

## Kap. 6.2: Methodologische Probleme und Möglichkeiten bei der längsschnittlichen Analyse motorischer Entwicklungsverläufe

Wolfgang Schneider

### 1. Einleitung

Der motorischen Entwicklung kommt insbesondere in den ersten Lebensjahren eine besondere Bedeutung zu. Frühe Bewegungserfahrungen nehmen bedeutsamen Einfluß auf die kindliche Gesamtentwicklung (vgl. Schilling 1982; Scheid 1989). Die Bedeutung des motorischen Entwicklungsbereiches für die frühe Kindheit scheint darin begründet, daß Kinder in diesem Funktionsbereich relativ früh eigene Kompetenz und Autonomie ausbilden. Aber auch in späteren Lebensphasen lassen sich enge Beziehungen zwischen motorischen Aktivitäten und Ausprägungen von Persönlichkeitsmerkmalen nachweisen (vgl. Zimmer 1981). Aspekte der motorischen, kognitiven, affektiven und sozialen Entwicklung bleiben eng miteinander verknüpft, so daß Veränderungen in einem Bereich auch Auswirkungen auf die anderen haben. Sportliche Erfolge von Kindern und Jugendlichen führen oft zu besserer sozialer Anpassung und Stellung innerhalb der "peer group" (vgl. Singer 1985; Bierhoff-Alfermann 1986).

Obwohl die Zahl empirischer Forschungsarbeiten zur motorischen Entwicklung innerhalb der letzten beiden Jahrzehnte stark zugenommen hat, sind eindeutige Aussagen über motorische Entwicklungsverläufe im Kindes- und Jugendalter und ihre Beziehung zur Sozial- und Persönlichkeitsentwicklung insofern nur schwer möglich, als die überwältigende Mehrzahl von Studien querschnittlich angelegt ist: Unterschiedliche Kinder aus unterschiedlichen Altersstufen werden im Hinblick auf bestimmte Komponenten der motorischen Entwicklung miteinander verglichen (vgl. Arbinger 1990). Aus diesen Vergleichen werden dann Rückschlüsse auf den allgemeinen Entwicklungsverlauf der betreffenden Merkmale gezogen, wobei unterstellt wird, daß sich vergleichbare Befunde ergeben hätten, wenn immer wieder dieselben Kinder zu unterschiedlichen Zeitpunkten untersucht worden wären. Aus Entwicklungsunterschieden (querschnittlicher Ansatz) wird demnach auf Entwicklungsveränderungen (längsschnittlicher Ansatz) geschlossen, die eigentlich wesentlich mehr interessieren.

Die Unzulänglichkeit querschnittlicher Ansätze für die Rekonstruktion intraindividuell-er Veränderungen ist in der entwicklungspsychologischen Forschung seit geraumer Zeit bekannt (vgl. *Wohlwill* 1973; *McCall* 1977). Die Notwendigkeit längsschnittlicher Analysen wird so auch von all denjenigen Forschern eingesehen, die sich näher mit Entwicklungsveränderungen beschäftigt haben. Dennoch fällt auf, daß immer noch relativ wenige Längsschnittstudien vorliegen. Das Gebiet der motorischen Entwicklung stellt dabei keine Ausnahme dar.

Die wenigen verfügbaren Längsschnittstudien im Bereich der motorischen Entwicklung (z.B. *Diem* et al. 1980; *Zimmer* 1981; *Rieder* et al. 1982) bieten zwar wertvolle Informationen zu intraindividuellen Veränderungen in motorischen Kompetenzen und ihrer Relation zu Persönlichkeitsmerkmalen und erschließen damit wertvolles Terrain, schöpfen die Möglichkeiten längsschnittlicher Forschungsmethoden jedoch bei weitem nicht aus.

Neben untersuchungstechnischen Gründen sind es oft auch methodologische Fragen, die Forscher von Längsschnittuntersuchungen abhalten. Als typische Probleme von Längsschnittuntersuchungen seien genannt: Fragen der Designplanung, des Datenschwundes, der Reliabilität von Veränderungsmessungen oder die Abhängigkeit beobachtbarer Merkmalsveränderungen vom Ausgangswert. Eine Übersicht über methodologische Probleme bei Längsschnittuntersuchungen geben *Helmreich* (1977) und *Petermann* (1978). Mit Blick auf die motorische Entwicklung finden sich eine Reihe methodologisch interessanter Beiträge in dem Sammelband von *Willimczik/Grosser* (1979).

Der vorliegende Beitrag will Probleme und Möglichkeiten längsschnittlicher Ansätze bei der Analyse motorischer Entwicklungsverläufe aufzeigen. Es werden zunächst praktische und konzeptuelle Probleme kurz skizziert, bevor dann auf methodologische Fragen Bezug genommen wird. Methodische Probleme betreffen einmal die Frage, wie Veränderungen im motorischen Entwicklungsstand zuverlässig erfaßt werden können. Zum anderen ist damit die Schwierigkeit angesprochen, Entwicklungsunterschiede in motorischen Kompetenzen zu prognostizieren und zu erklären. Der Schwerpunkt der Ausführungen wird deshalb einmal auf der Analyse intraindividuell-er Veränderungen, zum anderen auf der Analyse interindividuell-er Unterschiede liegen. Die Möglichkeiten der längsschnittlichen Methode werden jeweils anhand von Daten zur motorischen Entwicklung von Kindern und Jugendlichen demonstriert.

## 2. Praktische und konzeptuelle Probleme von Längsschnittstudien

### 2.1 Praktische Probleme

Offensichtliche praktische Probleme von Längsschnittstudien liegen einmal im Kostenfaktor, zum anderen in der normalerweise zu erwartenden Aufwendigkeit und schließlich in dem Belohnungsaufschub, d.h. in der langen Wartezeit bis zum Vorliegen interpretierbarer und publizierbarer Befunde (vgl. *Harway et al.* 1984; *Schneider* 1989). Das Kostenproblem wird evident, wenn es sich um Langzeitprojekte mit umfangreichen Stichproben handelt. Üblicherweise sind diese Projekte auch insofern aufwendig, als die Ausfallquote (drop-out-Quote) bei den Probanden niedrig gehalten werden muß, um die Repräsentativität der Stichprobe zu sichern. Der hierzu erforderliche besonders große Aufwand bei der Betreuung der Probanden wird als "Stichprobenpflege" gekennzeichnet. Schließlich müssen die mit Längsschnittstudien betrauten Wissenschaftler in der Regel längere Zeit darauf warten, bis die Früchte ihrer Arbeit publiziert und damit einem breiteren Publikum vorgestellt werden können. Als problematisch kann es sich hierbei erweisen, daß die Attraktivität der Ausgangsfragestellung nach einigen Jahren so nicht mehr vorhanden ist, was gerade bei der in den Verhaltens- und Sozialwissenschaften erkennbaren Orientierung an Modetrends fatale Folgen für die Publizierbarkeit von Befunden hat.

### 2.2 Konzeptuelle Probleme

Eine grundlegende Schwierigkeit bei der Diskussion von Problemen, wie sie bei Längsschnittstudien auftreten können, ist darin zu sehen, daß es keine präzise Definition von Längsschnittstudien gibt. Der Begriff wird nicht zur Beschreibung einer spezifischen Methode verwendet, sondern beschreibt eine ganze Palette unterschiedlicher Ansätze, die von Einzelfallstudien in Zeitreihen-Experimenten bis zu breit angelegten Langzeitstudien mit mehreren tausend Probanden reichen kann (vgl. *Baltes/Nesselroade* 1979). Dies gilt natürlich auch für den Bereich der motorischen Entwicklung. In einem kürzlich erschienenen Sammelband zum Beitrag von Längsschnittstudien für die Beschreibung und Erklärung motorischer Entwicklung in der frühen und späten Kindheit (*Kalverboer et al.* 1993) ist die Breite dieser Palette gut dokumentiert: Während Untersuchungen zur motorischen Entwicklung bei Föten nur eine eindeutig begrenzte Zeitspanne umfassen und auch die Phase des Fertigkeitserwerbs von kleinen Kindern beim Umgang mit Löffeln auf wenige Monate beschränkt ist, verlangen etwa Studien zum Einfluß bestimmter Risikofaktoren bei der Geburt auf die motorische Entwicklung im frühen und mittleren Kindesalter komplexe Versuchsanlagen, die sich über einen Zeitraum von zehn Jahren und mehr erstrecken können (vgl. *Schneider* 1993a).

Ein weiteres konzeptuelles Problem bei vielen Längsschnittstudien ist darin zu sehen, daß unterschiedliche Ziele im Mittelpunkt stehen. So kann es einmal um die Beschreibung und Erklärung von allgemeinen "Entwicklungsfunktionen" gehen, etwa Kurven zur Abbildung der körperlichen Wachstumsgeschwindigkeit oder zur Körperkraft, während in anderen Längsschnittstudien die Frage der individuellen Unterschiede besonders interessiert. Hier geht es demzufolge um das Problem, inwieweit die Personen der Stichprobe ihren relativen Rangplatz in der Bezugsgruppe über verschiedene Zeitpunkte hinweg beibehalten. Behandeln Studien zur Entwicklungsfunktion die Frage der *Kontinuität versus Diskontinuität* von Entwicklungsmustern, so gehen Untersuchungen zum Problem von individuellen Differenzen auf Aspekte der längsschnittlichen *Stabilität versus Labilität* von Personunterschieden ein. Beide Fragen können empirisch korreliert sein, sind jedoch konzeptuell unabhängig voneinander zu sehen (vgl. *Appelbaum/McCall* 1983). Schon vor Beginn einer Längsschnittstudie sollte den Autoren von daher klar sein, welches der beiden Ziele sie verfolgen wollen.

### 3. Methodologische Probleme und Möglichkeiten von Längsschnittstudien

Im folgenden soll zunächst auf einen Mythos bzw. ein Gerücht eingegangen werden, das die Längsschnittforschung lange belastet hat (vgl. *Rogosa* 1988). Es geht hier um die Frage, wie sich Veränderungen erfassen lassen. Im Anschluß daran werden methodische Ansätze diskutiert, die für die Analyse allgemeiner motorischer Entwicklungstrends und die Erkundung von interindividuellen Unterschieden geeignet sind. Zur Illustration werden Daten aus Längsschnittprojekten zu bestimmten Aspekten motorischer Entwicklung ausgewertet, an denen der Verfasser beteiligt war. Es handelt sich dabei zum einen um die Münchner Längsschnittstudie LOGIK (Longitudinalstudie zur Genese individueller Kompetenzen; siehe *Weinert/Schneider* 1989), zum anderen um eine Längsschnittstudie zur motorischen Entwicklung von Tennistalenten (*Rieder et al.* 1983; *Schneider et al.* 1993).

#### 3.1 Probleme mit der Veränderungsmessung.

##### 3.1.1 Differenzwerte als Indikatoren für Veränderung

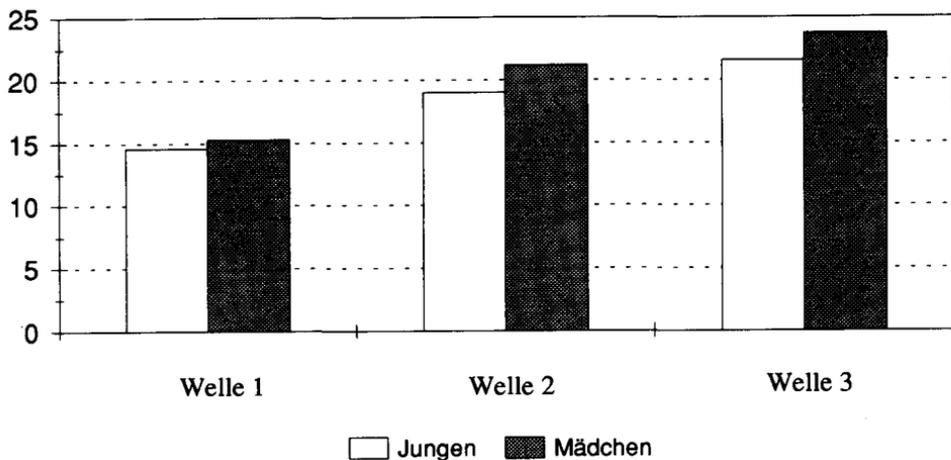
Seit der klassischen Arbeit von *Cronbach/Furby* (1970) wird davor gewarnt, Differenzwerte als Indikatoren für Veränderung zu benutzen. Vielfach wurde dabei nicht beachtet, daß sich diese Warnung nicht auf die Analyse von Gruppendaten bezog. So läßt sich beispielsweise leicht belegen, daß Varianzanalysen mit Meßwiederholungen, die auf Prätest- und Posttest-Rohwerten basieren, und Varianzanalysen,

die auf den Differenzen dieser Werte aufbauen, zu identischen Ergebnissen kommen (Maxwell/Howard 1981; Nunnally 1982).

Die Kritik bezog sich vielmehr auf die Angemessenheit von Differenzwerten bei der Erfassung intraindividuelle Veränderungen. Hier wurde insbesondere mit der Unreliabilität von Differenzwerten argumentiert, die die Analyse intraindividuelle Wachstumskurven enorm erschweren, ja in den meisten Fällen unmöglich machen sollte. Diese Lehrmeinung wurde erst durch neuere Arbeiten (z.B. Bryk/Raudenbush 1987; Rogosa 1988) grundsätzlich widerlegt. In diesen Studien wurde etwa herausgestellt, daß die Reliabilität von Differenzwerten nicht generell niedrig ist: Niedrige Werte resultieren lediglich dann, wenn sich die individuellen Wachstumsraten für die Probanden einer gegebenen Stichprobe nur wenig voneinander unterscheiden. Der Reliabilitätskennwert gibt demnach im wesentlichen die Genauigkeit an, mit der sich Individuen einer Stichprobe aufgrund ihrer Differenzwerte auf einer geschätzten Wachstumskurve in eine Rangfolge bringen lassen; dies unabhängig davon, wie präzise die Schätzung erfolgt. Niedrige Reliabilität deutet also nicht notwendig auf mangelnde Präzision der Messung hin und schließt von daher eine sinnvolle Erfassung intraindividuelle Veränderung nicht aus. Die Reliabilitätskennwerte fallen immer dann respektabel aus, wenn sich große interindividuelle Unterschiede in den intraindividuelle Zuwachsraten feststellen lassen.

Im folgenden wird anhand der LOGIK-Daten zu zeigen versucht, wie sich intraindividuelle Veränderungen in Merkmalen motorischer Entwicklung und interindividuelle Unterschiede in diesen Veränderungen im Rahmen eines Modells überprüfen lassen, das die Autoren als Hierarchical Linear Modeling (HLM) bezeichnet haben (vgl. Bryk/Raudenbush 1987). Die Daten wurden bei einer Stichprobe von annähernd 200 Kindern erhoben, die den Motoriktest für vier- bis sechsjährige Kinder (MOT 4-6) von Zimmer/Volkamer (1984) insgesamt dreimal im jährlichen Abstand absolvierten, als sie 4, 5 und 6 Jahre alt waren. Dieser Test erfaßt insbesondere Aspekte der Feinmotorik, wobei eine ganze Reihe motorischer Fertigkeiten wie etwa das Gefühl für Balance, die Bewegungsgenauigkeit und koordinatorische Fähigkeiten abgedeckt werden (vgl. Bös 1987). Für jede der 17 Aufgaben werden maximal 2 Punkte vergeben, was in einem Maximalwert von 34 Punkten resultiert. Abbildung 1 gibt die Durchschnittswerte (Rohpunkte) für die drei Meßzeitpunkte wieder, wobei die Ergebnisse einmal nach Geschlecht, zum anderen danach unterschieden werden, ob es sich um die jüngeren oder älteren Kinder der Stichprobe (Altersstreuung: ca. 5 Monate) handelte.

### Geschlechtsunterschiede im Motoriktest MOT4-6



### Altersunterschiede im Motoriktest MOT4-6

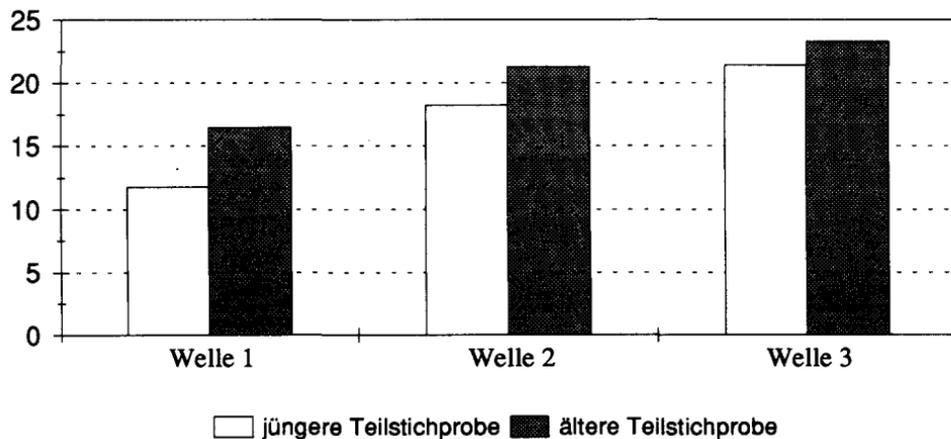


Abb. 1: Entwicklungsveränderungen im Motoriktest MOT 4-6 als Funktion der Geschlechts- und Altersgruppenzugehörigkeit (nach Schneider 1993a)

Traditionelle Verfahren zur Analyse von Gruppenunterschieden (zweifaktorielle Varianzanalysen mit Meßwiederholungen) mit Geschlecht als unabhängiger Variable ergaben Haupteffekte des Zeitfaktors und des Geschlechts. Alle Kinder steigerten ihre Leistungen mit zunehmendem Alter, wobei Mädchen allgemein besser abschnitten als Jungen. Signifikante Geschlechtsunterschiede ergaben sich hier allerdings nur für den zweiten und dritten Meßzeitpunkt. Ähnliche Befunde ergaben sich für die Varianzanalyse mit den beiden Altersgruppen als unabhängigem Faktor. Die Leistungsunterschiede zwischen den jüngeren und älteren Kindern der Stichprobe waren allerdings zu allen Meßzeitpunkten signifikant.

### 3.1.2 Analyse intraindividuelle Veränderungen mit HLM

Die wesentlichen Vorzüge dieses Verfahrens lassen sich wie folgt zusammenfassen (siehe *Bryk/Raudenbush* 1987, für technische Details): Über das Verfahren ist es einmal möglich, die durchschnittliche Wachstumskurve für eine vorgegebene Stichprobe zu schätzen. Darüber hinaus besteht die ungleich interessantere Option, individuelle Wachstumskurven für jeden Probanden zu schätzen und die Abweichung zum allgemeinen Kurvenverlauf zu bestimmen. Das Computerprogramm erlaubt zusätzlich eine Schätzung der Reliabilität sowohl für das Eingangsniveau eines gegebenen Merkmals als auch die Schätzung der Reliabilität der Veränderungswerte. Es ist schließlich möglich, Schätzungen der Korrelation zwischen Eingangsniveau und Veränderungsrate vorzunehmen, und theoretisch relevante Hintergrundmerkmale in die Analyse mit einzubeziehen, die individuelle Unterschiede in den Wachstumsraten erklären können.

Für die hier relevanten LOGIK-Daten wurde im ersten Schritt die mittlere Wachstumskurve geschätzt. Es wurde mit dem einfachsten Ansatz begonnen, der in diesem Fall in einem linearen Wachstumsmodell besteht. Es zeigte sich, daß das lineare Trendmodell besser mit den Daten übereinstimmte als Polynome höherer Ordnung, was uns dazu veranlaßte, das lineare Modell beizubehalten. Nachdem die individuellen Wachstumsraten auf der Basis dieses Modells bestimmt waren, ließen sich die individuellen Abweichungen von der durchschnittlichen Wachstumskurve näher explorieren. Die dazu verwendete Chi-Quadrat-Prüfstatistik ergab einen signifikanten Koeffizienten, was besagt, daß sich die Kinder in ihrem Eingangsstatus signifikant voneinander unterschieden. Andererseits ergaben sich keine Anhaltspunkte dafür, daß sich die individuellen Wachstumskurven im Verlauf der Studie bedeutsam voneinander unterschieden. Die Reliabilitäten der geschätzten Eingangs- und Veränderungswerte fielen dementsprechend aus: Während sich für die Reliabilität der Eingangsniveau-Schätzungen ein Koeffizient von .74 ermitteln ließ, betrug die Reliabilität der individuellen Wachstums- bzw. Veränderungsdaten lediglich .07.

Über HLM läßt sich schließlich auch der Frage nachgehen, welche Hintergrundvariablen die beobachtete Parametervarianz in den Eingangs- und Veränderungskennwerten erklären können. Die Ergebnisse der Varianzanalysen bewogen uns dazu, das Geschlecht und das Alter der Kinder als Hintergrundmerkmale einzubeziehen. Weiterhin erschien die sprachgebundene Intelligenz der Kinder theoretisch wichtig zu sein, da einige Subtests des Motoriktests MOT 4-6 relativ detaillierte Instruktionen erfordern, die für Kindergartenkinder nicht immer leicht zu verstehen sind. Diese Vermutung wurde durch signifikante bivariate Korrelationen zwischen der sprachlichen Intelligenz und den Ergebnissen im MOT 4-6 insbesondere zu Beginn der Studie gestützt.

Tabelle 1 gibt wieder, in welchem Ausmaß das Geschlecht, die Altersgruppenzugehörigkeit und die sprachliche Intelligenz individuelle Unterschiede in den motorischen Entwicklungs-Kennwerten erklären konnten. Die Analyse wurde dabei separat für das Eingangsniveau und die Veränderungsrate durchgeführt. Wie sich aus Tabelle 1 ablesen läßt, konnten über alle drei erfaßten Hintergrundmerkmale individuelle Unterschiede in den Eingangskennwerten erklärt werden, was in den signifikanten Z-Werten zum Ausdruck kommt (die Parameter-Koeffizienten in der ersten Spalte lassen sich als übliche Regressionkoeffizienten deuten). Demgegenüber erwies sich lediglich das Geschlecht im Hinblick auf die Erklärung individueller Veränderungsrate als bedeutsam: Die Veränderungsrate fielen für Mädchen größer als für Jungen aus.

Es bleibt die Frage offen, ob durch den Einbezug der drei gewählten Hintergrundvariablen die gesamte Parametervarianz in den Eingangskennwerten und Veränderungsrate aufgeklärt werden konnte. Für die beobachtete Varianz in den Eingangsniveau läßt sich dies nicht behaupten, wie der immer noch hochsignifikante Chi-Quadrat-Koeffizient von 673.74 ( $p < .001$ ) andeutet. Es gibt also noch andere bedeutsame Merkmale, die in unserer Untersuchung nicht erfaßt wurden. Im Vergleich dazu erwies sich der entsprechende Chi-Quadrat-Wert für die Wachstumskurven-Parameter als nicht mehr signifikant, wenn der Einfluß der drei Hintergrundmerkmale berücksichtigt worden war (142.97,  $p = .07$ ). Die Hypothese, daß diese drei Variablen die vorgefundene Varianz in den Wachstums-Parametern ausreichend erklären können, kann aufgrund dieser Ergebnisse nicht zurückgewiesen werden.

Zur Illustration der Befunde zur motorischen Entwicklung wurde das theoretisch einfachste Wachstumskurven-Modell gewählt, das lineare Entwicklungsverläufe unterstellte. Wie schon oben angedeutet, hatte sich dieses Modell im direkten Vergleich mit Polynomen höherer Ordnung (z.B. quadratischen oder kubischen Wachstumsmo-

dellen) als aussagekräftiger erwiesen. Das lineare Modell konnte die beobachtete Datenstruktur- und Varianz am besten erklären.

Tab. 1: Effekte von Geschlecht, Altersgruppe und verbalem IQ auf die motorischen Wachstumsparameter (nach Schneider 1993a)

Effekte	Strukturkoeffizient	Standardfehler	z	p
Effekt von Geschlecht auf Eingangsniveau	1.64	0.54	3.01	< .01
Effekt von Altersgruppe auf Eingangsniveau	2.35	0.55	4.27	< .01
Effekt des Verbal-IQs auf Eingangsniveau	-0.02	0.01	-2.46	< .05
Effekt von Geschlecht auf Veränderungsrate	0.80	0.36	2.24	< .05
Effekt von Altersgruppe auf Veränderungsrate	-0.59	0.36	-1.61	.11
Effekt des Verbal-IQs auf Veränderungsrate	-0.002	0.005	-0.35	.72

Insgesamt gesehen kann HLM als ein flexibles und elegantes statistisches Verfahren gelten, mit dem sich einmal die Reliabilität von Eingangskennwerten und Veränderungsrate, zum anderen die Korrelation zwischen Eingangsniveaus und Wachstumskurven schätzen läßt. Weiterhin eignet sich die Prozedur dazu, zukünftige indivi-

duelle Veränderungsraten vorherzusagen (vgl. *Bryk/Raudenbush* 1987 für spezifische technische Details). Obwohl das Verfahren multiple Meßzeitpunkte und ausreichend große Stichproben voraussetzt, kann es insofern als flexibel eingestuft werden, als Anzahl und Verteilung der Meßzeitpunkte über die einzelnen Personen hinweg variieren kann. Ein spezifischer Vorteil von HLM scheint darin zu bestehen, daß es die jeweiligen Stärken des verfügbaren Datensatzes optimal nutzt: Für den Fall, daß die individuellen Wachstumskurven reliabel sind, werden diese besonders gewichtet. Sind die Schätzungen der individuellen Wachstumskurven dagegen weniger reliabel, wird das Modell stärker auf Informationen über die durchschnittlichen Wachstumskurven und verfügbare Hintergrundmerkmale zurückgreifen.

### 3.2 *Zur längsschnittlichen Analyse individueller Unterschiede.*

Längsschnittstudien, die sich vorrangig mit individuellen Unterschieden zwischen Personen befassen, konzentrieren sich auf Fragen der Stabilität von Merkmalen bzw. auf Fragen der Vorhersagbarkeit (vgl. *Schneider* 1991). Im Hinblick auf den Bereich der motorischen Entwicklung interessiert hier beispielsweise die Frage, ob die Rangfolge einer Population von Kindern in der Entwicklung feinmotorischer Fertigkeiten über die Zeit hinweg als invariant anzunehmen ist, was auf interindividuell vergleichbare Entwicklungstempi hindeutet, oder ob sich im vorgegebenen Zeitintervall nur mäßige Stabilitäten der Rangfolge zeigen. Ein weiteres typisches Beispiel betrifft die Frage, inwieweit frühe Risikofaktoren bzw. Verzögerungen in der Ausbildung motorischer Fertigkeiten langfristig negative Auswirkungen auf diesen Kompetenzbereich haben können. Lange Zeit griff man ausschließlich auf klassische regressionsstatistische Ansätze zurück, um solche Fragen zu beantworten. Die längsschnittlichen Datenauswertungen basierten dabei auf Korrelations- oder Kovarianzmatrizen, über die sich die Konsistenz individueller Unterschiede über die Zeit erfassen läßt. Auch Weiterführungen des traditionellen Ansatzes gehen in der Regel von dieser Datenbasis aus, erlauben jedoch zusätzliche Aussagen. Im folgenden werden zunächst die traditionellen Modelle kurz skizziert, bevor auf methodologische Weiterentwicklungen eingegangen wird.

#### 3.2.1 *Probleme klassischer Ansätze*

##### *Multiple Regressionsanalysen*

Wie schon oben angedeutet, stellt die multiple Regressionsanalyse die Standardprozedur für alle Fragestellungen dar, die sich mit der Analyse von interindividuellen Unterschieden in intraindividuellen Veränderungen beschäftigen. Sie wird auch dann bevorzugt herangezogen, wenn es um die Vorhersage von individuellen Unterschieden in einem bestimmten Bereich durch individuelle Unterschiede

in einem anderen Bereich geht (etwa die Vorhersage von Rechtschreibleistungen durch frühe feinmotorische koordinative Fähigkeiten). Das Verfahren erscheint also prinzipiell für die längsschnittliche Analyse individueller Unterschiede im motorischen Bereich geeignet zu sein. Im klassischen Modell der Regressionsanalyse wird der Wert einer abhängigen Variablen (dem Kriterium) durch die gewichtete Summe von  $N$  unabhängigen Variablen (den Prädiktoren) so zu schätzen versucht, daß die Summe der quadrierten Abweichungen, also der Differenzen zwischen dem aktuellen Kriteriumswert und seinem Schätzwert ein Minimum ergibt (vgl. *Pedhazur* 1982; *Bortz* 1989).

Obwohl sich Regressionsmodelle in den Volks- und Wirtschaftswissenschaften durchaus bewährt haben und auch in den Sozialwissenschaften sehr populär geworden sind, ergeben sich in vielen Anwendungsfällen der Verhaltens- und Sozialwissenschaften methodische Probleme. Ein gravierendes Dilemma besteht hier beispielsweise in Meßfehler- und Multikollinearitätsproblemen. Was das Meßfehlerproblem angeht, so wird im Modell lediglich das Kriterium als (fehlerbehaftete) Zufallsvariable angenommen, während für die Prädiktoren fehlerfreie Messungen unterstellt werden. Diese Annahme ist in den Sozialwissenschaften nur äußerst selten erfüllt. Ein zweites Problem besteht darin, daß Prädiktoren in der Regel nicht unabhängig voneinander sind (wie im Modell unterstellt), sondern meist beträchtlich miteinander korrelieren. Sind die Interkorrelationen zwischen den Prädiktoren insgesamt hoch, spricht man von Multikollinearität. Es resultieren dann ernsthafte Probleme für die Schätzung der Modellparameter (vgl. *Pedhazur* 1982; *Schneider* 1994).

Ein letztes methodisches Problem des Regressionsansatzes ist darin zu sehen, daß nichts über den Zusammenhang zwischen einzelnen Prädiktoren ausgesagt werden kann: Sie weisen im Modell allesamt den gleichen Erklärungsstatus auf.

#### *Pfadanalysen mit beobachteten Variablen*

Diese Weiterentwicklung des multiplen Regressionsansatzes bietet die Möglichkeit, auch im Bereich der Prädiktoren weiter zwischen unabhängigen (exogenen) und abhängigen (endogenen) Variablen zu differenzieren. Die Pfadanalyse bietet also die Möglichkeit, komplexe Zusammenhänge zwischen Prädiktoren in ihrer Einwirkung auf ein gegebenes Kriterium genauer zu überprüfen. Wenn es auch damit gegenüber multiplen Regressionsanalysen eine deutliche Verbesserung darstellt, weist das Verfahren dennoch einige schon für das Regressionsmodell diskutierte Probleme auf. So geht auch die Pfadanalyse von der unrealistischen Vorannahme aus, daß die Merkmale meßfehlerfrei erhoben wurden, und daß die Residuen von unterschiedlichen Gleichungen voneinander unabhängig sind. Gerade letzteres trifft kaum für Längsschnittstudien zu, in denen die gleichen Personen mehrfach mit den gleichen Instru-

menten getestet werden. Aus diesem Grund scheinen für die Analyse längsschnittlicher Daten Strukturgleichungsmodelle mit latenten Variablen ungleich besser geeignet zu sein.

### 3.2.2 Möglichkeiten von Kausalmodellen mit latenten Variablen

Im Unterschied zu dem oben beschriebenen pfadanalytischen Ansatz differenziert diese Version von Strukturgleichungsmodellen zwischen latenten (nicht beobachteten) und gemessenen Merkmalen. Strukturmodelle dieses Typs können als Kombination faktoren- und regressionsanalytischer Ansätze aufgefaßt werden: Während über ein (faktorenanalytisches) Meßmodell die Beziehung zwischen den gemessenen Variablen und den latenten Konstrukten definiert wird, dient das Strukturmodell dazu, die Abhängigkeiten zwischen den latenten Variablen festzulegen. Das eigentliche Kausalmodell wird also auf der Ebene der latenten Variablen spezifiziert, die den Forscher in der Regel auch stärker interessiert als die der zugrundeliegenden gemessenen Merkmale.

Die Vorzüge und Möglichkeiten unterschiedlicher Versionen von Strukturgleichungsmodellen mit latenten Variablen sind in einschlägigen Standardwerken genauer beschrieben (vgl. *Bentler* 1985; *Möbus/Schneider* 1986; *Jöreskog/Sörbom* 1989; *Lohmöller* 1989), auf die an dieser Stelle verwiesen wird. Im folgenden soll es darum gehen, die Vorteile dieser Ansätze für die längsschnittliche Analyse von Motorikdaten zu veranschaulichen.

#### *Illustrationsbeispiel 1: Zusammenhang zwischen intellektueller und motorischer Entwicklung*

Die schon oben näher beschriebene LOGIK-Studie scheint gut dazu geeignet, die Einsatzmöglichkeiten von Strukturgleichungsmodellen mit latenten Variablen für die Überprüfung der Frage zu demonstrieren, wie sich der Zusammenhang zwischen motorischen und Intelligenzmerkmalen über die Kindergartenzeit hinweg verändert. In Anlehnung an *Bös* (1987) wurde unterstellt, daß feinmotorische Aufgaben (z.B. Balancier-Übungen) komplexe Tätigkeiten darstellen, die nicht nur motorische Regulationen sondern auch kognitive Informationsverarbeitungsfähigkeit erfordern. Eine genauere Beschreibung des Untersuchungsansatzes und seiner Realisierung findet sich bei *Schneider* (1993a). Hier sei nur vermerkt, daß zwei unterschiedliche theoretische Ansätze kombiniert wurden: Die Entwicklungstrend-Hypothese geht davon aus, daß der ursprünglich enge Zusammenhang zwischen (fein-)motorischen Kompetenzen, wie sie im MOT 4-6 erfaßt werden, und der Intelligenz von Kindergartenkindern sich im Lauf der Zeit verringert. Danach soll die Fähigkeit zur Kontrolle komplexer Bewegungsabläufe dann nicht mehr von mentalen Fähigkeitsmerk-

malen abhängen, wenn die motorischen Handlungsabläufe zunehmend mehr automatisiert werden.

Die Hypothese reziproker Kausalität besagt demgegenüber, daß sich über die Kindergartenperiode hinweg gegenseitige (reziproke) Beeinflussungen der beiden untersuchten Konstrukte nachweisen lassen sollten. Im Hinblick auf die "kausale Prädominanz" wurde vermutet, daß die Einflüsse von früher Intelligenz auf spätere feinmotorische Fertigkeiten größer sein sollten als die von früher feinmotorischer Fertigkeit auf nachfolgende Intelligenz. In Übereinstimmung mit der Entwicklungstrend-Hypothese wurde angenommen, daß die gegenseitigen Beeinflussungen beider Bereiche gegen Ende der Kindergartenzeit vernachlässigbar sein sollten.

Zur Schätzung und Testung des Modells wurde das Strukturgleichungssystem EQS (Equations; vgl. *Bentler* 1985) herangezogen, das sich auch dann verwenden läßt, wenn die Normalverteilung der Variablen (wie im vorliegenden Fall) nicht immer gegeben ist. Abbildung 2 zeigt das Ergebnis der Modell-Schätzung, wobei aus Gründen der Übersichtlichkeit lediglich das Strukturmodell, d.h. das Beziehungsgeflecht zwischen den latenten Variablen wiedergegeben ist. Die Pfeile geben Strukturkoeffizienten an, sind also im Sinne von Regressionskoeffizienten zu interpretieren.

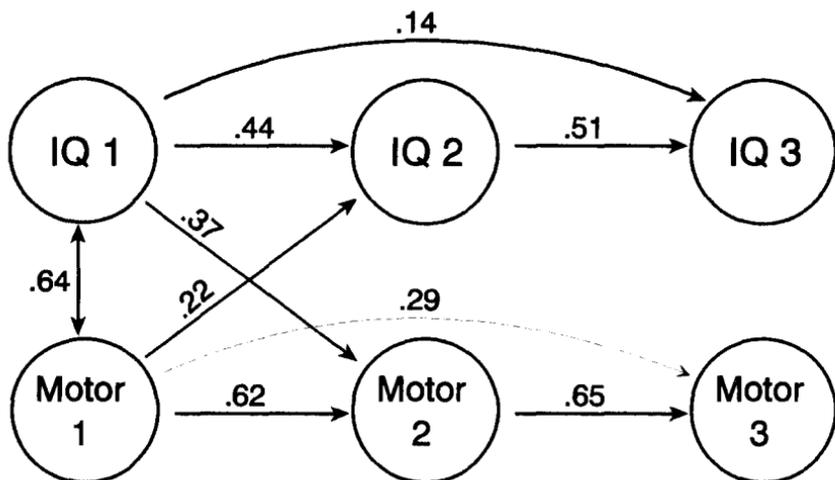


Abb. 2: Strukturgleichungsmodell für die Motorik- und Intelligenzdaten der LOGIK-Studie (nach *Schneider* 1993a)

Wie aus Abbildung 2 hervorgeht, ist der Effekt der Intelligenz zum ersten Meßzeitpunkt (also im Alter von 4 Jahren) auf die motorische Koordinationsfähigkeit zum zweiten Meßzeitpunkt (ein Jahr später) deutlich höher als der korrespondierende Effekt der frühen motorischen Fähigkeit auf die ein Jahr später erhobene Intelligenz. Dieser Befund paßt gut zur hier formulierten Hypothese der reziproken Kausalität. Die Ergebnisse bestätigen weiterhin die Annahme der Entwicklungstrend-Hypothese, da die Korrelationen zwischen den motorischen und intellektuellen Fähigkeiten über die Zeit hinweg kontinuierlich abnahmen (von  $r = .64$  zum ersten Meßzeitpunkt auf  $r = .25$  im letzten Kindergartenjahr). Das in Abbildung 2 wiedergegebene Modell erwies sich mit den Daten als gut kompatibel und ließ sich auch durch weitere Modifikationsversuche nicht mehr verbessern.

Was die Stabilität individueller Unterschiede über die Zeit hinweg betrifft, so ergaben sich für beide Konstrukte mittelhohe Retest-Korrelationen, die im Bereich von  $.45$  bis  $.65$  variierten. Daraus läßt sich schließen, daß die individuellen Unterschiede in der Intelligenz wie auch in der motorischen Koordinationsfähigkeit für den betroffenen Altersbereich (4 bis 6 Jahre) noch nicht zeitstabil sind. Wenn auch die Leistungsranfolge zu unterschiedlichen Zeitpunkten im wesentlichen ähnlich ist, gibt es in beiden Bereichen doch immer noch einige Verschiebungen in der Platzierung.

*Illustrationsbeispiel 2: Die Bedeutung allgemeiner motorischer Fähigkeiten für die Vorhersage bereichsspezifischer Fertigkeiten*

Anhand dieses Beispiels soll demonstriert werden, wie sich Strukturgleichungsmodelle mit latenten Variablen für Fragestellungen heranziehen lassen, bei denen es um die Prognose spezifischer motorischer Fertigkeiten zu einem späteren Zeitpunkt aufgrund früher erhobener allgemeiner motorischer Fähigkeiten und anderer Hintergrundmerkmale geht. Eine genauere Darstellung des Untersuchungsansatzes findet sich bei *Schneider et al.* (1993).

Zum besseren Verständnis des zugrundeliegenden Ansatzes sei darauf hingewiesen, daß es hierbei um die Reanalyse der Tennistalent-Studie von *Rieder et al.* (1983) ging, bei der über einen Untersuchungszeitraum von 5 Jahren insgesamt etwas mehr als 100 jugendliche Tennistalente (Mitglieder des C-Kaders) im Hinblick auf ihre motorische und psychologische Entwicklung hin untersucht wurden. Die im Projekt erhobenen Merkmale reichten von Hintergrundvariablen wie dem Engagement der Eltern oder körperlichen/psychologischen Auffälligkeiten der Tennistalente über basale motorische Fähigkeiten (Ausdauer, Schnelligkeit) zu tennisspezifischen Fertigkeiten (Zielschläge, Rahmen-Schläger-Wechselstest).

In der Reanalyse von *Schneider et al.* (1993) wurde analog zu neueren kognitionspsychologischen Ansätzen zur Entwicklung von Expertise ein Strukturgleichungsmodelle

dell mit latenten Variablen spezifiziert. In diesem Modell wurde davon ausgegangen, daß Hintergrundmerkmale wie etwa das Ausmaß der elterlichen Unterstützung, die spezifische Motivation der Jugendlichen (ausgedrückt durch Ergebnisse in Motivationstests und über die Präferenz des Tennis gegenüber schulischen Aktivitäten), und die Konzentrationsfähigkeit direkte und indirekte Einflüsse auf die Entwicklung tennisspezifischer Fertigkeiten haben, wobei indirekte Einflüsse über basale motorische Fähigkeiten vermittelt werden. In Analogie zu den kognitiven Expertisemodellen wurde weiterhin angenommen, daß basale motorische Fähigkeiten im Vergleich zu tennisspezifischen Fertigkeiten einen deutlich geringeren Effekt auf die spätere Leistung im Tennis (operationalisiert durch den Rangplatz zu unterschiedlichen Zeitpunkten) haben sollten.

Das zugehörige Strukturmodell zur Vorhersage der Rangplätze in den Jahren 1982 und 1989 ist in Abbildung 3 wiedergegeben. Wie schon bei Abbildung 2 wird aus Gründen der Übersichtlichkeit nur auf die Beziehungsstruktur zwischen den latenten Variablen eingegangen. Da die zugrundeliegende Personenstichprobe vergleichsweise klein ausfiel, wurde zur Schätzung des Modells das Verfahren LVPLS (Latent Variables Partial Least Squares) von *Lohmöller* (1989) verwendet, das auch in solchen Situationen einsetzbar ist.

Wie aus Abbildung 3 hervorgeht, ließen sich die oben formulierten Annahmen im wesentlichen bestätigen. So zeigten insbesondere Merkmale wie der Zeitaufwand der Eltern und die frühe Entscheidung für das Tennis bedeutsame direkte Effekte auf den späteren Rangplatz im Tennis. Damit scheinen individuelle Unterschiede in der Betreuung und im Trainingsaufwand schon zu einem relativ frühen Zeitpunkt langfristige Auswirkungen auf spätere Karriereverläufe zu haben. Individuelle Unterschiede in der Konzentration, Motivation und Trainingsintensität beeinflussen zusammen mit den allgemeinen motorischen Fähigkeiten die tennisspezifischen Fertigkeiten, wobei der direkte Effekt der motorischen Fähigkeiten vergleichsweise gering ausfällt. In Übereinstimmung mit der theoretischen Vorhersage erklären Hintergrundmerkmale wie die elterliche Betreuung und die Trainingsintensität zusammen mit den tennisspezifischen Fertigkeiten den Löwenanteil der Varianz im Rangplatz-Kriterium. Auffallend bleibt schließlich, daß sich aus der Kenntnis der Rangplätze im Jahr 1982 die 7 Jahre später erzielten Rangplätze relativ gut vorhersagen lassen (die entsprechende Korrelation beträgt .70). Daraus läßt sich ableiten, daß die individuellen Unterschiede in der "Tennis-Expertise" schon im frühen Jugendalter recht stabil sind und es in der Folge auch bleiben.

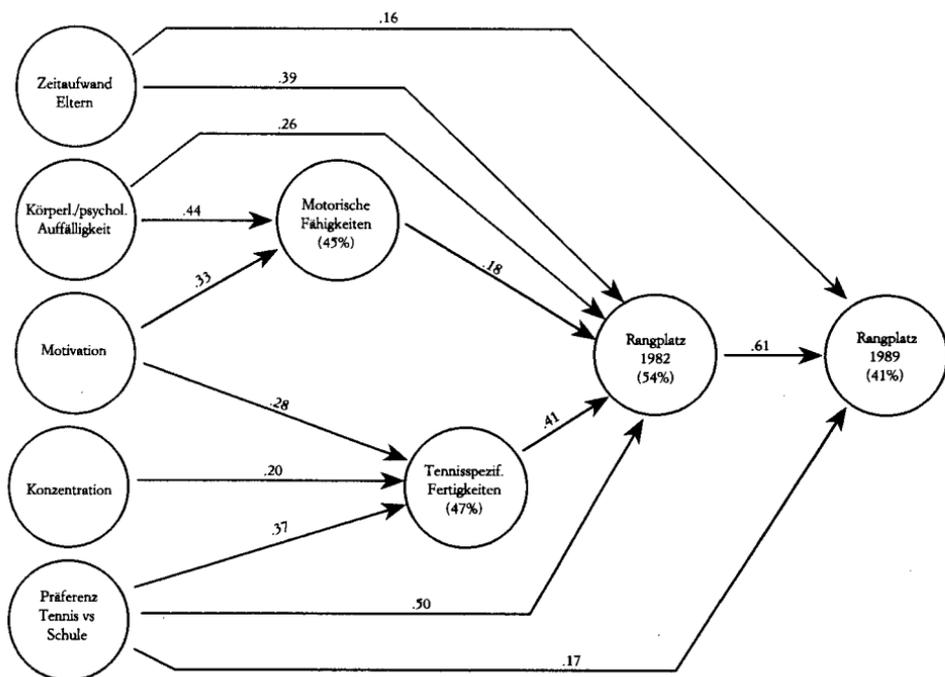


Abb. 3: Kausalmodell zur Leistungsvorhersage im Tennis (nach *Schneider et al.* 1993)

### 3.2.3 *Schlußfolgerungen*

Die beiden Beispiele sollten verdeutlicht haben, daß Strukturgleichungsmodelle mit latenten Variablen im Vergleich zu traditionellen regressionsanalytischen Verfahren eine Reihe von Vorzügen aufweisen. So wird der Forscher veranlaßt, ein theoretisch begründbares Modell zu formulieren, das sich anhand von empirischen Daten schätzen läßt. Es kann weiterhin zwischen der Reliabilität von gemessenen Variablen und der Stabilität struktureller Beziehungen unterschieden und die Anpassung des Modells an die zugrundeliegenden Daten genau überprüft werden (vgl. *Bentler* 1985; *Jöreskog/Sörbom* 1989). Unterschiedliche Vorstellungen über längsschnittliche Verlaufsmuster lassen sich schließlich leicht über alternative Modellspezifikationen gegeneinander testen.

Diesen vielfältigen Vorzügen stehen einige noch ungelöste Probleme (etwa bei der Modellbewertung) gegenüber, die hier nur erwähnt und an anderer Stelle (z.B. *Schneider* 1991; 1993b) näher ausgeführt werden. Ungeachtet dieser Probleme bleibt jedoch festzuhalten, daß die beschriebenen Strukturgleichungsmodelle mit latenten Variablen eine äußerst wertvolle Ergänzung des sozialwissenschaftlichen Methoden-Arsenals darstellen und insbesondere für die Analyse von interindividuellen Unterschieden in intraindividuellen Entwicklungsverläufen unersetzbar geworden sind.