

## Kausalmodelle in der Pädagogischen Psychologie: Ergänzende Analysen zu zwei einschlägigen Untersuchungen

Wolfgang Schneider

Universität Heidelberg

In the present secondary analysis the results of studies by Simons et al. (1975) and Schwarzer (1979) were critically reanalyzed. The central point of the reanalysis concerned the test of the causal models used, especially the question, if the technique of path analysis was correctly applied. Whereas some modifications seemed necessary in both studies, there was no reason to question their main results.

Während die Technik der Pfadanalyse in soziologischen Untersuchungen schon seit geraumer Zeit dazu eingesetzt wird, kausale Hypothesen über Merkmalszusammenhänge zu überprüfen, ist diese Weiterführung des regressionsanalytischen Ansatzes in der pädagogischen Effektforschung vergleichsweise selten und erst in jüngerer Zeit gezielter verwendet worden.

Trotz der inzwischen verfügbaren Einführungen in das Verfahren (vgl. z. B. Anderson, 1974; Anderson & Evans, 1978; Hummell & Ziegler, 1976; Opp & Schmidt, 1976; Weede, 1977 u. a. m.) weisen Kausalanalysen empirischer Daten nicht selten spezifische Mängel bzw. Unvollkommenheiten auf, die sich wie folgt charakterisieren lassen: der Forscher begnügt sich damit, ein (mehr oder weniger) theoriegeleitetes rekursives<sup>1)</sup> Kausalmodell so zu konstruieren, daß es vollständig spezifiziert (bzw. „gerade identifiziert“) genannt werden kann; in diesem Modell sind demnach alle als kausal vorgeordnet gedachten Variablen zu allen nachgeordneten über Pfade in Beziehung gesetzt. Der Rechengvorgang ist zumeist mit der Schätzung der Pfad- bzw. Regressionskoeffizienten des vollständigen Modells abgeschlossen, die dann als Interpretationsgrundlagen und Kriterien für die Ablehnung oder Beibehaltung der gewählten kausalen Hypothesen dienen.

---

1) Wechselwirkungs- bzw. *Feedback-Prozesse* sind bei dieser immer noch gebräuchlichsten Modellkonstruktionsvariante ausgeschlossen.

Bei einer solchen Vorgehensweise wird übersehen, daß vollständig spezifizierte Modelle zur Hypothesentestung unbrauchbar sind, da sich für diesen Spezialfall die Ausgangs-Korrelationsmatrix durch die geschätzten Pfadkoeffizienten unabhängig davon, welche Kausalordnung der Variablen nun auch immer gewählt worden ist, perfekt reproduzieren läßt (vgl. Pedhazur, 1975). Es wird damit evident, daß Koeffizientenschätzungen im vollständig spezifizierten Kausalmodell keinerlei Aussagen über die Adäquanz des betreffenden Modells erlauben.

Hypothesentestungen sind nur dann möglich, wenn sogenannte überidentifizierte Modelle vorliegen, wenn also (entweder *a priori* oder nach Inspektion der Koeffizienten des vollständigen Modells) bestimmte Pfadkoeffizienten gleich Null gesetzt werden. Von Modell-Falsifikation kann dann gesprochen werden, wenn der Vergleich von Ausgangs-Korrelationsmatrix und der über die geschätzten Pfadkoeffizienten reproduzierten Korrelationsmatrix bedeutsame Abweichungen (beispielsweise nach einer praktisch bewährten Faustregel  $\leq .10$ )<sup>2</sup> ergibt. Aufgrund der theoretisch gegebenen Möglichkeit, mehrere Alternativmodelle für einen umschriebenen Gegenstandsbereich zu konstruieren, sind solche Lösungen zu präferieren, die sich einmal als maximal reduziert und damit sparsam herausstellen, zum anderen aber gleichzeitig relativ hohe Varianzaufklärungs-Quoten für die modell-inhärenten (endogenen) Variablen erbringen.

Angesichts der oben näher bezeichneten Mißverständnisse bzw. Unklarheiten im Umgang mit Kausalanalysen schien es durchaus zweckmäßig, auch die wenigen, z. T. aber vielbeachteten Beispiele pfadanalytischer Modellkonstruktion in deutschsprachigen pädagogisch-psychologischen Publikationen daraufhin zu untersuchen, ob hier tatsächlich Modellüberprüfungen vorgenommen wurden und — falls ja — ob letztere zu optimalen, d. h. zu maximal reduzierten und aussagekräftigen Lösungen geführt hatten.

Besonders gut geeignet schienen dabei die inhaltlich eng miteinander verknüpften Arbeiten von Simons, Weinert & Ahrens (1975) und Schwarzer (1979) zur Problematik der Schulerfolgsprognose vor allem deshalb zu sein, weil hier die Frage nach der korrekten Wahl des Modell-Typs (rekursives vs. *Feedback-System*) bzw. nach möglichen alternativen Variablen-Konstellationen innerhalb des Modells vernachlässigt werden konnte: durch die Annahme eines hierarchischen Lernmodells (Simons et al.) bzw. die Verwendung von Längsschnittdaten (Schwarzer) dürfte die Konstruktion von rekursiven Modellen als durchaus gerechtfertigt gelten. Zum besseren inhaltlichen Verständnis sollen die wesentlichen Annahmen der beiden Studien noch einmal kurz rekapituliert werden.

---

2) Die als Alternative zu einer derartigen Faustregel angebotenen inferenzstatistischen Verfahren (*likelihood-ratio-Tests*) werden hier in Anlehnung an Weede (1977) wegen Zweifel an der Häufigkeit ihrer Anwendungsvoraussetzungen vernachlässigt.

In der Untersuchung von Simons et al. wurde die Brauchbarkeit von globalen Maßen der Lernleistung für die Vorhersage des Schulerfolgs in Frage gestellt und stattdessen die Berücksichtigung spezifischer Vorkenntnisstrukturen für die Prädiktion komplexerer Kriteriumsleistungen gefordert. In Anlehnung an das Lernkomponentenmodell Gagnés ließen sich Mathematikaufgaben für 4. Klassen auf der Basis des Schwierigkeitsniveaus hierarchisch in drei Vorkenntnis- und eine Kriteriumsebene gliedern<sup>3</sup>). Die so definierten vier Vorkenntnisstufen als endogene (d. h. innerhalb des Kausalsystems zu erklärende) Variablen sowie Ängstlichkeit und Allgemeine Intelligenz als vorgeordnete exogene (d. h. nicht im Modell erklärbare) Variablen bildeten die Grundlage für ein rekursives kausales Bedingungsmodell. Die Hypothesen, denen zufolge bei zunehmender Aufgabenkomplexität Intelligenzeinflüsse ab- und Ängstlicheitseinflüsse zunehmen sowie substantiell direkte Einflüsse nur für unmittelbar benachbarte Kenntnisebenen beobachtbar sein sollten, ließen sich über das Kausalmodell im wesentlichen bestätigen.

Dieser von Simons et al. über eine Querschnittsanalyse vorgenommene Vergleich der Relevanz von allgemeinen Schülerpersönlichkeitsmerkmalen und fachspezifischen Vorkenntnissen im Hinblick auf die Schulerfolgsprognose wurde von Schwarzer (1979) dadurch sinnvoll ergänzt, daß die Hypothese des abnehmenden Einflusses von Intelligenzkenntwerten bei gleichzeitig zunehmender Bedeutung von Vorkenntnisvariablen über eine auf drei Jahre befristete, methodisch sicherlich adäquatere Längsschnittstudie bei Schülern der Orientierungsstufe (4.—6. Klasse) überprüft wurde. In das Kausalmodell gingen neben Intelligenz- und Ängstlichkeitskenntwerten die Zensuren des 4. bzw. Testwerte des 5. und 6. Schuljahrs für die Fächer Mathematik und Deutsch ein (letztere wurden bei der Sekundäranalyse aus Vergleichbarkeitsgründen ausgespart). Auch hier fielen die Befunde im wesentlichen hypothesenkonform aus, was im Prinzip als weitere Bestätigung der Ergebnisse von Simons et al. gewertet wurde.

### *Methode der Sekundäranalyse*

Im Unterschied zu den Berechnungen von Simons et al. und Schwarzer, die alle auf der Basis von üblichen Regressionsprozeduren durchgeführt worden waren, wurde für die Re-Analyse das spezielle Pfadanalyse-Programm YPFAD (Ziegler, 1972) verwendet.

Der Vorteil dieses Programms kann darin gesehen werden, daß die für die Modelltestung wesentliche Differenzmatrix (Unterschiede zwischen der ursprünglichen und der durch das Modell implizierten Korrelationsmatrix) sowie die Summe der Abweichungsquadrate (SAQ) zusätzlich enthalten sind.

Im Anschluß an den zur Kontrolle wichtigen Koeffizienten-Vergleich für die vollständig spezifizierten Modelle wurde in der Sekundäranalyse jeweils geprüft, ob und inwieweit sich die bei Simons et al. sowie Schwarzer vorgelegten Kausalmodelle weiter reduzieren ließen, ohne als falsifiziert gelten zu müssen<sup>4</sup>).

3) vgl. zur Problematik dieser Prozedur bzw. zur Entwicklung von Alternativmethoden (Kleiter & Petermann, 1977).

4) Das vollständige Pfadmodell wurde hierbei solange um die nichtsignifikanten Pfade reduziert, bis Modell-Falsifikation im Sinne der oben erwähnten ‚Faustregel‘ vorlag.

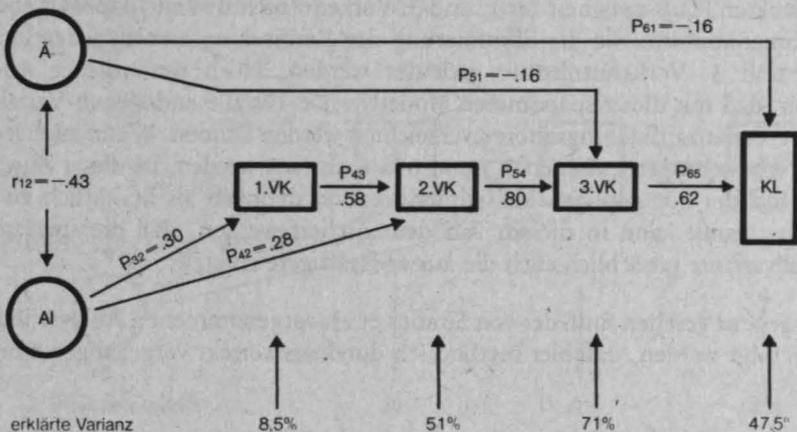
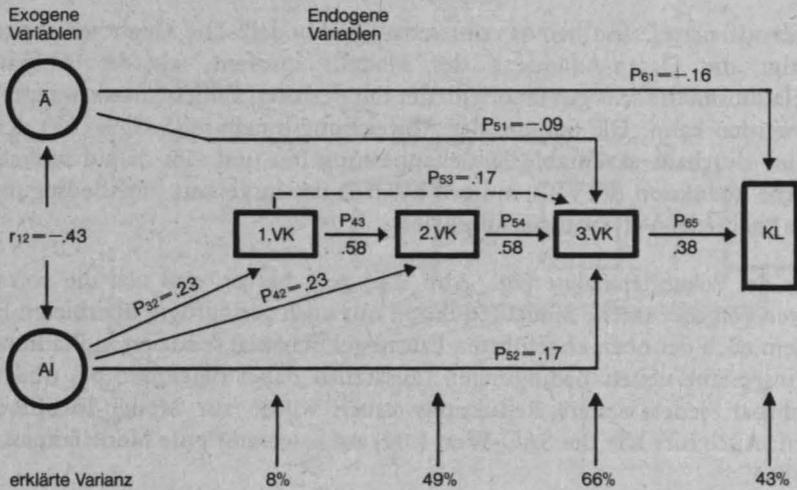


Abb. 1. Pfadanalyse zur differentialpsychologischen Analyse von Rechenleistungen

a) Überidentifiziertes Kausalmodell von Simons et al. (1975)

b) Ergebnis der Sekundäranalyse

 $\bar{A}$  = Schulungst AI = Allgemeine Intelligenz

VK = Vorkenntnisebene KL = Kriteriumsleistung

### Ergebnisse

a) Replikation zu Simons et al. (1975). Wie aus Abb. 1a abzulesen ist, handelt es sich bei der von Simons et al. dargelegten Ergebnisversion um ein

überidentifiziertes, also bereits reduziertes Pfadmodell. Die Hypothesentestung bestätigt die Daten-Adäquanz des Modells insofern, als die Ausgangs-Korrelationsmatrix sehr gut innerhalb der tolerierbaren Fehlerschranken reproduziert werden kann. Die Summe der Abweichungsquadrate ( $SAQ = .03$ ) deutet auf eine durchaus akzeptable Modellanpassung hin und läßt darauf schließen, daß eine Reduktion des vollständigen Modells um insgesamt fünf Bedingungen (Pfade) keinerlei Verzerrungen impliziert.

Aus der Sekundäranalyse (vgl. Abb. 1b) geht hervor, daß sich die von den Autoren vorgenommene Modellreduktion nur noch geringfügig überbieten läßt. Bei dem nach der oben angeführten Faustregel maximal reduzierten Pfadmodell sind insgesamt sieben Bedingungen (zusätzlich dabei die Pfade  $p_{52}$  und  $p_{53}$ ) verzichtbar; jeder weitere Reduktionsversuch würde zur Modell-Falsifikation führen. Auch hier läßt der SAQ-Wert (.04) auf insgesamt gute Modellanpassung schließen.

Rein inhaltlich betrachtet steht das in Abb. 1b aufgeführte Bedingungsmodell mit den zugrundeliegenden Hypothesen noch besser in Einklang: der Wegfall des direkten Pfads zwischen der 1. und 3. Vorkenntnisstufe kann dabei als ebenso annahmekonform wie die Eliminierung der Verbindung zwischen der Intelligenz und 3. Vorkenntnisstufe gedeutet werden. Noch wesentlicher scheint jedoch, daß mit dieser sparsameren Modellversion für alle endogenen Variablen höhere Varianzaufklärungsquoten verzeichnet werden können. Wenn auch lediglich Verbesserungen zwischen 0,5 und 6% registriert werden, ist dieser Zuwachs aufgrund der nur geringen Modellmodifikation dennoch als beachtlich zu bewerten. Damit kann in diesem Fall demonstriert werden, daß die sparsamere Modellvariante tatsächlich auch die aussagekräftigere darstellt.

Insgesamt gesehen muß der von Simons et al. vorgenommenen Analyse jedoch bescheinigt werden, daß hier methodisch durchaus korrekt vorgegangen worden ist<sup>5)</sup>.

b) *Sekundäranalyse zu Schwarzer (1979)*. Aus Abb. 2a läßt sich demgegenüber unschwer entnehmen, daß für das von Schwarzer konstruierte rekursive kausale Bedingungsmodell keinerlei Hypothesentestungen durchgeführt worden sind. Es erstaunt dabei besonders, daß es dem Autor offensichtlich bewußt darum ging, „Überidentifikation zu vermeiden“ (vgl. Schwarzer, 1979, S. 4), so daß von daher das vollständig spezifizierte Modell präferiert wurde.

---

5) Auf die Frage, inwieweit bei der Querschnittsanalyse funktioneller Hierarchien Pfadanalysen überhaupt sinnvolle Aussagen über die Prognoserichtung ermöglichen, bzw. inwieweit Korrelationskoeffizienten als Symmetriemaße für Analysen dieser Art geeignet sind, wird hier nicht näher eingegangen. Ausführliche Erörterungen hierzu finden sich bei Möbus (1980).

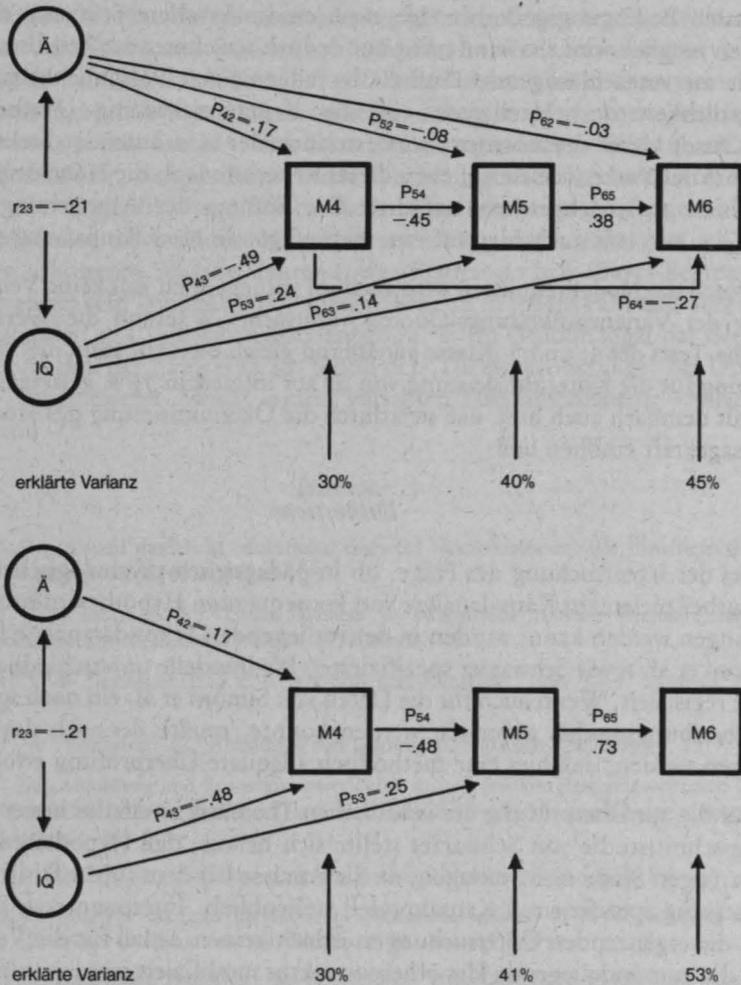


Abb. 2. Pfadanalyse zur sequentiellen Prädiktion des Schulerfolgs  
 a) Gerade identifiziertes Kausalmodell von Schwarzer (1979)  
 b) Überidentifiziertes Modell der Sekundäranalyse

$\ddot{A}$  = Ängstlichkeit     $IQ$  = Intelligenz  
 $M4$  = Mathematik-Note der 4. Klasse  
 $M5$  und  $M6$  = Mathematik-Testwerte der 5. und 6. Klasse

Die Hypothesentestung (vgl. Abb. 2a) macht jedoch deutlich, daß bei einer Eliminierung von vier Bedingungen immer noch von einer Modellbestätigung ausgegangen werden kann. Ähnlich wie bei der Re-Analyse zu den Simons et al.-Daten zeigt sich auch hier, daß durch die Straffung und Ökonomisierung des

kausalen Bedingungsgeflechts eine noch eindrucksvollere Stützung der Hypothesen möglich wird. So kann nicht nur der mit zunehmender Zeitdistanz immer mehr zu vernachlässigende Einfluß der allgemeinen Persönlichkeitsmerkmale Ängstlichkeit und Intelligenz auf die Kriteriumsleistung (Mathematiktest 6. Klasse) klarer demonstriert, sondern auch der zunehmende direkte kausale Einfluß der Vorkenntnisse auf eben dieses Kriterium (vgl. die Höhe des Pfadkoeffizienten  $p_{65}$ ) nachgewiesen werden. Die Summe der Abweichungsquadrate ( $SAQ = .04$ ) läßt auch hier auf eine befriedigende Modellanpassung schließen.

Durch die Modellreduktion wird auch im vorliegenden Fall keine Verschlechterung der Varianzaufklärungs-Quoten verursacht. Während die Werte für die Mathe-Tests der 4. und 5. Klasse annähernd gleich bleiben, kann die Varianzaufklärung für die Kriteriumsleistung von 45 auf immerhin 53 % gesteigert werden. Es gilt demnach auch hier, daß sich durch die Ökonomisierung des Modells seine Aussagekraft erhöhen ließ.

### *Diskussion*

Bei der Untersuchung der Frage, ob in pädagogisch-psychologischen Anwendungsbeispielen zur Kausalanalyse von konsequenten Hypothesentestungen ausgegangen werden kann, wurden in der vorliegenden Sekundäranalyse für die bei Simons et al. sowie Schwarzer spezifizierten Pfadmodelle unterschiedliche Ergebnisse registriert. Wenn auch für die Daten von Simons et al. ein noch sparsameres Beschreibungsmodell gefunden werden konnte, mußte dennoch davon ausgegangen werden, daß hier eine methodisch adäquate Überprüfung erfolgt war.

Für die zur Überprüfung der inhaltlichen Thematik zweifellos besser geeignete Längsschnittstudie von Schwarzer stellte sich heraus, daß Hypothesentestungen im strengen Sinne nicht vorlagen, da die Analyse bei dem (nicht falsifizierbaren) vollständig spezifizierten Kausalmodell stehenblieb. Interessanterweise ergaben hier die ergänzenden Untersuchungen jedoch keinen Anlaß für die Vermutung, daß die zugrundeliegende Hypothesenstruktur modifiziert werden müßte; ebenso wie bei Simons et al. ließen sich die theoretischen Vorannahmen durch das reduzierte Modell noch pointierter und insgesamt eindeutiger bestätigen.

Auch die Befunde der Sekundäranalysen legen es demnach nahe, bei der Prognose des Schulerfolgs davon auszugehen, daß die Bedeutsamkeit allgemeiner Persönlichkeitsmerkmale mit zunehmender Lernerfahrung zugunsten von spezifischen Vorkenntnissen geringer wird.

Anzumerken bleibt jedoch, daß die im Hinblick auf die Modelltyp-Entscheidung vorteilhafte inhaltliche Bestimmung der unabhängigen Variablen in beiden Studien Probleme aufwirft, die die generelle Interpretierbarkeit der Befunde betreffen. Die relativ hohen Interkorrelationen zwischen den Prädiktoren haben in Form des schon aus der Regressionsrechnung bekannten Phänomens der

‚Multikollinearität‘ (multiple lineare Abhängigkeit) insbesondere Auswirkungen auf Größe, Vorzeichen und Standardschätzfehler der Pfadkoeffizienten (vgl. im Detail Opp & Schmidt, 1976, S. 168 ff.). Die im Rahmen der Re-Analyse nebenbei ermittelten Multikollinearitätskennwerte erreichten bzw. überschritten bei Schwarzer in zwei (M4 und M5), bei Simons et al. sogar in drei Fällen (1., 2. und 3. Vorkenntnisstufe) den als Konvention vorgeschlagenen Grenzwert von 0.60. Es bleibt allerdings die Frage offen, ob bei den hier angesprochenen hierarchischen Modellen die bei hoher Multikollinearität angeblich ‚therapeutisch‘ verwendbaren robusteren Schätzverfahren (*ridge*-Regression bzw. Bayes-Regression) viel erfolgreicher sein können: problematisch scheint in jedem Fall, daß die erklärenden Variablen z. T. erheblich überlappen, z. T. im Grunde sogar das Gleiche messen, ohne daß der hier (etwa im Sinne der ‚Therapievorschläge‘ bei Opp & Schmidt) angemessene Schritt einer Variablen-Zusammenfassung vollzogen werden kann.

### Literatur

- Anderson, J. G., Causal models in educational research: Nonrecursive models. *American Educational Research Journal*, 1978, 15, 81—97.
- Anderson, J. G. & Evans, F. B., Causal models in educational research Recursive models. *American Educational Research Journal*, 1974, 11, 29—39.
- Hummell, H. J. & Ziegler, R. (Hg.), Korrelation und Kausalität, Bde. 1—3. Stuttgart: Enke, 1976.
- Kleiter, E. F. & Petermann, F., Abbildung von Lernwegen. München: Oldenbourg, 1977.
- Möbus, C., Zur Abbildung von Entwicklungsverläufen mittels Strukturgleichungssystemen (Test-Retest-Situationen). In G. Rudinger (Hg.), *Methoden der Entwicklungspsychologie*. Stuttgart: Kohlhammer, 1980.
- Opp, K.-D. & Schmidt, P., Einführung in die Mehrvariablenanalyse. Reinbek: Rowohlt, 1976.
- Pedhazur, E. J., Analytic methods in studies of educational effects. In F. N. Kerlinger (Hg.), *Review of research in education*, Vol. III, Itasca, Ill.: Peacock, 1975.
- Schwarzer, R., Sequentielle Prädiktion des Schulerfolgs. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 1979, 11, 1—11.
- Simons, H., Weinert, F. E. & Ahrens, H.-J., Untersuchungen zur differentialpsychologischen Analyse von Rechenleistungen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 1975, 7, 153—169.
- Weede, E., Hypothesen, Gleichungen und Daten. Spezifikations- und Meßprobleme bei Kausalmodellen für Daten aus einer und mehreren Beobachtungsperioden. Kronberg, Ts.: Athenäum, 1977.
- Ziegler, R., Theorie und Modell. München: Oldenbourg, 1972.

Anschrift des Verfassers:

Dr. Wolfgang Schneider, Psychologisches Institut der Universität Heidelberg,  
6900 Heidelberg, Hauptstraße 47—51