

Zeitschrift: Schweizer Schule
Herausgeber: Christlicher Lehrer- und Erzieherverein der Schweiz
Band: 77 (1990)
Heft: 10

Artikel: Was Intelligenztests über Generationenunterschiede lehren
Autor: Schallberger, Urs
DOI: <https://doi.org/10.5169/seals-533602>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. [Siehe Rechtliche Hinweise.](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. [Voir Informations légales.](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. [See Legal notice.](#)

Download PDF: 25.05.2025

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

Was Intelligenztests über Generationenunterschiede lehren

Urs Schallberger

Neuere Untersuchungen zeigen, dass das Durchschnittskind der achtziger Jahre in bestimmten Aufgabentypen von Intelligenztests deutlich mehr leistet, als dies beim Durchschnittskind der fünfziger Jahre der Fall war. Der folgende Artikel berichtet etwas genauer über diese Befunde, ihre Hintergründe und mögliche Bedeutung.

Einleitung

«Es ist nicht mehr so wie früher!» Diesen Ausspruch kennen wir aus den verschiedensten Zusammenhängen. Im allgemeinen soll damit ausgedrückt werden, dass sich ein bestimmter Sachverhalt (das Wetter, die Jugend, die Qualität von Waren, die Höflichkeit usw.) in der Gegenwart schlechter darstellt als eben «früher». Soweit es dabei um Aspekte menschlichen Verhaltens geht, sind damit *Generationen-* oder – was dasselbe bedeutet – *Kohortenunterschiede* angesprochen. Das Problem bei solchen Behauptungen ist, dass deren Richtigkeit oder Falschheit meist nur sehr schwer zu untersuchen ist. Denn Daten, die einen objektiven Vergleich zwischen den Generationen erlauben würden, sind rar. Dies gilt insbesondere auch für viele Vermutungen über Generationenunterschiede im Bereich schulischen Verhaltens. Die wenigen vorhandenen wissenschaftlichen Vergleiche zeigen zudem erstaunlicherweise, dass oft entsprechende Klagen der Grundlage entbehren. Manche der Klagen scheinen damit einen komplexeren Hintergrund zu haben, als im Wortlaut zum Ausdruck

kommt, beispielsweise können sie auch Ausdruck veränderter Anforderungen sein. Wie beispielsweise die sprachlichen Anforderungen gestiegen sind, zeigt sehr überzeugend ein in dieser Zeitschrift erschienener Artikel über Generationenunterschiede im Lesen und Schreiben (Sieber, 1989). Daraus ist allerdings nicht etwa zu schliessen, dass es immer so sein muss.

Die Frage nach Generationenunterschieden im Bereich schulischen Verhaltens, nach ihrem Ausmass, ihren Hintergründen und Konsequenzen bleibt damit ein interessantes Forschungsproblem. Und gerade auf dem Hintergrund der erwähnten schwierigen Datenlage scheint es sinnvoll zu sein, auch Informationen unter diesem Aspekt zu betrachten, die nicht unmittelbar dem schulischen Kontext entstammen.

Solche Informationen finden sich unter anderem im Zusammenhang mit dem Studium von *Generationenunterschieden in Intelligenztestleistungen*. Sie stehen im Zentrum dieses Artikels. Zuvor muss allerdings noch kurz auf die Grundidee von Intelligenztests eingegangen werden, damit das folgende richtig eingeordnet werden kann.

1. Was sind Intelligenztests?

Intelligenztests sind im Prinzip standardisierte Arbeitsproben, die unter einigermaßen normalen Umständen einen gewissen Rückschluss auf die intellektuelle Leistungsfähigkeit einer Person in einem bestimmten Zeitpunkt und bezogen auf bestimmte intellektuelle Anforderungen erlauben. Diese vielleicht etwas kompliziert anmutende «Definition» ist hier nötig, weil sich um Intelligenztests – übrigens selbst unter Psychologen – viele falsche Vorstellungen ranken, die zum Teil in eine Überschätzung, zum Teil in eine Unterschätzung ihrer Leistungskraft ausmünden. *Überschätzt* werden diese Tests etwa, wenn man glaubt, damit «die» Intelligenz eines Menschen «messen» zu können. «*Intelligenz*» ist ein Begriff mit einem höchst umfassenden und nur unscharf abgrenzbaren Bedeutungs-horizont. Er fasst eine Reihe in unserer Kultur hoch bewerteter Fähigkeiten

im intellektuellen, aber auch im sozialen und praktischen Bereich zusammen, von denen recht unklar ist, wie sie untereinander zusammenhängen. Intelligenztests betreffen dagegen – entsprechend der in ihnen zusammengefassten Aufgabentypen – ein nur ganz eng begrenztes Spektrum von Fähigkeiten. Aber auch in diesem eingeschränkten Sinne kann von «Messung» im alltäglichen Wortsinn keine Rede sein. «Messbar» sind höchstens die Intelligenztestleistungen und nicht die mit dem Intelligenzbegriff gemeinten Fähigkeiten. Bereits aus der Alltagserfahrung wissen wir, dass Leistungen noch von mehr und anderem abhängig sind, als von der eigentlich geforderten Fähigkeit, z.B. von der Motivation, situativen Bedingungen, von überdauernden oder momentanen Arbeitsstörungen etc. Der Schluss von Intelligenztestleistungen auf die Leistungsfähigkeit ist daher nicht so einfach wie etwa jener von der Höhe einer Quecksilbersäule auf die Temperatur, sondern eher vom Typus eines qualitativen Schätzvorgangs, wie er zum Beispiel auch notwendig ist, um von den aktuellen Schulleistungen eines Kindes auf dessen schulische Zukunft zu schliessen.

Unterschätzt werden Intelligenztests dann, wenn man aus den genannten und andern Gründen die Aussagekraft von Intelligenztests gänzlich in Frage stellt. Denn es gibt heute eine praktisch unübersehbare Fülle von empirischen Befunden, die zeigen, dass bereits die Intelligenztestleistungen eine sehr relevante Information vermitteln. So lehren z.B. Längsschnittstudien, dass die im späten Kindesalter feststellbaren Leistungsunterschiede zwischen Menschen über den ganzen Lebenslauf in erheblichem Ausmass stabil bleiben. Weiter erlauben Intelligenztestergebnisse gute Prognosen über schulische und berufliche Leistungen. Diese Prognosen sind zwar weit davon entfernt, perfekt zuzutreffen. Dies wäre aber auch dann der Fall, wenn wir die «wahre Intelligenz» eines Menschen kennen würden. Wie ja schon festgehalten wurde: Konkrete Leistungen sind noch von mehr und anderem abhängig als von den eigentlich angesprochenen Fähigkeiten!

Im folgenden beschränken wir unsere Überlegungen ganz auf die Ebene der Intelligenztestleistungen, also auf den unmittelbar

beobachtbaren Sachverhalt. Das Mass für eine individuelle Testleistung besteht in der Anzahl gelöster Aufgaben. Diese Zahl – genannt *Rohwert* – sagt aber erst etwas aus, wenn man sie mit einem Standard vergleicht und in einem *Normwert* ausdrückt. Als solcher Normwert dient bei vielen (nicht allen) Intelligenztests der sogenannte «Intelligenz-Quotient» (IQ). Dieser berühmte-berüchtigte Kennwert ist eine rein statistisch definierte Grösse: Im Rahmen der Testentwicklung, die aus mehreren Phasen besteht (siehe z.B. Schelten, 1980), wird auch eine sogenannte *Normierung* oder *Eichung* durchgeführt. Diese besteht darin, dass der Test in einer für den künftigen Anwendungsbereich repräsentativen Stichprobe angewendet wird. Hauptergebnis einer Eichung ist die Verteilung der in dieser Stichprobe gefundenen Rohwerte, die graphisch dargestellt in der Regel wie in Abbildung 1 aussieht:

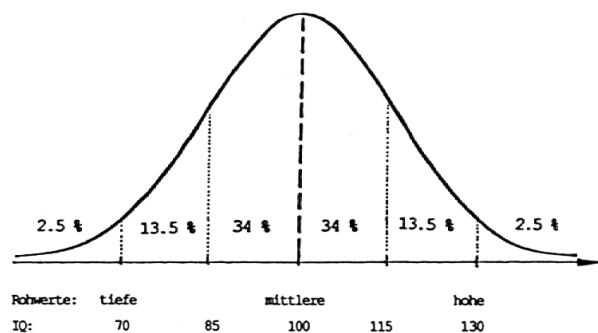


Abb. 1: Definition des IQ durch die Verteilung der Rohwerte

Die Rohwerte massieren sich in einem Mittelbereich und werden immer seltener, je weiter sie von diesem Mittelbereich entfernt sind. Auf der Basis dieser Verteilung wird dann zuhanden der künftigen Testbenützer eine *Normentabelle* erstellt, in der jedem Rohwert ein bestimmter Normwert zugeordnet wird. Werden IQs als Normwerte benützt, wird dem Mittelwert (Durchschnitt) der gefundenen Rohwertverteilung (willkürlich) der Normwert (IQ) 100 zugeteilt. Die Zuordnung der übrigen Rohwerte zu IQs erfolgt so, dass die in Abbildung 1 eingetragenen prozentualen Verhältnisse entstehen. Mit Hilfe dieser Normentabelle kann dann der Testbenützer für jeden möglichen Rohwert einen zugehörigen IQ bestimmen. Dies ermöglicht – sofern man mit der statistischen Definition

des IQs vertraut ist – eine einfache Interpretation individueller Testleistungen: Erreicht ein Kind einen Rohwert, dem gemäss Normentabelle der IQ von 98 entspricht, ist für den informierten Testleiter sofort klar, dass es sich dabei um ein ungefähr durchschnittliches Ergebnis handelt; wird ein IQ von 133 erreicht, verweist dies – auch ohne Hilfe von Abbildung 1 – auf eine deutlich überdurchschnittliche Leistung, die nur noch von wenigen Altersgenossen übertroffen wird. Erst diese Normentabellen machen also die Intelligenztests zu dem, was sie sind: Zum nützlichen Hilfsmittel, um mit einem relativ beschränkten Aufwand einen einigermaßen objektiven Vergleich der geprüften (getesteten) Leistung eines Kindes oder eines Erwachsenen mit seiner Referenzgruppe zu erhalten.

2. Generationenunterschiede in Intelligenztestleistungen

Aus der oben dargestellten Logik der Bestimmung von IQs ergibt sich trivialerweise, dass sich in einer neuen Repräsentativstichprobe auf der Ebene der IQs wieder ein Mittelwert von 100 ergeben müsste. Würde der Mittelwert grösser oder kleiner als 100 ausfallen, dann würde dies bedeuten, dass sich die durchschnittliche Leistungsfähigkeit der Probanden in der seit der Eichung verflossenen Zeit verbessert oder verschlechtert hat, dass also ein Generationenunterschied vorliegt. Solche Repräsentativuntersuchungen sind aber sehr aufwendig; entsprechend gibt es keine Studie dieser Art im deutschen Sprachraum. Basierend auf derselben Grundidee kann man nun aber auch mittels einer einfacher durchzuführenden Methode nach Generationenunterschieden fragen. Voraussetzung für diese einfachere Methode ist, dass zwei inhaltlich praktisch identische Intelligenztests zur Verfügung stehen, die zu zwei verschiedenen Zeitpunkten geeicht worden sind. Testet man dann dieselben Personen mit beiden Tests, widerspiegelt die IQ-Differenz das Ausmass der «Aufwertung» oder «Abwertung» derselben Leistung im Gefolge einer allgemeinen Leistungsabnahme oder -zunahme zwischen den Eichungen der beiden Tests. Die so gefundene

IQ-Differenz ist also ein Mass für den Leistungsunterschied der beiden Generationen, denen die Eichstichproben entstammen.

Ein solcher Testvergleich war im deutschen Sprachraum erstmals möglich, als im Jahre 1983 eine revidierte Version des sog. «Hamburg-Wechsler-Intelligenztests für Kinder» (HAWIK, Hardesty & Priester, 1956), genannt HAWIK-R (Tewes, 1983) erschien. Die beiden Tests sind konzeptionell identisch; der eine ist ca. 1955, der andere ca. 1981 geeicht worden. Inzwischen liegen sechs Vergleichsuntersuchungen vor, die Rückschlüsse auf Leistungsverbesserungen bzw. -verschlechterungen im Zeitraum zwischen den beiden Eichungen (also von ca. 1955 bis ca. 1981) zulassen. Wertet man diese Untersuchungen vergleichend aus (Schallberger, 1987, 1990), dann ergeben sich folgende Hauptbefunde:

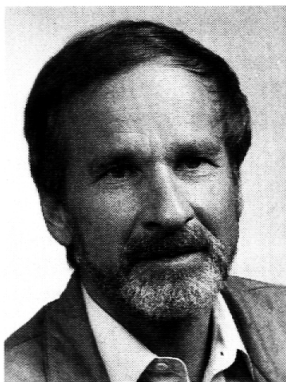
- 1.) Die Leistung im Gesamtest hat eine massive Zunahme erfahren. Sie beträgt – in IQ-Punkten ausgedrückt – ca. 15 Punkte. 15 IQ-Punkte, das entspricht etwa dem Unterschied zwischen ...

 - ... einem 10jährigen und einem 12jährigen Durchschnittskind;
 - ... dem Durchschnittskind und einem wegen leichter geistiger Behinderung IV-berechtigten Kind derselben Altersgruppe;
 - ... dem durchschnittlichen Mittelschüler und dem durchschnittlichen Angelernten in einer repräsentativen Stichprobe von Jugendlichen (Schallberger, 1988).

Die drei Beispiele illustrieren, dass der festgestellte Generationenunterschied von einer ganz erstaunlichen Grössenordnung ist. Würde sich zum Beispiel ein Durchschnittskind von ca. 1951 in eine heutige schulpyschologische Praxis verirren, müsste es damit rechnen, dass man sich ernsthaft mit seiner Sonderschulbedürftigkeit befassen würde. Diese Grössenordnung hat bisher kaum jemand für möglich gehalten. Entsprechend gibt es Autoren, die versuchen, jene IQ-Unterschiede auf andere Weise zu erklären als durch Generationenunterschiede. Die dabei verwendeten Argumente sind allerdings nicht überzeugend (vgl. dazu Schallberger, 1990).

2.) Intelligenztests enthalten in der Regel nicht nur einen, sondern mehrere verschiedene Aufgabentypen. So umfasst zum Beispiel der HAWIK(-R) elf verschiedene Arten von Aufgaben. Und die Leistungszunahme im Gesamttest widerspiegelt einfach die durchschnittliche Leistungszunahme in all diesen Aufgabentypen. Eine natürliche Folgefrage ist, ob es zwischen den Aufgabentypen Unterschiede gibt. Dies ist tatsächlich der Fall: Es gibt Typen von Aufgaben, bei denen es keine oder praktisch keine Generationenunterschiede gibt und andere, bei denen die Unterschiede so gross sind, dass sie – bestünde der Test nur aus solchen Aufgaben – IQ-Unterschiede von deutlich über 20 IQ-Punkten ergäben. Leider sind die Befunde im Bereich der einzelnen Aufgabentypen nicht so gut abgesichert wie im Gesamttest. Doch lässt sich immerhin folgendes festhalten:

- Es gibt keinen einzigen Aufgabentyp, bei dem ein Generationenunterschied im Sinne einer Leistungs*abnahme* feststellbar wäre. Die Leistung der Kinder zu Beginn der achtziger Jahre ist in allen Aufgabentypen mindestens so gut wie jene der Kinder in den fünfziger Jahren.
- Praktisch gleich geblieben sind die Leistungen im Bereich der Merkfähigkeit (Kurzzeitgedächtnis), ebenso in Aufgaben, die in Form von Sätzchenaufgaben das rechnerische Denken prüfen.



Urs Schallberger, Dr. phil., geboren 1941 in Uznach SG. Ausbildung zum Berufsberater und Studium der Psychologie, Mathematik und Philosophie. Gegenwärtig Oberassistent und Lehrbeauftragter am Psychologischen Institut der Universität Zürich. In seiner Forschung befasst er sich unter anderem mit methodischen und inhaltlichen Problemen von Intelligenztests.

- Mittlere Zunahmen – etwa in der Grössenordnung des Gesamttests – finden sich in sprachlichen Aufgaben verschiedener Art, beispielsweise in einem Wortschatztest, aber auch in Aufgaben, in denen das Sprachverständnis zumindest mitgeprüft wird. Ebenfalls eine mittlere Leistungszunahme ergibt sich bei einer psychomotorischen Aufgabe, deren Lösung auch in erheblichem Ausmass Konzentrationsfähigkeit voraussetzt.
- Auffällige bis extreme Leistungszunahmen finden sich in verschiedenen Aufgabentypen, bei denen es vor allem darauf ankommt, abstrakt-logisch zu denken, bzw. Wesentliches von Unwesentlichem zu unterscheiden. Ein Beispiel ist das Finden von Oberbegriffen zu zwei vorgegebenen Begriffen. Andere Beispiele sind mehr spielerischer Natur: So geht es etwa bei einer Aufgabe darum, in schematischen Zeichnungen bekannter Gegenstände fehlende wichtige Teile zu identifizieren; oder bei einer andern, die einzelnen Zeichnungen einer Bilder-geschichte in die richtige Reihenfolge zu bringen.
- Fasst man die Aufgaben nach dem Kriterium in zwei Gruppen zusammen, ob explizit auch Bildungsinhalte (Sprache, Rechnen) mitgeprüft werden oder nicht, fällt auf, dass in den Aufgaben, die keine derartigen Bildungsinhalte voraussetzen, insgesamt deutlich grössere Leistungszunahmen festzustellen sind als in jenen Aufgaben, die sprachliche und/oder rechnerische Anforderungen stellen. Oder anders gesagt: Je mehr Schulwissen zur Lösung der Aufgaben nötig ist, umso geringer ist die Leistungszunahme.

Am Rande sei noch erwähnt, dass ähnliche Beobachtungen über Generationenunterschiede in Intelligenztestleistungen auch in einer Reihe anderer moderner Industrieländer gemacht worden sind (Flynn, 1987).

3. Das Problem der Erklärung

Konfrontiert mit diesen Befunden stellt sich nun natürlich sofort die Frage, wie diese zum Teil phänomenalen Leistungs-

nahmen zu erklären sind. Darüber kann im heutigen Zeitpunkt nur spekuliert werden. Im wesentlichen scheint es vier mögliche Erklärungen zu geben, die sich aber gegenseitig nicht ausschliessen. Es ist sogar zu vermuten, dass diese vier Erklärungen alle gleichzeitig eine gewisse, wenn auch beschränkte Gültigkeit beanspruchen können, und dass sie erst *zusammen* eine wirklich *hinreichende* Erklärung ergeben:

1.) Eine erste Erklärung bezieht sich auf Äusserlichkeiten, zum Beispiel die Zunahme der Testgewandtheit der Kinder oder auf grössere Vertrautheit der Kinder mit testartigen Aufgaben. Dies mag sicher eine Rolle spielen, erklärt aber kaum die volle Differenz. Zum Beispiel wäre auf einem solchen Hintergrund nicht einzusehen, warum es so grosse und systematische Unterschiede zwischen den verschiedenen Aufgabentypen gibt. Weiter müsste man einen gewissen «Plateau-Effekt» erwarten; wie Trainingsstudien zeigen, ist der überhaupt mögliche Effekt steigender Testgewandtheit oder -vertrautheit in seiner Grösse begrenzt.

2.) Eine zweite Erklärung beruft sich auf sozioökonomische Veränderungen in modernen Gesellschaften. Es ist ja schon lange bekannt, dass deutliche Leistungsunterschiede zwischen Kindern aus verschiedenen sozialen Schichten festzustellen sind. Nun haben sich aber in den letzten Jahrzehnten die Lebensbedingungen vor allem in den unteren Schichten gewaltig verbessert, ebenso das Ausmass der Teilhabe an Kultur und Bildung. Durch diese sozioökonomischen Verschiebungen sei auch – so lautet das Argument – das durchschnittliche intellektuelle Leistungsniveau in der Gesamtheit der Bevölkerung angestiegen. Auch diese Erklärung reicht nicht aus. Hauptgrund ist, dass in allen Schichten und Bildungsgruppen Leistungszunahmen aufgetreten sind. Beim Motor dieser allgemeinen Leistungszunahme muss es sich also um etwas handeln, das alle Bevölkerungskreise gleichermassen betrifft, und nicht nur einen Teil.

3.) Eine dritte Erklärung verweist im Zusammenhang mit einem solchen generellen Faktor auf das Schulsystem; dieses sei

besser geworden, bewirke eine breitere Förderung der intellektuellen Leistungsfähigkeit. Ein Problem bei dieser Erklärung ist, dass bereits bei Kindern im Vorschulalter Leistungszunahmen festzustellen sind. Weiter lässt diese Erklärung nicht gut verstehen, warum – wie im 2. Abschnitt hervorgehoben wurde – die explizit schulabhängigen Leistungen deutlich weniger angestiegen sind als jene, die kein Vorwissen dieser Art erfordern. Flynn (1984) stellt für die USA sogar fest, dass es ein Kontinuum zwischen Leistungszunahme und -abnahme gibt, wenn man die Aufgabentypen nach ihrer Bildungsabhängigkeit ordnet: Die grössten Zunahmen ergeben sich bei sprachfreien, von expliziten Bildungsinhalten völlig unabhängigen Aufgabentypen. Hingegen zeigen in den USA Aufgabentypen, deren Lösung Bildungsinhalte voraussetzen, sogar Abnahmen.

4.) Es ist also auch noch eine vierte Erklärung notwendig, nämlich die, dass sich in den letzten Jahrzehnten in den alltäglichen Lebensbedingungen, in den alltäglichen Entwicklungsbedingungen etwas massiv verändert haben muss, etwas, was die Entwicklung der intellektuellen Aktivitäten unserer Kinder und Jugendlichen wesentlich stärker stimuliert als jene früherer Generationen. Mögliche Konkretisierungen für diese Pauschalaussage findet man natürlich rasch auf allen Ebenen, von den Spielsachen über die Medien bis hin zu den vielfältigen kognitiven «Reizen» der modernen Lebenswelt, die fast permanent wirksam sind; bei einer Eisenbahn- oder Autofahrt, beim Gang durch eine Einkaufsstrasse, etc. Die Rolle dieser basalen und allgegenwärtigen Stimulierung, dieses im Unterschied zu bewusst konzipierten Erziehungs- und Bildungsmaßnahmen eigentlich «nicht-intendierten» Trainings, scheint bisher unterschätzt worden zu sein.

4. Generationenunterschiede in Intelligenztestleistungen und die Schule

Ich bin kein «Schulmann». Trotzdem will ich zum Abschluss kurz zwei Vermutungen darüber wagen, was das Gesagte für die Schule bedeuten könnte.

Zunächst kann ich mir nicht vorstellen, dass die Entwicklungen, die hinter den dargestellten Leistungszunahmen in Intelligenztests stehen, spurlos am Geschehen in den Klassenzimmern vorbeigegangen sind. Verglichen mit den Altersgenossen der fünfziger Jahre verfügt das Kind der achtziger Jahre nicht nur über mehr Alltagswissen, es ist auch geistig beweglicher und leistet intellektuell deutlich mehr. Damit muss sich das Geschehen im Klassenzimmer radikal verändert haben. Dass man diese Veränderung nicht oder nicht in vollem Ausmass realisiert hat, hängt wohl damit zusammen, dass sie sich in einem schleichenden Prozess realisierte – über all die Jahre in ganz kleinen Schritten erfolgte. Es ist aber undenkbar, dass eine durchschnittliche Leistungszunahme im Gegenwert von 15 IQ-Punkten nicht zu deutlichen Veränderungen geführt hat.

Zum andern sagen unsere Befunde allerdings nichts darüber, *wie* diese Veränderungen im Schulzimmer aussehen. Was einzig festzustehen scheint: Die Zunahme der intellektuellen Leistungen hat sich nicht direkt und proportional in einer analogen Zunahme des schulischen Wissens und Könnens niedergeschlagen. Soweit der Intelligenztest, um den es hier geht, diesbezüglich Rückschlüsse zulässt, sind ja die Leistungszunahmen im sprachlich-rechnerischen Bereich klar geringer als im abstrakt-logischen Bereich. Möglicherweise bedeutet dies, dass die Kinder zwar geistig «vifer» geworden sind, diese grössere «Vifheit» aber nicht unbedingt schön brav in eine entsprechend grössere Lernbereitschaft umsetzen. Heisst dies, dass sie vielleicht auch kritischer geworden sind? Mehr nach dem Sinn von Schulinhalt fragen? Dem eigenen Urteil mehr Gewicht zuschreiben? Versucht man sich dies konkret vorzustellen, dann drängt sich die Vermutung auf, dass das Unterrichten trotz oder gerade wegen des beschriebenen Generationenunterschieds nicht einfacher geworden ist!

Literaturverzeichnis

- Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, 95, 29–51.
- Flynn, J. R. (1987). Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*, 101, 171–191 (siehe dazu auch: Editorial note, Correction to Flynn. *Psychological Bulletin*, 101, 427.)
- Hardesty, F. P. & Priester, H. J. (1956). Handbuch für den Hamburg-Wechsler-Intelligenztests für Kinder (3. Auflage 1966). Bern: Huber.
- Schallberger, U. (1987). HAWIK und HAWIK-R: Ein empirischer Vergleich. *Diagnostica*, 33, 1–13.
- Schallberger, U. (1988). Berufsausbildung und Intelligenzentwicklung. In K. Häfeli, U. Kraft & U. Schallberger (Hrsg.), *Berufsausbildung und Persönlichkeitsentwicklung. Eine Längsschnittstudie* (S. 148–167). Bern: Huber.
- Schallberger, U. (1990). Das Ausmass des IQ-Gewinns im deutschen Sprachraum von ca. 1956 bis ca. 1983. Zürich: Psychologisches Institut der Universität, Abteilung Angewandte Psychologie.
- Schelten, A. (1980). Grundlagen der Testbeurteilung und Testerstellung. Heidelberg: Quelle & Meyer (UTB 995).
- Sieber, P. (1989). Neue Anforderungen an die Sprache in Schule und Öffentlichkeit. *Schweizer Schule*, 76, Nr. 10.
- Tewes, U. (1983). HAWIK-R. Handbuch und Testanweisung. Bern: Huber.