniveaus. Hier bestätigt sich einmal mehr die Tatsache, dass Kinder aus verschiedenen Herkunftsschichten auch bei gleicher Begabung ungleiche Bildungschancen haben.

Überlegungen im Anschluss an die Befunde

Die Forschungsbefunde weisen darauf hin, dass die Selektion, so wie sie bei uns heute üblicherweise praktiziert wird, manche Zielvorstellungen, denen sie explizit oder implizit zu

dienen vorgibt, nicht oder nur unzureichend verwirklicht. Es gelingt offenbar nur unvollkommen, «begabungsgerecht» auszulesen, d. h. eine Gruppierung der Schüler wirklich nach Fähigkeits- und leistungsbezogenen Kriterien vorzunehmen. Es gelingt auch nur unzureichend, Gruppen von Schülern zu bilden, die bezüglich Lern- und Leistungsvoraussetzungen relativ homogen sind. Damit wird es aber schwierig, die Aufteilung der Schüler auf verschiedene Ausbildungswege mit den weitreichenden Konsequenzen, die dies für die weitere Laufbahn und die Lebenschancen der Schüler hat, zu legitimieren, wenn nicht überzeugend belegt werden kann, dass sie durch eindeutige Fähigkeits- und Leistungsunterschiede gerechtfertigt ist.

Eine der Hauptschwierigkeiten im Zusammenhang mit der Selektion und eine der Ursachen für die geschilderten Sachverhalte scheint darin zu liegen, dass es nicht möglich ist, zum Zeitpunkt der Selektion einigermaßen verlässliche Aussagen über die weitere Entwicklung der Schüler zu machen. Die Gründe für den insgesamt nur mässigen Zusammenhang zwischen früheren und späteren Schulleistungen sind weniger in Mängeln der Diagnoseinstrumente und der zu einem gegebenen Zeitpunkt vorgenommenen Beurteilungen zu suchen, sondern vielmehr im Fehlen von Möglichkeiten, künftige Veränderungen vorzusehen. Bedeutsame Voraussetzungen beim Schüler für das Lernen und den Schulerfolg, wie zum Beispiel Fähigkeiten und Motivationen, verändern sich unter dem Einfluss verschiedenster Bedingungen. Je länger die betrachteten Zeiträume, desto wahrscheinlicher sind solche Veränderungen. Aber auch die Anforderungen der Schule bleiben nicht konstant. Sie ändern sich als Folge der Einführung neuer Fächer (z. B. einer Fremdsprache), innerhalb der einzelnen Fächer als Folge von Verlagerungen in den Zielen und Inhalten des Unterrichts, als Folge von Lehrerwechseln u. a. m. Da es nicht möglich ist, künftige Entwicklungen und Veränderungen systematisch zu erfassen und bei der Selektionsentscheidung mitzubetrachten, ist man darauf angewiesen, im wesentlichen vom aktuellen Leistungs- und Fähigkeitsniveau direkt auf das künftige zu schliessen: es wird als konstant angenommen. Dabei wird offensichtlich das Ausmass an möglichen und tatsächlichen Veränderungen unterschätzt.

Wenn es zutrifft, dass die Hauptschwierigkeit der Selektion mit der fehlenden Möglichkeit zusammenhängt, Veränderungen in den Fähigkeiten und Leistungen vorzusagen, darf man sich von Bemühungen um neue Beurteilungsverfahren und -instrumente nur beschränkte Verbesserungen erhoffen. Diese beziehen sich ja auch nur auf die Erfassung früherer und aktueller Leistungen. Der Kern des Problems der Ausleseentscheidungen wird davon nicht berührt. Es wird dann auch deutlich, in welcher Weise sich die Art der Selektion und die damit verbundenen Probleme aus den Schulstrukturen ergeben. Ein Schulsystem, das die Schüler in einem bestimmten Zeitpunkt auf verschiedene, voneinander getrennte Schultypen aufteilt, erfordert eine *punktueller* Selektion, die sich auf *langfristige Entwicklungsprognosen* stützen müsste, da es sich um *globale* und *überdauernde* Zuweisungen zu unterschiedlichen Ausbildungstypen und schulischen Anforderungsniveaus handelt und spätere Korrekturen in den Zuordnungen nur noch beschränkt möglich sind. Es wurde gezeigt, dass dies kaum zu leisten ist. Es bieten sich grundsätzlich zwei Richtungen an, in die sich Schulstrukturen und Formen der Unterrichtsdifferenzierung entwickeln können, so dass die Problematik der Selektion verändert und verringert wird. Es wäre dies einerseits die Schaffung von differenzierteren und durchlässigeren Strukturen, die wiederholte Zuweisungsentscheidungen zulassen, die jeweils nur für einen Teil des Unterrichts und für beschränkte Zeit gelten: dies ist zum Beispiel bei Formen der *fachbezogenen Leistungsdifferenzierung* der Fall. Die andere Möglichkeit bestünde darin, auf äussere Differenzierung, also auf die Aufteilung von Schülergruppen, zu verzichten zugunsten von *innerer Differenzierung*, d. h. von Individualisierung innerhalb von heterogen zusammengesetzten, überdauernden Lerngruppen. Schulsystemvergleichende Untersuchungen haben gezeigt, dass sich bei anderen Schulstrukturen wesentlich andere Formen, andere Probleme und andere Auswirkungen der Selektion ergeben (FEND 1982, HEDINGER 1983).

Bei der Selektion handelt es sich also nicht um ein technisches Problem, das auf der Ebene der Verfahren und Instrumente zu lösen wäre. Es ist unlösbar mit den Strukturen der Schule, aber auch mit den grundlegenden Zielvorstel-

lungen und mit den Funktionen der Schule für Individuum und Gesellschaft verknüpft. Das Schulsystem erfüllt zweifellos eine wichtige gesellschaftliche Funktion, wenn es die Heranwachsenden selektioniert, d. h. in unterschiedliche Bildungswege einweist und dort in unterschiedlicher Weise für verschiedenartige Berufs- und andere Lebensbereiche qualifiziert. Die Schule muss sich aber immer darauf befragen lassen, ob sie diese Auslesefunktion in einer Weise wahrnimmt, die, dem jeweiligen Kenntnisstand entsprechend, Vorstellungen und Anforderungen der sozialen Gerechtigkeit genügt. Sie muss sich auch der Frage stellen, wie sich die *gesellschaftliche Funktion* der Selektion zu *pädagogischen Zielsetzungen* verhält. Auch aufgrund von Befunden der pädagogischen Forschung wird ernsthaft in Zweifel gezogen, dass die heutige Schule mit der von ihr praktizierten Selektion das pädagogisch zentrale Ziel der optimalen, den individuellen Voraussetzungen des einzelnen angemessenen Förderung *möglichst aller Schüler* in ihren Fähigkeiten und ihrer Persönlichkeit befriedigend verwirklichen kann. Ich habe oben darauf hingewiesen, wie Schulstrukturen und die damit verbundene Selektion Grenzen und Erschwernisse für positive individuelle Entwicklungen darstellen können.

Der Wunsch, das pädagogische Ziel der optimalen Förderung jedes einzelnen noch besser zu verwirklichen, die Probleme und Unzulänglichkeiten der Selektion zu vermindern und die Qualität der Bildung und Ausbildung und das Niveau der dabei erreichten Qualifikationen zu erhalten oder noch zu heben, müssen auch in Zukunft wichtige Motoren für Bemühungen in Forschung, Politik und Schulpraxis um Weiterentwicklung unserer Schulen bleiben.

Anmerkung

¹ Wir verweisen hier auf einen ausführlichen Bericht zum gleichen Thema: Silvia Wytttenbach, Urs K. Hedinger: Schülerbeurteilung und Selektion. Amt für Unterrichtsforschung und -planung, Bern 1982. Der Bericht enthält u. a. eine Zusammenfassung von Untersuchungen zu Ausleseverfahren, die in verschiedenen Schweizer Kantonen angewendet werden. Ferner weitere Literaturhinweise.

Literaturhinweise

- Fend, H.: Gesamtschule im Vergleich. Bilanz der Ergebnisse des Gesamtschulversuchs. Weinheim 1982.
- Halbherr, P.: Leistungsnormen im Übergangsbereich Schule-Beruf. Bern 1980.
- Halbherr, P.: Die Selektion in die Oberstufe. In: Berufsberatung und Berufsbildung, Heft 4/1981.
- Hedinger, U. K.: Selektion und Schullaufbahn in unterschiedlichen Differenzierungssystemen, 1983 (im Druck).
- Ingenkamp, K. H.: Die Fragwürdigkeit der Zensurenggebung. Weinheim 1971.
- Preisig, E.: Ein Übertrittsverfahren auf dem Prüfstand. In: Bildungsforschung und Bildungspraxis, 1979, Heft 2/3.
- Rosenthal, R. / Jacobson, L.: Pygmalion im Klassenzimmer. Weinheim 1971.
- Ullrich, H. / Woebcke, M.: Notenelend in der Grundschule. München 1981.
- Wytttenbach, S. / Hedinger, U. K.: Schülerbeurteilung und Selektion. Amt für Unterrichtsforschung und -planung, Bern 1982.
- Zielinski, W.: Die Beurteilung von Schülerleistungen. In: Weinert et al.: Funkkolleg Pädagogische Psychologie, Bd. 2. Frankfurt a.M. 1974.

Was hält denn eigentlich diese verflixte Selektion?*

Anton Strittmatter

* Der Aufsatz stützt sich grösstenteils auf ein Referat zur «Selektivität der Primarschule», gehalten vom Verfasser am Europarat-Seminar «Primary Education in Western Europe» vom 9. bis 12. November 1982 in Vaduz FL.

Kaum ein Thema wird an öffentlichen Elternabenden so häufig und so heftig diskutiert wie das der Selektion. Kaum ein Thema wird von den Primarlehrern so häufig als das belastendste bezeichnet wie das der Selektion. Kaum ein