

## Mythen in der Jugend- und Schulleistungsforschung: Möglichkeiten und Grenzen populationsorientierter Untersuchungen

### Myths in Research on Youth and Academic Achievement: Benefits and Limitations of Population- Based Studies

**Zusammenfassung:** Ziele wissenschaftlicher Beschreibung und Erklärung werden auch in der Jugend- und Schulleistungsforschung in der Regel anhand populationsorientierter Untersuchungen verfolgt. In diesem Beitrag werden Möglichkeiten und Begrenzungen quer- und längsschnittlicher Untersuchungen an Populationsstichproben und ein alternatives Vorgehen diskutiert. Dargestellt wird zunächst, dass Querschnittuntersuchungen Aufschlüsse zum Entwicklungs- und Leistungsstand von Jugendlichen und zu Verknüpfungen von Entwicklungsstand und anderen Variablen liefern, aber im Gegensatz zu manchen Rezipientenerwartungen keine Schlüsse auf Entwicklungsprozesse und Ursache-Wirkungs-Verhältnisse zulassen. In der Literatur wird deshalb empfohlen, Längsschnittstudien durchzuführen. Bei der Durchführung solcher Studien aber wird in der Regel zu wenig beachtet, dass übliche stichprobenbasierte, nomothetisch orientierte Längsschnittsdesigns zwar Aussagen zu Entwicklungen von aggregierten Werten in Stichproben und zu Bedingungen von Entwicklungsunterschieden zwischen Jugendlichen ermöglichen, hingegen ebenfalls keine Evidenz zu Entwicklungsprozessen und den Bedingungen solcher Prozesse bei einzelnen Jugendlichen oder Populationen von einzelnen Jugendlichen liefern. Es wird gezeigt, dass dies gravierende Fehlschlüsse zur Folge haben kann. Als Alternative bieten sich Strategien an, die idiographische und nomothetische Perspektiven so miteinander verknüpfen, dass sowohl Befundvalidität für Einzelfälle wie auch nomothetische Generalisierbarkeit über Individuen hinweg sichergestellt werden.

**Abstract:** Studies on youth and academic achievement take no exception in using population-oriented designs for purposes of scientific description and explanation. In this article, we discuss benefits and limitations of cross-sectional and longitudinal studies using samples of adolescents, as well as an alternative strategy. Cross-sectional surveys produce evidence on the developmental status of adolescents, and on associations between developmental status and other variables. However, in contrast to public expectations, they do not allow any conclusions about developmental processes and the causes and effects of such processes. Therefore, the use of sample-based, nomothetically oriented longitudinal designs is recommended in the methodological literature. However, the limitations of such designs are rarely noticed. Sample-based longitudinal studies typically produce evidence on the development of aggregated sample statistics over time, and on antecedents of inter-individual developmental differences. In contrast to usual practices of interpreting the findings of such studies, however, they do not allow any inferences on developmental processes and their antecedents in individual adolescents or in populations of individual adolescents. Therefore, alternative strategies are recommended which combine idiographic and nomothetic perspectives in such a way that the resulting evidence is valid both for individual adolescents, and for populations of individuals.

Beschreibung und Erklärung zählen auch in der Jugendforschung zu den wesentlichen Zielen wissenschaftlicher Tätigkeit. Für die Erforschung des Jugendalters ist es zentral, Beschreibungen des Entwicklungsstandes von Jugendlichen zu liefern sowie zugrunde liegende Entwicklungsprozesse und die Bedingungen und Wirkungen solcher Prozesse zu analysieren. Dies gilt für Analysen der Kompetenz- und Leistungsentwicklung, der Persönlichkeits- und Sozialentwicklung und der politischen Entwicklung gleichermaßen.

Welche Untersuchungsformen sind geeignet, solche Aufschlüsse zu liefern? Sieht man die Methodenliteratur durch, so zeigt sich, dass heute ein reichhaltiges Arsenal an Werkzeugen zur Verfügung steht, um auch komplexen Fragestellungen multivariater, dynamischer und in mehreren Ebenen organisierter Entwicklungs- und Bedingungsprozesse gerecht zu werden (z.B. Bryk/Raudenbush 1987; van Eye 1990; Lütkepohl 1991; Nesselrode/Cattell 1988). Eine andere Frage aber ist, inwieweit die Jugendforschung auf diese Werkzeuge in hinreichendem Maße zurückgreift und sie in gegenstandsangemessener Weise verwendet. In diesem Beitrag soll dargestellt werden, dass diese Frage keineswegs trivial ist. Dies gilt selbst für methodisch sophistierte Untersuchungen, die sich mehr-ebenen- und strukturgleichungsanalytischer Verfahren bedienen, da solche Verfahren zu Fehlschlüssen der Inferenz von Stichproben auf Individuen verführen können. Offenbar geschieht dies so häufig, dass ganze Teildisziplinen der Jugend- und Entwicklungsforschung auch heute noch durch empirisch kaum begründbare Schlussfolgerungen gekennzeichnet sind.

Diskutiert werden soll deshalb, welche Möglichkeiten populationsorientierte Querschnitt- und Längsschnittstudien bieten, die heute die empirische Jugendforschung dominieren, welche Grenzen aber gleichzeitig einer validen Interpretation der resultierenden Befunde gezogen sind. Anschließend wird auf Alternativen zu herkömmlichen Vorgehensweisen eingegangen, welche einige der Grenzen dieser Vorgehensweisen zu überwinden vermögen. Dabei konzentriert sich die Diskussion auf Untersuchungsstrategien quantifizierender Art, obschon die Argumentation grundsätzlich auch auf qualitativ orientierte Erhebungen übertragbar wäre. Ein besonderes Augenmerk soll auf vergleichende Evaluationsstudien zu Kompetenzen und schulbezogenen Leistungen von Jugendlichen gerichtet werden, da einige Studien dieses Typus sich heute – im Unterschied zu den meisten anderen Unternehmungen der deutschen Jugendforschung – international vergleichbarer, repräsentativer Stichproben bedienen und nationalen Begrenzungen insoweit nicht unterliegen (dies unterscheidet sie u.a. auch von den Shell-Jugendstudien). Zwei Studien dieser Art sind TIMSS (Third International Mathematics and Science Study; vgl. Baumert/Lehmann 1997) und die gegenwärtig laufenden OECD-Studien PISA (Programme for International Student Assessment), die repräsentative Befunde zum Stand der Kompetenzentwicklung von Jugendlichen

liefern und heute auch in Deutschland wesentliche Bereiche der öffentlichen Diskussion zu Jugend- und Bildungsproblemen dominieren.

## Populationsorientierte Querschnittuntersuchungen wie TIMSS, PISA und die Shell-Jugendstudien: Möglichkeiten und Grenzen

Populationsorientierte Querschnittuntersuchungen an Stichproben von Jugendlichen liefern zunächst deskriptive Aufschlüsse zum *Entwicklungsstand* in der jeweiligen Population. Voraussetzung hierfür ist, dass die Stichproben hinreichend groß und repräsentativ genug angelegt sind, um solche Schlüsse zuzulassen. Inhaltlich kann es sich dabei insbesondere um Schlüsse zu Erscheinungsformen, Strukturen und Verteilungen der erhobenen Entwicklungsvariablen handeln (also z.B. zu Inhalten, Strukturen, durchschnittlichen Ausprägungen und Variationen von Einstellungen zu politischer Gewalt).

TIMSS und PISA zählen zu den wenigen Jugendstudien in Deutschland, welche die genannten Stichprobenvoraussetzungen mit hinreichender Näherung erfüllen (vgl. zu den Stichprobenproblemen von Jugendstudien Hoffmann-Lange 2001). Eine erste zentrale Leistung dieser Studien liegt in einer Deskription nationaler und internationaler Variation schulbezogener Leistungen, der populationsorientierte Validität zuzusprechen ist. Schließen lässt sich, dass die erfassten Kompetenzen und Leistungen von Jugendlichen der jeweiligen Altersstufen (insbesondere 15-Jährige bzw. Schüler neunter Klassen) international wie national erheblich variieren, wobei das durchschnittliche Abschneiden deutscher Jugendlicher eher dem Mittelfeld wirtschaftlich entwickelter Nationen zuzuordnen ist – in den Augen von Bildungspolitik und Öffentlichkeit ein enttäuschender Befund, dessen Stellenwert allerdings kaum zu überschätzen ist, da er wesentliche Impulse zur Eröffnung einer neuen Bildungsdebatte in Deutschland geliefert hat. In analoger Weise gilt für die ebenfalls repräsentativ angelegten deutschen Shell-Studien, dass sie eine Deskription des Entwicklungsstandes von Jugendlichen liefern, wobei in diesen Studien Aspekte der Persönlichkeitsentwicklung sowie der sozialen und politischen Entwicklung im Vordergrund stehen.

Über eine Deskription des Entwicklungsstandes bezüglich einzelner Entwicklungsvariablen hinaus können populationsorientierte Querschnittuntersuchungen deskriptive Aufschlüsse zu den *korrelativen Verknüpfungen* unterschiedlicher Variablen liefern. So lassen sich z.B. aus TIMSS und PISA populationsorientierte Schlüsse zu Beziehungen zwischen Leistungsvariablen und ihren mutmaßlichen Bedingungen und Wirkungen ziehen (wie z.B. der Intelligenz und Motivation von Schülern, der Qualität des schulischen Unterrichts oder der sozioökonomischen Stellung des Elternhauses). Bei solchen korrelati-

ven Verknüpfungen handelt es sich um Beziehungen interindividuell-korrelativer Art, also um Beziehungen von Leistungsunterschieden zwischen Jugendlichen zu Unterschieden in Intelligenz, Schulunterricht, Elternhaus etc. Bei ihrer Interpretation ist zu bedenken, dass sich aus ihnen nicht auf entsprechende Verknüpfungen beim einzelnen Individuum zurückschließen lässt (auf diesen kritischen Punkt wird im nächsten Abschnitt ausführlicher eingegangen).

Kompetenz- und Leistungsstudien an Jugendlichen liefern damit ein Systemmonitoring des Bildungswesens, das eine Deskription von Output-, Prozess- und Kontextvariablen und darüber hinaus Evidenz zu korrelativen Beziehungen zwischen diesen Variablen liefert. Analoges gilt für querschnittliche Untersuchungen zu anderen Entwicklungsbereichen, also z.B. die Shell-Jugendstudien. Häufig aber werden an Querschnittstudien Erwartungen gerichtet, die über solche Deskriptionsleistungen hinausgehen. So wird z.B. mit der Durchführung von Studien wie TIMSS und PISA gerne unterstellt, dass die erfassten schulischen Lernleistungen von Jugendlichen relevant seien für Leistungen jenseits der Schulzeit: Schulische Lernleistungen interessieren, weil ihnen ein prädiktiver Stellenwert für die zukünftige Leistungsfähigkeit in nachschulischer Ausbildung, Beruf und Wirtschaft zugeschrieben wird. Ferner wird häufig erwartet, dass aus korrelativen Verknüpfungen zwischen Leistungen und anderen Variablen auf die Ursache-Wirkungs-Beziehungen geschlossen werden könne, die dem Zustandekommen dieser Leistungen zugrunde liegen. So wird beispielsweise auf der Basis von TIMSS-Befunden häufig angenommen, dass Unterschiede des Mathematikunterrichts in verschiedenen Nationen ursächlich für korrelativ zugeordnete Leistungsunterschiede von Schülern dieser Nationen seien. Schließlich wird häufig auch erwartet, dass aus korrelativen Verknüpfungen dieser Art nicht nur auf Leistungsbedingungen, sondern auch auf Handlungsmöglichkeiten zur Leistungsoptimierung geschlossen werden könne.

Querschnittstudien aber können Erwartungen dieser Art grundsätzlich nicht erfüllen: Da die Erhebung unterschiedlicher Variablen in solchen Studien zeitlich nicht sequenziert ist, lassen sich Schlüsse zu Prognose-, Ursache- und Wirkungsbeziehungen zwischen diesen Variablen nicht ziehen. Querschnittlich-korrelative Zusammenhänge sind bekanntlich offen gegenüber alternativen kausalen Interpretationen (für einen bivariaten Korrelationszusammenhang zweier Variablen A und B gelten die folgenden Interpretationsmöglichkeiten: A bedingt B; B bedingt A; beide werden durch eine oder mehrere Variablen C bedingt; mehrere dieser Möglichkeiten treffen gleichzeitig zu). Verknüpfungen zwischen Unterrichtsmerkmalen und Mathematikleistungen beispielsweise können auf Unterrichtseffekte zurückgehen; es kann sich aber z.B. auch um Effekte von Schülerleistungen auf das Unterrichtsverhalten von Lehrern oder um gemeinsame Bedingtheit von Unterricht und Leistungen durch kulturspezifische Lehr-Lern-Orientierungen von Schülern und Lehrern handeln. Dies ist zwar methodisch trivial, scheint aber auch in der Fachdiskussion häufig nicht hinreichend beachtet zu werden.

## Populationsorientierte Längsschnittstudien: Der Königsweg der Jugendforschung?

Da Querschnittstudien Entwicklungsprozesse und Bedingungs-Wirkungs-Beziehungen nicht angemessen abzubilden vermögen, wird in der Methodoliteratur die Durchführung längsschnittlicher Studien zur Gewinnung von Entwicklungs- und Bedingungsaussagen empfohlen (vgl. Pekrun/Fend 1991). Bei Einbeziehung hinreichend repräsentativer Stichproben erlauben es solche Studien ähnlich wie repräsentativ angelegte Querschnittstudien, populationsvalide Schlüsse zu ziehen.

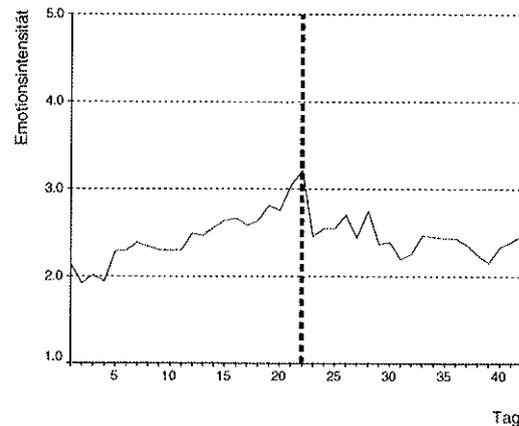
Entscheidend ist aber auch bei solchen populationsorientierten Längsschnittunternehmungen, Möglichkeiten und Grenzen realistisch einzuschätzen. Erwartet wird meist, dass solche Studien – eine entwicklungsangemessene Taktung von Messzeitpunkten und Operationalisierung von Variablen vorausgesetzt – tatsächlich in der Lage seien, Entwicklungsprozesse bei Jugendlichen abzubilden und die Ursachen solcher Prozesse aufzuklären. Werden populationsorientierte Längsschnittstudien solchen Erwartungen gerecht?

Eine nähere Analyse zeigt, dass dies nur bedingt der Fall ist. Zu diskutieren ist dies zunächst für die *Deskription von Entwicklungsprozessen*. Das typische Vorgehen zur Beschreibung der Veränderung einer Entwicklungsvariablen über die Zeit sieht in populationsorientierten Längsschnittstudien so aus, dass die Werte der einzelnen Individuen innerhalb der Messzeitpunkte des Längsschnitt durch Mittelwertbildung über Individuen aggregiert werden und anschließend der Verlauf der Mittelwerte über die Zeit analysiert wird. Was auf diese Weise gewonnen wird, ist eine Beschreibung der Entwicklung von Durchschnittswerten in der Stichprobe über die Zeit. Üblicherweise wird nun gefolgert, dass aus einer solchen Stichprobenbeschreibung auf die tatsächliche Entwicklung bei den einzelnen Jugendlichen der betreffenden Stichprobe bzw. Population zurückgeschlossen werden könne. Eine solche Folgerung aber ist methodisch nicht zulässig und kann inhaltlich zu falschen Schlüssen führen: Veränderungen von Durchschnittswerten können zu Veränderungen bei den Individuen der Stichprobe kongruent sein, müssen dies aber nicht (vgl. Rogosa 1988; Schmitz 2000). Hat man Pech, so gibt es kein einziges Individuum in der Stichprobe, dessen Entwicklung der Entwicklung der Durchschnittswerte entspricht.

Die Abbildungen 1 und 2 zeigen dies exemplarisch am Beispiel von Befunden aus einer Tagebuchstudie an  $N = 72$  jungen Erwachsenen im Examen (Hofmann 1997; Pekrun/Hofmann 1999). Untersucht wurde die Entwicklung von prüfungsbezogenem Erleben und Verhalten im Verlauf von sechs Wochen vor und während universitärer Abschlussprüfungen. Dargestellt sind die Veränderungen der Stichprobenmittelwerte der täglichen prüfungsbezogenen Angstwerte (Abbildung 1) einerseits und die Veränderungen der zugrunde liegenden Individualwerte (Abbildung 2) andererseits. Die Veränderung der

Durchschnittswerte folgt einem plausiblen Muster: leichter Anstieg der Werte bis zum Beginn des Prüfungszeitraums, leichter Abfall in den ersten drei Wochen während dieses Zeitraums. Dennoch ist dieses Verlaufsmuster keineswegs repräsentativ für die zugrunde liegenden Entwicklungen bei den einzelnen Examenskandidaten: Es handelt sich hier um eine Fülle von unterschiedlichen individuellen Entwicklungsverläufen, die von der stichprobenbezogenen Information kaum abgebildet werden. Dabei finden sich Individualentwicklungen, deren Typus der Entwicklung der Durchschnittswerte entspricht (Sukzession von An- und Abstieg), mit Abweichungen im Steigungswinkel oder in der Variation zwischen einzelnen Tagen. Bei anderen Teilnehmern der Stichprobe aber folgt die Entwicklung gänzlich anderen Mustern (z.B. gleichbleibende Angstwerte oder ausschließlich an- oder absteigende Werte).

Abbildung 1: Prüfungsangst: Durchschnittlicher Verlauf

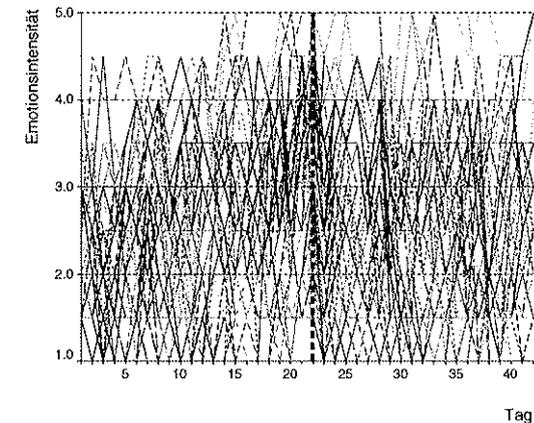


Es wäre also irreführend, hier anzunehmen, aus der Stichprobeninformation könne auf die tatsächlichen Veränderungen bei den Individuen der Stichprobe zurückgeschlossen werden. Vielmehr setzen valide Schlüsse auf die Individualentwicklung voraus, diese direkt zu analysieren. Ähnliches dürfte auch für Entwicklungsverläufe in vielen anderen Entwicklungsbereichen gelten, und zwar insbesondere dann, wenn diese Verläufe komplexen Entwicklungsbedingungen unterliegen, die in ihrer Zusammensetzung interindividuell und temporal variieren.

Über die Beschreibung von Entwicklungsprozessen hinaus haben Längsschnittstudien typischerweise zum Ziel, die solchen Prozessen zugrunde liegenden *Bedingungs-Wirkungs-Beziehungen* zu ermitteln. Ähnlich wie bei der Abbildung von Entwicklungsprozessen wird auch hier in der Regel so vorgegangen, dass über Individuen aggregierte Statistiken zur Analyse von

Bedingungsbeziehungen verwendet werden. Ein typisches Vorgehen beinhaltet, die synchronen und diachronen Korrelationen der jeweiligen Bedingungs- und Wirkungsvariablen in der Stichprobe innerhalb von Messzeitpunkten und über Messzeitpunkte hinweg zu berechnen und zueinander ins Verhältnis zu setzen, beispielsweise anhand von dynamisch-bedingungsanalytischen Strukturgleichungsmodellierungen (z.B. Modellierung mit LISREL). Ebenso wie im querschnittlichen Fall handelt es sich bei den verwendeten Stichprobenkorrelationen um interindividuelle Korrelationen, also jeweils um Verknüpfungen der interindividuellen Werteverteilung einer Variablen mit der interindividuellen Werteverteilung einer zweiten Variablen. Eine Korrelation dieser Art beinhaltet Aussagen über die Verknüpfungen von Personunterschieden auf der ersten Variablen mit Personunterschieden auf der zweiten Variablen; auf solchen Korrelationen basierende Bedingungsanalysen liefern Schlüsse zu den Bedingungsbeziehungen zwischen solchen Personunterschieden. Es handelt es sich hier also grundsätzlich um Verknüpfungs- und Bedingungsansagen differenziellpsychologischer Art (bzw. – bei Stichproben von überindividuellen Einheiten wie Schulen, Gemeinden oder Regionen – um Aussagen differenziellsoziologischer Art).

Abbildung 2: Prüfungsangst: Individuelle Verläufe

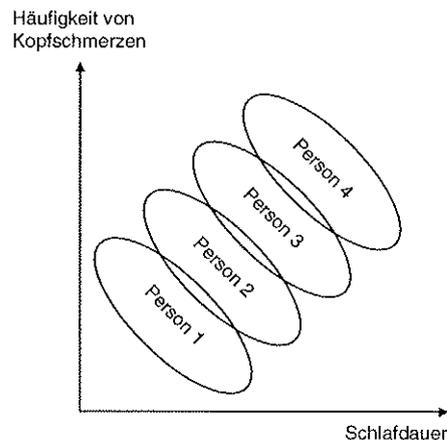


Nun ist auch in der Jugendforschung im Grundsatz gut bekannt, dass aus Beziehungen zwischen interindividuellen Verteilungen nicht auf individuelle, intrapsychische Funktionsmechanismen bei einzelnen Jugendlichen oder Populationen von einzelnen Jugendlichen geschlossen werden kann. Dennoch werden die Resultate entsprechender Korrelations- und Bedingungsanalysen allzu häufig so interpretiert, als seien solche Schlüsse möglich (hierzu Asendorpf 1995; Valsiner 1986). Ein klassisches Beispiel sind Analysen zu den

Bedingungsbeziehungen zwischen Fähigkeitsselbstkonzept und Leistungen bei Jugendlichen (vgl. Krapp 1997; Pekrun 1997). Bedingungsbeziehungen zwischen diesen beiden Variablen werden üblicherweise so untersucht, dass beide Variablen über die Zeit hinweg mehrfach erhoben werden und die syn- und diachronen korrelativen Beziehungen zwischen ihnen anhand von Bedingungsanalysen analysiert werden. Aus so ermittelten Effekten der Leistungs- auf die Selbstkonzeptvariable (oder umgekehrt) wird dann geschlossen, Leistung bedinge das Selbstkonzept (bzw. umgekehrt), aus fehlenden Effekten, dass kein Einfluss vorhanden sei (z.B. Marsh/Yeung 1997).

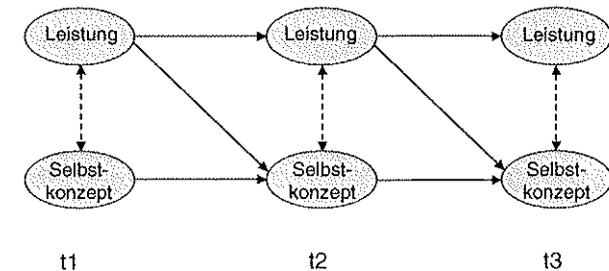
Eine Interpretation dieser Art aber kann zu gravierenden Fehlschlüssen führen. Grundsätzlich gilt nämlich, dass aus Beziehungen zwischen interindividuellen Verteilungen von Variablen in Stichproben von Personen nicht auf die Beziehungen zwischen diesen Variablen innerhalb von Personen geschlossen werden kann (vgl. Robinson 1950; Schmitz 2000). Das von B. Schmitz und E. Skinner zitierte medizinische Beispiel des Zusammenhangs von Schlafdauer und Migränehäufigkeit macht dies besonders deutlich (Schmitz/Skinner 1993; siehe die schematische Darstellung in Abbildung 3): Die interindividuelle Korrelation von Schlafdauer und Migränehäufigkeit ist typischerweise positiv. Hieraus könnte man nun folgern, langer Schlaf begünstige Migräne (oder auch umgekehrt). Diese Folgerung aber wäre irreführend: Wie die *intraindividuellen* Korrelationen von Schlafdauer und dem Auftreten von Migräne über die Zeit zeigen, kann – ganz im Gegenteil – eine *verkürzte* Schlafdauer mit Migräne verknüpft sein (alternativ finden sich intraindividuelle Nullkorrelationen für diese beiden Variablen).

Abbildung 3: Schlafdauer und Häufigkeit von Migränekopfschmerzen: Intraindividuelle vs. interindividuelle Korrelation



Solche Fälle einer fehlenden Übereinstimmung von inter- und intraindividuellen Zusammenhängen dürften auch im Bereich der Jugendforschung häufig sein. So ist z.B. ein typisches Resultat der zitierten leistungsbezogenen Selbstkonzeptforschung bei Jugendlichen, dass Effekte von Selbstkonzeptvariablen auf Leistungsvariablen in längsschnittlich-bedingungsanalytischen Untersuchungen gering bleiben und sich nicht signifikant von Null unterscheiden (vgl. Abbildung 4). In Übereinstimmung mit heute üblichen Interpretationen solcher Befunde könnte man geneigt sein, aus solchen fehlenden interindividuell-bedingungsanalytischen Effekten des Selbstkonzepts zu schließen, die Selbsteinschätzung von Jugendlichen habe keinen Einfluss auf ihre Leistungen. Dies aber wäre ganz offenkundig ein Fehlschluss. Anzunehmen ist vielmehr, dass Einschätzungen der eigenen Leistungsmöglichkeiten Tag für Tag zu entsprechenden Gefühls- und Motivationslagen beitragen, die ihrerseits Einfluss auf die individuellen Lern- und Leistungsanstrengungen nehmen; und zwar ganz unabhängig davon, ob sich dies auch – wie im interindividuell-bedingungsanalytischen Design untersucht – in einer Veränderung der Leistungsunterschiede zwischen verschiedenen Jugendlichen von einem zum nächsten Messzeitpunkt niederschlägt.

Abbildung 4: Leistung und Fähigkeitsselbstkonzept: Typisches Resultat von längsschnittlichen Bedingungsanalysen (schematische Darstellung)



Ein analoges Beispiel sind die Einflüsse von Anlage- und Umweltfaktoren auf die Entwicklung phänotypischer Merkmalsausprägungen bei Kindern und Jugendlichen. Auch hier finden sich typischerweise Divergenzen interindividueller und intraindividuelle Wirkzusammenhänge, die Schlüsse von Zusammenhängen der ersteren auf Zusammenhänge der letzteren Art zu Fehlschlüssen werden lassen. Folgt man beispielsweise aus hoher Heritabilität von Intelligenz in bestimmten Populationen (also einer starken Bedingtheit von Intelligenzunterschieden zwischen Individuen durch genotypische Unterschiede), Umwelteinflüsse spielten nur eine geringe Rolle bei der Entstehung von Intelligenz, so wäre dies ein offenkundiger Fehlschluss: Ohne hinrei-

chend günstige physische und psychosoziale Umweltbedingungen werden genotypische Potenziale kaum in phänotypische Entwicklung umgesetzt, wie sich z.B. an den Folgen von Mangel- und Fehlernährung ablesen lässt; und zwar unabhängig davon, ob sich das Wirken dieser Bedingungen auch in resultierenden Unterschieden der phänotypischen Intelligenz verschiedener Individuen niederschlägt (leiden alle Kinder einer Population unter dem betreffenden Mangel, so kann die Intelligenzentwicklung der ganzen Population rückständig bleiben; vgl. auch Montada 1995).

Es zeigt sich also, dass populationsorientierte Längsschnitte in der Lage sind, Entwicklungstrends von Stichproben- bzw. Populations-Durchschnittswerten abzubilden und darüber hinaus Aussagen zu Bedingungsbeziehungen zwischen interindividuellen Verteilungen von Entwicklungsvariablen zu liefern. Hingegen lässt sich aus den Befunden so angelegter Studien – entgegen üblichen Interpretationsgewohnheiten weiter Bereiche auch der Jugendforschung – weder auf die tatsächlichen Entwicklungsverläufe noch auf die Ursachen solcher Verläufe bei einzelnen Jugendlichen bzw. Populationen von einzelnen Jugendlichen schließen. Insbesondere gilt, dass übliche längsschnittliche Bedingungsanalysen auf der Basis interindividueller Korrelationen Aussagen zu den Bedingungsbeziehungen liefern, die dem Zustandekommen von *Unterschieden* zwischen Jugendlichen zugrunde liegen; nicht gewinnen aber lassen sich Aussagen zu den grundsätzlich *intraindividuellen* Bedingungsbeziehungen, die Entwicklungsverlauf und resultierende Merkmalsausprägungen beim einzelnen Jugendlichen bzw. Populationen von einzelnen Jugendlichen bedingen.

### Problemlösungen durch idiographisch-nomothetische Forschungsstrategien

Welche Auswege bieten sich an, wenn man Rückschlüsse auf Entwicklungen und Bedingungsbeziehungen gewinnen möchte, die über Populationstrends und Personunterschiede hinaus auch den Einzelfall erklären? Antworten auf diese Frage dürften für die Zukunft von Psychologie, Sozialwissenschaften und Pädagogik und damit auch für die Zukunft der Jugendforschung entscheidend sein. So bedeutsam aggregierte Statistiken zunächst auch erscheinen mögen – wertvoll werden sie unter der Perspektive allgemeiner Entwicklungs- und Funktionsfragestellungen in der Regel erst dann, wenn sie möglichst viele, im Optimalfall sogar alle einzelnen Personen der jeweiligen Population zu erklären vermögen. Valide Schlüsse dieser Art aber sind offenbar kaum zu gewinnen, wenn man im Sinne der beschriebenen stichproben- und populationsorientierten Strategien als erstes aggregierte Werte gewinnt und erst anschließend auf die Ebene der zugrunde liegenden Einzelfälle

zurückschließt. Vielmehr wäre es offensichtlich notwendig, diese Einzelfälle direkt zu untersuchen und Generalisierungen erst dann anzustreben, wenn der Einzelfall analysiert ist.

Notwendig wäre es also, zunächst idiographischen Strategien zu folgen und erst anschließend in nomothetischer Absicht zu generalisieren. Im Sinne dieser von Windelband (1894) eingeführten begrifflichen Differenzierung handelt es sich um *Idiographie*, wenn Einzelfälle analysiert werden, ohne dass ein Anspruch auf Generalisierbarkeit erhoben wird; und um *Nomothetik*, wenn Schlüsse auf Populationen von Einzelfällen gezogen werden sollen, wobei diese Populationen typischerweise unabgeschlossen groß sind (z.B. nach oben offene Zahl aller Individuen unserer Spezies, die jemals gelebt haben und noch leben werden, als Bezugspopulation einer nomothetischen allgemeinen Psychologie, oder aller Jugendlichen als Bezugspopulation einer nomothetischen Jugendforschung). Dabei wäre es ein Missverständnis, idiographisches Vorgehen als unwissenschaftlich anzusehen und nur nomothetischen Strategien wissenschaftliche Dignität zuzusprechen: Diejenigen Natur-, Sozial- und Geisteswissenschaften, deren Gegenstände einzelfallartig organisiert sind, tragen per Gegenstandsdefinition idiographischen Charakter (Astronomie von Sonnensystemen, Geologie spezifischer Erdformationen, Geographie und Geschichte einzelner Länder und Nationen etc.). Auch für Jugend- und Leistungsstudien wie TIMSS, PISA oder die Shell-Jugendstudien gilt, dass sie mit ihrer Deskription spezifischer Kompetenz-, Einstellungs- und Verhaltensvariationen in definierten einzelnen Nationen zu bestimmten epochalen Zeitpunkten zunächst idiographischen Zwecken folgen. Mit der Analyse von allgemeineren Gesetzmäßigkeiten der Variation von Variablen und der Verknüpfungen von Variationen verschiedener Variablen können sie allerdings gleichzeitig auch nomothetischen Zielsetzungen dienen.

In einem ersten, idiographischen Schritt könnten Entwicklungsprozesse und ihre Bedingungen also zunächst direkt am Einzelfall, d.h. bei einzelnen Jugendlichen untersucht werden (bzw. – bei Betrachtung überindividueller Einheiten – bei einzelnen Schulen, Gemeinden, Regionen etc.). Eine solche Analyse würde beinhalten, zum einen den intraindividuellen Verlauf der jeweiligen Merkmale über die Zeit hinweg zu betrachten, zum anderen intraindividuelle Beziehungen zwischen diesen Merkmalen und möglichen Bedingungen oder Wirkungen ihrer Entwicklung zu analysieren, beispielsweise anhand synchroner und diachroner intraindividuelle Korrelationen, die anschließend – analog zum Vorgehen bei interindividuellen Korrelationen auf Stichprobenebene – anhand von längsschnittlich-bedingungsanalytischen Methoden auf die zugrunde liegenden Bedingungsbeziehungen hin untersucht werden können. Kritisch ist dabei die Zahl der Messzeitpunkte: Um reliable Verknüpfungs- und Effektmaße zu gewinnen, muss diese hinreichend hoch sein (analog zu den Forderungen an die Größe von Personenstichproben im populationsorientierten Vorgehen). Zu den geeigneten Erhebungsmetho-

den zählen u.a. Tagebuchverfahren, Beobachtung und apparative Methoden der Registrierung im Feld.

Allerdings sollten die Ansprüche der Jugendforschung gleichzeitig auch – wie von traditionellen populationsorientierten Vorgehensweisen angestrebt – über die Analyse von Einzelfällen hinausreichen: Wissenschaftlich ist der einzelne Jugendliche vor allem dann von Interesse, wenn sich von ihm auf möglichst viele andere Jugendliche und damit auf Populationen von Jugendlichen schließen lässt. Erreichen lässt sich dies mit Strategien, die idiographische und nomothetische Vorgehensweisen miteinander verknüpfen (vgl. Jaccard/Dittus 1990). Kern solcher Strategien ist, (1) eine hinreichend große Zahl von Einzelfällen in idiographischer Weise getrennt voneinander, aber anhand derselben Vorgehensweisen zu analysieren und anschließend (2) zu prüfen, inwieweit die Resultate über die untersuchten Einzelfälle hinweg Generalisierbarkeit zeigen. Soweit dies der Fall ist, lässt sich dann im Sinne des üblichen nomothetischen Vorgehens Generalisierbarkeit auch über die untersuchte Stichprobe von Einzelfällen hinaus folgern (sofern die Stichprobe hinreichend populationsrepräsentativ angelegt ist).

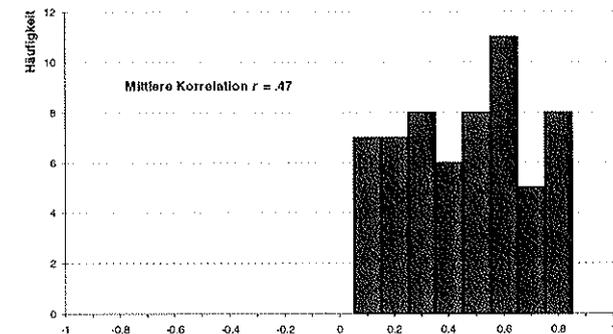
Gewinnt man auf der Basis einer solchen Strategie über Personen aggregierte Durchschnittswerte für Entwicklungsprozesse oder kausale Effekte, so sind diese Werte – anders als die direkt aggregierten Werte bei traditionellen Strategien – repräsentativ nicht nur für die betreffende Stichprobe bzw. Population, sondern auch für die einzelnen Individuen dieser Stichprobe. Voraussetzung hierfür ist, dass hinreichende Generalisierbarkeit von Prozessen bzw. Bedingungsbeziehungen über Individuen gegeben ist. Eine solche Generalisierbarkeit muss sich nicht auf Gesamtpopulationen aller Jugendlichen beziehen, sondern kann auf Teilpopulationen eingeschränkt sein (z.B. weibliche vs. männliche Jugendliche, Jugendliche mit unterschiedlich hohem Selbstvertrauen). Auch in einem solchen Fall lassen sich nomothetische Ansprüche aufrechterhalten. Unter nomothetischen Perspektiven problematisch wäre erst der Fall, dass sich unterschiedliche Individuen in so idiosynkratischer Weise unterschiedlich entwickeln bzw. so unterschiedliche Mechanismen des psychischen und sozialen Funktionierens zeigen, dass sich Gemeinsamkeiten über Individuen hinweg auch anhand geeigneter Definitionen von Teilpopulationen nicht mehr ausmachen lassen (erst in einem solchen Fall wären nomothetische Ansprüche der Jugendforschung grundsätzlich in Frage zu stellen; vgl. auch Tschacher 1997).

Beispiele für ein idiographisch-nomothetisches Vorgehen zur Analyse von Variablenverknüpfungen liefert die bereits zitierte Tagebuchstudie an Examenskandidaten (Hofmann 1997; Pekrun/Hofmann 1999). Gleichzeitig zeigen die Befunde dieser Studie einmal mehr, zu welcher irreführenden Schlussfolgerungen traditionelle, interindividuell-populationsorientierte Korrelationsanalysen verführen können. Untersucht wurden in dieser Studie neben der Entwicklung von prüfungsbezogenen Emotionen wie antizipatorischer Prüfungsfreude und Prüfungsangst u.a. auch die Entwicklung der

Lernmotivation über die einbezogenen 43 Tage vor und während des Prüfungszeitraums hinweg. Für Freude und Angst wurden nun getrennt für die 72 Examenskandidaten jeweils die intraindividuellen Korrelationen mit der Lernmotivation über die Zeit hinweg berechnet. Im Anschluss an dieses idiographische Vorgehen wurde die Verteilung der so ermittelten intraindividuellen Korrelationen über Individuen analysiert, um – im Sinne einer nomothetischen Prüfung allgemeiner Hypothesen zu Emotionswirkungen – Möglichkeiten überindividueller Generalisierbarkeit zu prüfen. Zusätzlich wurden die interindividuellen Korrelationen von Freude und Angst mit der Lernmotivation berechnet, nachdem zunächst die Werte für Freude, Angst und Motivation innerhalb der Personen über Zeitpunkte hinweg aggregiert worden waren (durch Mittelwertbildung), um geeignete Individualwerte zu gewinnen (emotionale und motivationale "Quasi-traits"; Hofmann 1997).

Die Abbildungen 5 und 6 zeigen die resultierenden Verteilungen intraindividuell korrelierender Korrelationen in der Personenstichprobe. Wie man sieht, ergeben sich für Freude Korrelationen mit der Lernmotivation, die ausnahmslos positiv sind und nur in ihrer Höhe variieren (siehe Abbildung 5; die Höhe der Korrelation hing insbesondere von der intraindividuellen Reliabilität der verwendeten Emotions- und Motivationswerte ab). Es handelt sich hier um eine Verknüpfung, die auf funktionale Homogenität der zugrunde liegenden Bedingungsbeziehungen über Individuen hinweg schließen lässt; mithin ist Generalisierbarkeit zu folgern. Gleichzeitig entspricht diese intraindividuelle Verknüpfung auch der interindividuellen Korrelation der zugeordneten aggregierten Individualwerte ( $r = .70$ ).

Abbildung 5: Freude und Lernmotivation: Intraindividuelle Korrelationen

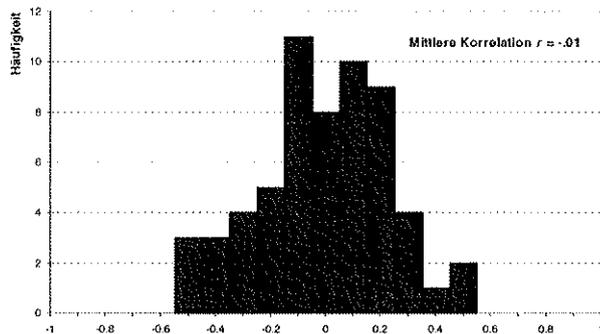


Ein anderes Bild aber zeigt sich für Prüfungsangst (Abbildung 6): Während Angst und Lernmotivation bei manchen Examenskandidaten negativ korrelieren, ergeben sich für einige Personen Nullkorrelationen und für einige Personen positive Beziehungen. Für den Zusammenhang von Prüfungsangst und

Lernmotivation wäre es also irreführend, aus der interindividuellen Korrelation ( $r = -.38$ ) oder auch der durchschnittlichen intraindividuellen Korrelation ( $r = -.01$ ) auf die Angst-Motivations-Beziehung im Einzelfall schließen zu wollen. Zu schließen ist in diesem Fall vielmehr auf funktionale Heterogenität: Unterschiedliche Individuen sind durch unterschiedliche motivationale Bezüge ihrer prüfungsbezogenen Angst gekennzeichnet.

Soweit sich diese korrelativen Beziehungen funktional interpretieren lassen, handelt es sich bei Examenskandidaten mit negativen Korrelationen um Personen, deren Angst zu Motivationsverlusten führt, bei Kandidaten mit positiven Korrelationen hingegen um Personen, die motivational von ihrer Angst profitieren. Ungeachtet funktionaler Heterogenität könnte es auch hier gelingen, Generalisierbarkeit sicherzustellen, wenn Personen mit negativen, fehlenden und positiven intraindividuellen Korrelationen getrennt voneinander betrachtet werden. Wenn ermittelt werden könnte, dass sich diese Personengruppen bezüglich spezifizierbarer, theoretisch ableitbarer Merkmale so voneinander unterscheiden, dass interindividuell unterschiedliche Motivationsbezüge individueller Angst eine Erklärung fänden, so wäre – den Zielsetzungen nomothetischer Forschung entsprechend – Generalisierbarkeit bestätigt: Anhand der betreffenden Merkmale könnte vorhergesagt werden, welcher Gruppe eine Person zuzuordnen wäre und welche Motivationswirkungen für diese Person angenommen werden müssten.

Abbildung 6: Angst und Lernmotivation: Intraindividuelle Korrelationen



Aus vorliegenden Angsttheorien lassen sich hierzu eine Reihe von plausiblen Annahmen ableiten. So lässt sich beispielsweise theoretisch deduzieren, dass die jeweiligen individuellen Motivationslagen für den Angst-Motivations-Zusammenhang entscheidend sein könnten. Zu vermuten ist, dass Angst intrinsische Motivationsformen negativ beeinflusst, hingegen manche Formen von extrinsischer Motivation verstärkt oder sogar erst erzeugt (z.B. Motivation zu Misserfolgsmeidung). Annehmen lässt sich deshalb, dass unterschiedliche Bezüge von Angst zur (Gesamt-)Lernmotivation auf interindi-

duell unterschiedlichen Gewichtungen solcher Motivationsformen und damit der motivationsförderlichen und -hinderlichen Wirkungen von Angst beruhen können (vgl. Pekrun, 1992); eine Annahme, die anhand der Daten der zitierten Tagebuchstudie allerdings nicht mehr analysiert werden kann.

## Schlussfolgerung

In diesem Beitrag ist argumentiert worden, dass traditionelle populationsorientierte Untersuchungsdesigns der Jugendforschung kaum geeignet sind, zentrale Fragen der Beschreibung und Erklärung von Entwicklungsprozessen zu beantworten: Auf individuelle Entwicklungen und Mechanismen des psychischen und sozialen Funktionierens von Individuen lässt sich aus ihnen nicht zurückschließen. Dies gilt auch für stichprobenbasierte Längsschnittstudien, die häufig als Königsweg der Entwicklungsforschung deklariert worden sind. Im Grundsatz ist dieser Sachverhalt der scientific community seit längerem bekannt. Es ist deshalb nachgerade erstaunlich, dass entsprechende Warnungen seitens der sozialwissenschaftlichen und psychologischen Methodenlehre (z.B. Asendorpf 1995; Schmitz 2000; Sidman 1952) in der Jugendforschung wie in anderen Forschungsdisziplinen immer wieder ignoriert worden sind (manche Arbeiten der Autoren des vorliegenden Beitrags stellen hier keine Ausnahme dar). Beiträge mit individualistischen Fehlschlüssen aus Stichprobendaten finden sich auch heute noch in so gut wie jedem Jahrgang national und international führender Zeitschriften der Jugend- und Entwicklungsforschung.

Wissenschaftstheoretische und -soziologische Analysen zur Trägheit konventioneller Denk- und Analysemuster fänden hier sicher ein reiches Betätigungsfeld. Zu vermuten ist, dass das Weiterführen von methodisch problematischen Forschungs- und Interpretationsstrategien Mechanismen entspringt, die für ein hinreichendes Maß an Ökonomie sorgen sollen, um wissenschaftliches Arbeiten zu ermöglichen. In Gefahr sind hier nämlich kognitive ebenso wie finanzielle und zeitliche Ressourcen: Analysen unterschiedlicher Einzelfälle sind theoretisch anstrengend, da man Gefahr läuft, sich von simplifizierenden Annahmen verabschieden zu müssen; und sie werden finanziell und zeitlich spätestens dann teuer, wenn nicht nur ein einzelner Fall über viele Messzeitpunkte hinweg zu analysieren ist, sondern hinreichend große Stichproben von Einzelfällen, um – anders als in manchen Zweigen der qualitativ orientierten Jugendforschung üblich – wissenschaftliche Generalisierbarkeit von Beschreibungen und Erklärungen sicherzustellen.

Wenn aber die Jugendforschung valide Erkenntnisse zutage fördern soll, gibt es keinen anderen Weg, als gegenstandsangemessene Forschungsstrategien zu verwenden. Dass Strategien der beschriebenen idiographisch-nomo-

thetischen Art durchaus praktikabel sein können, wird mittlerweile durch eine Vielzahl einzelner Untersuchungen belegt (Beispiele finden sich in Schmitz 2000). Will die Jugendforschung Entwicklungsprozesse bei Jugendlichen beschreiben und erklären, so dürfte sie folglich gut daran tun, neben aggregierend-differenziell angelegten Vorgehensweisen traditioneller Art zukünftig in verstärktem Maße idiographisch basierte Strategien zum Einsatz zu bringen, ohne dabei Ansprüche auf Generalisierbarkeit aufzugeben.

## Literatur

- Asendorpf, Jens (1995): Persönlichkeitspsychologie: Das empirische Studium der individuellen Besonderheit aus spezieller und differentieller Perspektive. In: *Psychologische Rundschau* 46, 235-247
- Baumert, Jürgen/Lehmann, Rainer (1997): TIMSS – Mathematisch-naturwissenschaftlicher Unterricht im internationalen Vergleich. Opladen: Leske + Budrich
- Bryk, Anthony S./Raudenbush, Stephen W. (1987): Application of Hierarchical Linear Models to Assessing Change. In: *Psychological Bulletin* 101, 147-158
- Eye, Alexander van (Ed.) (1990): *Statistical Methods in Longitudinal Research*. New York: Academic Press
- Hoffmann-Lange, Ursula (2001): Der fragwürdige Beitrag von Jugendstudien zur Analyse von Trends in der politischen Kultur. In: Merkens, H./Zinnecker, J. (Hrsg.): *Jahrbuch Jugendforschung 1*. Opladen: Leske + Budrich, 187-210
- Hofmann, Hubert (1997): Emotionen in Lern- und Leistungssituationen – eine idiographisch-nomothetische Tagebuchstudie an Lehramtsstudenten im Examen. Unveröff. Dissertation. Universität Regensburg: Fakultät für Psychologie und Pädagogik
- Jaccard, James/Dittus, Patricia (1990): Idiographic and Nomothetic Perspectives on Research Methods and Data Analysis. In: Hendrick, C./Clark, M. S. (Eds.): *Research Methods in Personality and Social Psychology*. Newbury Park, CA: Sage, 312-351
- Krapp, Andreas (1997): Selbstkonzept und Leistung – Dynamik ihres Zusammenspiels: Literaturüberblick. In: Weinert, F. E./Helmke, A. (Hrsg.): *Entwicklung im Grundschulalter*. Weinheim: Beltz, 325-339
- Lütkepohl, H. (1991). *Introduction to Multiple Time Series Analysis*. Berlin: Springer
- Marsh, Herbert/Yeung, Alexander S. (1997): Causal Effects of Academic Self-concept on Academic Achievement: Structural Equation Models of Longitudinal Data. In: *Journal of Educational Psychology* 89, 41-54
- Montada, Leo (1995): Fragen, Konzepte, Perspektiven. In: Oerter, R./Montada, L. (Hrsg.): *Entwicklungspsychologie. Ein Lehrbuch*. Weinheim: Psychologie Verlags Union
- Nesselroade, John R./Cattell, Raymond B. (Eds.) (1988): *Handbook of Multivariate Experimental Psychology*. New York: Plenum
- Pekrun, Reinhard (1992): The Impact of Emotions on Learning and Achievement: Towards a Theory of Cognitive/Motivational Mediators. In: *Applied Psychology: An International Review* 41, 359-376
- Pekrun, Reinhard (1997): Selbstkonzept und Leistung – Dynamik ihres Zusammenspiels: Kommentar. In: Weinert, F. E./Helmke, A. (Hrsg.): *Entwicklung im Grundschulalter*. Weinheim: Beltz, 351-358
- Pekrun, Reinhard/Fend, Helmut (Hrsg.) (1991): *Schule und Persönlichkeitsentwicklung. Ein Resümee der Längsschnittforschung*. Stuttgart: Enke
- Pekrun, Reinhard/Hofmann, Hubert (1999): Lern- und Leistungsemotionen: Erste Befunde eines Forschungsprogramms. In: Jerusalem, M./Pekrun, R. (Hrsg.): *Emotion, Motivation und Leistung*. Göttingen: Hogrefe, 247-268
- Robinson, William S. (1950): Ecological Correlations and the Behavior of Individuals. In: *American Sociological Review* 15, 351-356
- Rogosa, David (1988): Myths About Longitudinal Research. In: Schaie, W./Campbell, R. T./Meredith, W./Rawlings, S. C. (Eds.): *Methodological Issues in Aging Research*. New York: Springer, 171-209
- Schmitz, Bernhard (2000): Auf der Suche nach dem verlorenen Individuum: Vier Theoreme zur Aggregation von Prozessen. In: *Psychologische Rundschau* 51, 83-92
- Schmitz, Bernhard/Skinner, Ellen (1993): Perceived Control, Effort, and Academic Performance: Interindividual, Intraindividual, and Multivariate Time-series Analyses. In: *Journal of Personality and Social Psychology* 64, 1010-1028
- Sidman, Murray (1952): A Note on Functional Relations Obtained from Group Data. In: *Psychological Bulletin* 49, 263-269
- Tschacher, Wolfgang (1997): *Prozeßgestalten. Die Anwendung der Selbstorganisationstheorie und der Theorie dynamischer Systeme auf Probleme der Psychologie*. Göttingen: Hogrefe
- Valsiner, Jaan (Ed.) (1986): *The Individual Subject and Scientific Psychology*. New York: Plenum
- Zinnecker, Jürgen (2001): Fünf Jahrzehnte öffentliche Jugend-Befragung in Deutschland. Die Shell-Jugendstudien. In: Merkens, H./Zinnecker, J. (Hrsg.): *Jahrbuch Jugendforschung 1*. Opladen: Leske + Budrich, 243-274
- Windelband, Wilhelm (1894): *Geschichte und Naturwissenschaft*. Straßburg: Heitz