
ZUR BESCHREIBUNG UND ANALYSE KURZ- UND LANGFRISTIGER
TESTINTELLIGENZVERÄNDERUNGEN
MIT
ZEITDISKRETEN UND ZEITKONTINUIERLICHEN DYNAMISCHEN MODELLEN

Claus Möbus

Summary (Description and analysis of short- and longtimed changes in intelligence test scores with time-discrete and time-continous dynamic models): Longitudinal studies or repeated measurement designs seldom include process models of the empirical phenomenon. Evaluation of change or intervention effects is nearly always done by static statistical methods (e.g. ANOVA). Static models, however, are not very adequate for the planning of optimal intervention schedules. Especially if the statistical results are nonsignificant, we should explain this finding in terms of dynamic processes: (a) was the intervention not effective at all? (b) was the intervention effective but too short? (c) was the change in the dependent variable only transitory, so that the intervals between observations were too long? It is the aim of this paper to show how some of these questions could be answered. We analyse a short-termed intelligence test coaching and crossvalidate the results with a nonexperimental longitudinal study. It can be shown that a simple crossvalidation of the experimental results fails. It is possible, however, to crossvalidate the results of a simulation study, which is based on the experimental findings. This confirms the hypothesis, that the experimental study was too short for stabilizing the effects on the dependent variables. It suggests itself, that many unsuccessful cross-validation studies could be explained in a similar fashion.

The simulation study was based on a time-discrete model (difference equation systems), whereas the study of intervention effects and the development of hypotheses at any time necessitates the development of time-continous models. This is especially true, when the intervals between measurements are very long. These models are based on differential equation systems with constant coefficients. It is shown, how to identify and estimate the parameters with panel or two-wave data. Identification is done via the LAPLACE-transformation of the system and the comparison of the input-output equation of the model with its state-variable representation.

1. Einleitung

Pawlik (1976) zählt unter die Zielsetzungen psychologischer Diagnostik u.a. die Veränderungsmessung und die Entscheidungs/Behandlungsoptimierung. Letzteres ist unerlässlich, wenn man an enge Rahmenbedingungen hinsichtlich der zur Verfügung stehenden Ressourcen gebunden ist. So läßt sich die zeitliche Dauer der Intervention (Therapie, Training, Instruktion etc.) und der den Prozeßverlauf abbildenden Messungen aus verschiedenen Gründen (Ermüdung der Pbn, Kosten der Datenerhebung etc.) nicht beliebig verlängern. Daher stellt sich oft die Aufgabe, bei fester Test- und Interventionsdauer die "Mixtur" der Trainingsteile optimal zu verändern. Läßt sich der Interventionseffekt nicht statistisch sichern, wird damit meist vorschnell auf die mangelnde Güte der Intervention geschlossen. Eine genaue Analyse zeigt aber einen weitaus komplizierteren Sachverhalt. Es gibt mindestens drei Erklärungsmöglichkeiten für die Nichtsignifikanz: (a) mangelnde Güte des Interventionsprogramms, (b) zu kurze zeitliche "Dosierung" eines an sich effektiven Programms, (c) Flüchtigkeit des Effekts: nach dem Absetzen der Intervention verschwindet der Effekt (liegen die Meßzeitpunkte weit auseinander, wird oft gar kein Effekt wahrgenommen). Die letzten beiden Erklärungen sowie die Planung der Optimierung können mit klassischen statistischen Verfahren (Varianz-, Faktorenanalyse etc.) nicht geleistet werden, weil sie als statische Modelle die zeitliche Prozesstruktur der empirischen Phänomene nur ungenügend modellieren. Selbst die für die Analyse zeitbezogener Daten gerne herangezogene Varianzanalysenvariante "Wachstumskurvenanalyse" (Khatrı, 1966; Timm, 1975) besteht im Prinzip nur aus einer Polynomannäherung der Mittelwertsverläufe. Eine inhaltliche Begründung für die Be-

schreibung von Entwicklungsverläufen mit Polynomen wird in der Regel nicht gegeben. Statt dessen werden technische Argumente ("leichtere Berechenbarkeit") angeführt.

Statt dessen wollen wir Längsschnittdaten in diskreter Zeit mit Differenzgleichungssystemen und in kontinuierlicher Zeit mit Differentialgleichungssystemen analysieren. Beide Modelltypen sind im Gegensatz zu varianzanalytischen Methoden in dem Sinne dynamisch, daß alle Variablen bzw. Variablenänderungen explizit Funktionen der Zeit sind. Mit den dynamischen Modellen läßt sich die Transferstruktur der Variablen untereinander und der Einfluß exogener Variabler abschätzen. Anschließend sind mit den so gewonnenen Parametern Prognosen und Simulationen der erwarteten Entwicklungsverläufe möglich. Im Gegensatz zu den von Buss (1973) und Flammer(1975) vorgeschlagenen Modellen ist dabei die Formulierung zeitbezogener Hypothesen möglich: Wir erwarten unter bestimmten Rahmenbedingungen zu einem vorher festgelegten Zeitpunkt einen Mittelwertvektor $\mu(t)$ mit den Werten a,b,c.. Diese Hypothese läßt sich dann multivariat (zB. mit Hotelling's T^2) testen. Dabei ist zu erwarten, daß wir "gegen uns arbeiten": Je größer die Prognosedistanz, desto größer die Wahrscheinlichkeit der Modellfalsifikation. Muß die Prognose falsifiziert werden, gibt es dafür u.a. fünf Erklärungen: (a) vom Forscher nicht bemerkte Änderung des in der Empirie geltenden "Gesetzes", (b) Formulierung eines von Anfang an falschen Modells für das Prozessphänomen, (c) Nichtbeachtung relevanter Variabler, (d) mangelnde Kontrolle der exogenen Randbedingungen oder deren falsche Vorhersage, (e) falsche Messung des Anfangszustandes (Pretest). Die Testung eines Modells entlang der Zeitachse verläuft wesentlich strenger als Modellkontrollen in einem nicht zeitlich aufgespannten Rahmen.

2. Empirische und theoretische Analyse

2.1 Zeitdiskrete dynamische Modelle und Prüfung des "Dosierungsproblems"

Im Rahmen einer Untersuchung zur Beeinflußbarkeit von Intelligenztestwerten am Beispiel des IST haben Hamouzova & Würthner (Anm.1) ein Kontrollgruppendesign mit 2 Meßzeitpunkten im Abstand von 8 Tagen durchgeführt. Nach den Pretestergebnissen wurden 3 parallelisierte Gruppen (N=36) gebildet. Gruppe A und B erhielten am 3. oder 4.Tag nach dem Pretest ein 3-stündiges programmiertes Coaching mit ähnlichen aber nicht identischen Aufgaben (aus MIT, LPS etc.). Gruppe C diente als Kontrollgruppe. A und B unterschieden sich nur in der Form der Rückmeldung auf nicht gelöste Aufgaben. Während in B nur über die Nichtlösung informiert wurde, erhielten die Pbn in A zusätzliche Denkstrategiehilfen. 8 Tage nach dem Pretest wurde die 2.Messung mit der Parallellform des IST durchgeführt. Die statistische Datenanalyse zeigte zwar für alle Gruppen einen signifikanten Zeit- jedoch einen nichtsignifikanten Schereneffekt zwischen den Gruppen. Wir wollten nun Aufschluß darüber erhalten, ob die 3-stündige Trainingszeit zu kurz war. Dazu nahmen wir an, daß die Daten des Test-Retest-Kontrollgruppendesigns Realisationen eines multivariaten autoregressivén Prozesses 1.Ordnung waren. Kompliziertere Modelle (Chow, 1975) konnten wegen der Beschränkung auf zwei Meßzeitpunkte nicht formuliert werden. Die Prozessgleichung lautet:

$$(1a) \quad E[X(t)|X(t-1)] = A X(t-1) + \underline{b} \quad \text{od. für die Mittelwerte} \quad (1b) \quad \underline{\mu}(t) = A \underline{\mu}(t-1) + \underline{b}$$

Die Parametermatrix A spiegelt die Einflüsse der Pretests auf die Posttests wieder. Sie stellt also die zeitdiskrete Transferstruktur dar. Dagegen repräsentiert der Parametervektor b die Stärke äußerer Einflüsse auf die Variablen X(t). Ähnliche Modelle finden sich bei Steyer (Anm.2). A und b wurden mit LISREL geschätzt. Jede Zeile von A und b enthält die Regressionskoeffizienten eines IST-Posttests auf alle Pretests. Während in der Gruppe A jeweils der zum Posttest parallele Pretest der beste Einzelprediktor war, "zerfaserte" sich die Matrix A in

den Gruppen B und C. Die Hauptdiagonale von \underline{A} enthielt nicht mehr in jedem Fall den größten Regressionskoeffizienten. Daher scheint nur durch das Lernen von Denkstrategien (in Gruppe A) eine Stabilisierung der Testvariablen im korrelativen Sinne eingetreten zu sein. Zur Simulation der erwarteten Entwicklungsverläufe wurde die Lösung des Differenzgleichungssystems

$$(2) \underline{\mu}(t) = \underline{A}^t \underline{\mu}(0) + \left[\sum_{i=0}^{t-1} \underline{A}^i \right] \underline{b} = \left\{ \begin{array}{l} \text{endogene Entwicklung} \\ \text{seit } t=0 \end{array} \right\} + \left\{ \begin{array}{l} \text{exogener Einfluß zwischen} \\ t=0 \text{ und } t-1 \end{array} \right\}$$

berechnet. Nach den Ergebnissen der Simulation würden wir deutliche Unterschiede zwischen den Gruppen erst im Laufe der Zeit (frühestens nach 2 Monaten bzw. 8 mal 3 Stunden Training) erwarten. Während in der Gruppe A stetige Verbesserungen zu vermuten wären, würden sich in der Kontrollgruppe 3 der 8 trainierten Tests nach anfänglicher Verbesserung wieder auf dem alten Niveau einpendeln. Wir vermuten daher, daß das Coaching effektiv (wenn auch verbesserungswürdig) aber zeitlich zu kurz dosiert war, um einen signifikanten Schereneffekt herauszubilden.

Um diese Vermutung zu unterstützen, kreuzvalidierten wir diese Ergebnisse an einer Längsschnittstudie (Amelang & Hoppensack, 1977; Amelang & Möbus, Anm.3), in der 1968, 1973¹⁾ und 1978²⁾ Hamburger Studenten untersucht wurden. Zwischen 1968 und 1973 lag bei allen Pbn das Studium. Rangkorreliert man die Mittelwertsdifferenzen der 6 IST-Subtests, die sowohl in der Längsschnittstudie als auch im kurzfristigen Coaching erhoben worden waren, bewegen sich alle Korrelationen im nichtsignifikanten Bereich. Erwartet man dagegen vom Studium einen Anregungseffekt auf den IST, der ebenso durch ein längeres Coaching erzielt werden kann, liegt es nahe, die Simulation heranzuziehen. Bringt man die Tests in eine Rangreihe, nach der sie in der Simulation die 20-Punkte-Lösungsgrenze (bis auf GE haben alle Subtests nur 20 Aufgaben) erreicht haben und korreliert die "Beeinflussbarkeits"rangreihe mit den Rangreihen der Mittelwertveränderungen der Längsschnittstudie, erhält man ein anderes Ergebnis der Kreuzvalidierung (Tab.1). Es scheinen Studium und ein längeres Training einen ähnlichen Anregungseffekt auf den IST zu haben. Nach der Simulation müßte das Training eine Dauer von ca. 13 mal 3 Stunden besitzen, um kreuzvalidierbare Ergebnisse zu liefern. Einfache Kreuzvalidierungen ohne Berücksichtigung von Zeit- und Dosierungsaspekten müssen scheitern oder sind zumindest Glückstreffer.

Tab.1: Rangkorrelationen zwischen den Mittelwertsdifferenzen der IST-Subtests in der Längsschnittstudie und den zwei Beeinflussbarkeitsmaßen im kurzzeitigen Trainingsprogramm (Maß 1: Mittelwertsdifferenzen zwischen Post- und Pretest; Maß 2: auf Grund der Simulation gehegte Erwartung, wie schnell alle Personen einen Untertest lösen=Schnelligkeit mit der ein Subtest in der Simulation die 20-Punkte Decke erreicht)

	Trainingsgruppe A		Kontrollgruppe C			
	Maß 1	Maß 2	Maß 1	Maß 2		
Längsschnittstudie	1968/73	.143	.885(s)	.442	-.142	(s)=signifikant auf .05 bei einseitiger Hyp.
	1973/78	-.085	-.086	.071	.085	

2.2 Zeitkontinuierliche Modelle und Überlegungen zum "Optimierungsproblem"

Zeitdiskrete Modelle erlauben Prognosen nur zu festen Zeitpunkten. Besonders störend ist das bei Untersuchungen mit großen Abständen zwischen den Messungen. So sind bei der Studie von Amelang et al. mit zeitdiskreten Modellen nur Prognosen im 5-Jahres-Rhythmus möglich. Da aber die Zeit keine "Löcher" besitzt, benötigen wir zeitkontinuierliche Modelle. Ein weiterer Grund für diese Forderung liegt in der Suche nach Parameterinvarianz. So haben Post-Pretestregressi-

onen die unangenehme Eigenschaft, je nach Zeitabstand zwischen den Messungen andere Parameterwerte anzunehmen, obwohl der zeitkontinuierliche Prozess unverändert bleibt. Optimierung der Intervention würde nun bedeuten, daß man die pauschal im Vektor \underline{b} zusammengefaßten exogenen Einflüsse aufspalten könnte in kontrollierbare und nichtkontrollierbare Variable. Dann ließen sich die Ergebnisse von Intervention und Modellprognosen über beliebige Zeiträume vergleichen. Die Schätzung der Systemparameter erfolgt entweder in linearisierter Approximation des Prozesses (über Differenzen und Mittelwerte) mit LISREL oder direkt mit nichtlinearer Regression via numerische Integration (Möbus & Nagl, im Druck). Das schwierigste Problem bereitet zur Zeit die Parameteridentifikation. Dazu müssen nach der LAPLACE-Transformation des Systems symbolisch Polynommatrizen hergeleitet werden. Anschließend muß ebenfalls nichtnumerisch die Lösbarkeit oder Überbestimmtheit eines nichtlinearen Gleichungssystems geprüft werden (Möbus & Nagl, im Druck). Zur Zeit läßt sich dieser Ansatz noch nicht zur Prüfung größerer Theorien verwenden, obwohl er wegen der Formulierung zeitbezogener Hypothesen attraktiv ist.

-
- 1) Die Untersuchung der 1.u.2.Welle wurde von der Stiftung Volkswagenwerk gefördert(AZ132558)
 - 2) Die Untersuchung der 3.Welle wurde ebenfalls von der Stiftung Volkswagenwerk gef.(AZ34291)
- Anm.1: Hamouzova,M.& Würthner,K., Der Einfluß zweier Trainingsmethoden auf die Intelligenztestleistung Erwachsener, unveröffentl.Dipl.arbeit am Psych.Inst.d.Uni Heidelberg 1976
- Anm.2: Steyer,R., Stochastische Prozesse und kausale Abhängigkeit, Vortrag auf dem XXII.Int. Kongreß für Psychologie, Leipzig, 6.-12.Juli,1980
- Anm.3: Amelang,M.& Möbus,C., Persönlichkeitsstruktur und Berufstätigkeit, Bericht an die Stiftung Volkswagenwerk (in Vorbereitung)

PROZESSANALYSE IN DER PSYCHOLOGISCHEN FORSCHUNG

Friedhelm Meier

Summary (Process analysis in psychological research): The issues of this discussion group had been the procedure and realization of time series analysis in various topics of psychological research, e.g. clinical or educational psychology.

Die quantitativen Methoden des Gruppenvergleichs sind zu einem festen Bestandteil in der psychologischen Forschung geworden. Sie haben aber zugleich die Fortentwicklung von Methoden behindert, die den Zeitverlauf eines Merkmals als kontinuierlichen Prozeß abbilden. Nicht zuletzt aufgrund der Erfordernisse der klinisch-psychologischen Therapiefor schung entstand aber in den vergangenen fünf Jahren eine Diskussion prozeßbezogener Methoden der Zeitreihenanalyse, die in anderen Wissenschaftsdisziplinen wie z.B. den Ingenieur- oder Wirtschaftswissenschaften schon länger genutzt werden. Im Mittelpunkt des Interesses stehen dabei einerseits die Spektralanalyse und deren verwandte Auswertungsmethoden im Frequenzbereich und andererseits die Zeitreihenanalyse nach dem ARIMA - Modell von Box & Jenkins (1970) im Zeitbereich. Der letztgenannte Analyseansatz scheint eine hohe Applikabilität für stochastische und somit psychologische Merkmalsprozesse aufzuweisen.

Bericht über den 32. Kongreß der Deutschen Gesellschaft für Psychologie in Zürich 1980

Band 1

**Bericht des Präsidenten, Sondervorträge, Metatheorie,
Methodologie, Grundlagen**

**Im Auftrag der Deutschen Gesellschaft für Psychologie
herausgegeben von**

Wolfgang Michaelis

**Institut für Psychologie
der Christian-Albrechts-Universität Kiel**

1981

**Verlag für Psychologie · Dr. C.J. Hogrefe
Göttingen · Toronto · Zürich**

Wolfgang Michaelis wurde am 20. 7. 39 in Itzehoe/Holstein geboren. Nach Humanistischem Gymnasium, verlängertem Wehrdienst und Studium der Klassischen Altertumswissenschaften wandte er sich 1964 der Psychologie zu, für die er 1967 das Diplom erwarb, gefolgt 1969 von der Promotion und 1977 von der Habilitation.

Ab 1967 gastierte er an der University of London, fertigte dort bei *Eysenck* eine experimentelle Dissertation an über Persönlichkeit und Konzeptbildung, arbeitete anschließend ab 1969 in einem Forschungsinstitut der Bundesmarine (Diagnostik und Human Engineering) und wechselte im Winter-Semester 1970/71 als wissenschaftlicher Assistent in den Universitätsbereich.

Michaelis steht dem Wissenschaftsbetrieb in der Psychologie sehr kritisch gegenüber. Anders jedoch als in der sog. Kritischen Psychologie ist seine Opposition nicht sozial-, sondern wissenschaftspolitisch begründet. Seine Interessenschwerpunkte liegen in Integrationsbemühungen in der (allgemeinen) Psychologie (u. a. den Lerntheorien) und in der nicht-kognitiven Steuerung des Psychoorganismus.

Bände 1 und 2 zusammen: ISBN 3 8017 0174 3

Band 1: ISBN 3 8017 0175 1

Band 2: ISBN 3 8017 0176 X

Alle Rechte, insbesondere das der Übersetzung in fremde Sprachen, vorbehalten.

© by Verlag für Psychologie · Dr. C.J. Hogrefe, Göttingen 1981

Printed in Germany

Herstellung: Dieterichsche Universitäts-Buchdruckerei W. Fr. Kaestner
Rosdorf/Kr. Göttingen