

1

Vol. 2

Psychologie  
des Alltagshandelns

---

*Psychology  
of Everyday Activity*

**Editor**  
P. Sachse

JOURNAL

**JOURNAL**  
**Psychologie des Alltagshandelns**  
*Psychology of Everyday Activity*  
Vol. 2 / No. 1, February 2009  
ISSN 1998-9970  
**innsbruck university press**

## Impressum

### Herausgeber / Editor

Pierre Sachse, Innsbruck (A)

### Redaktionsassistent / Editorial Assistant

Thomas Höge, Innsbruck (A)

### Mitherausgeber / Associate Editors

Dietrich Dörner, Bamberg (D)

Winfried Hacker, Dresden (D)

Hartmann Hinterhuber, Innsbruck (A)

Oswald Huber, Fribourg (CH)

Wolfgang G. Weber, Innsbruck (A)

Eberhard Ulich, Zürich (CH)

### Beirat / Advisory Board

Petra Badke-Schaub, Delft (NL)

Claudia M. Eckert, Cambridge (GB)

Jürgen Glaser, München (D)

Rüdiger von der Weth, Dresden (D)

Momme von Sydow, Göttingen (D)

Anton Wäfler, Olten (CH)

### Verlag / Publisher

*iup – innsbruck* university press (A)

[www.uibk.ac.at/iup](http://www.uibk.ac.at/iup)

### Grafisches Konzept / Art Direction

*iup – innsbruck* university press (A)

### Gestaltung / Layout

Carmen Drolshagen, Innsbruck (A)

### Herstellung / Produced

Fred Steiner, Rinn

Alle Rechte vorbehalten. / All rights reserved.

ISSN 1998-9970

# Inhalt

Zielklärung und Handlungsplanung beim Problemlösen in Gruppen .....	2
<i>Andreas Müller</i>	
The influence of a problem solving training on shared mental models of spatial planners .....	14
<i>Rinat Saifoulline, Rüdiger von der Weth, Walter Schönwandt, Christoph Hemberger and Jens-Peter Grunau</i>	
Employment Prospects of Temporary and Permanent Workers: Associations with Well-being and Work Related Attitudes .....	22
<i>Thomas Rigotti, Nele De Cuyper, Hans De Witte, Sabine Korek and Gisela Mohr</i>	
Kritik am Dogma „Simple Structure“ .....	36
<i>Suibert Ertel</i>	

# Zielklärung und Handlungsplanung beim Problemlösen in Gruppen<sup>1</sup>

Andreas Müller

Ludwig-Maximilians-Universität München / Institut und Poliklinik für Arbeits-, Sozial- und Umweltmedizin

## ZUSAMMENFASSUNG

Am Beispiel des Konstruierens technischer Produkte werden erfolgskritische Faktoren für die gemeinsame Bearbeitung komplexer Probleme in Gruppen handlungstheoretisch abgeleitet und empirisch überprüft. Kennzeichen von Konstruktionsaufgaben ist, dass Aufgabenziele zu Beginn meist nur abstrakt bzw. „lückenhaft umschrieben“ sind. Entsprechend wird davon ausgegangen, dass erfolgreiches Gruppenhandeln beim Konstruieren eine fortwährende Klärung der Aufgabenziele sowie der kooperativen Ausführung von Maßnahmen zur Realisierung dieser Ziele (Handlungsplanung) erfordert. Zur Prüfung dieser Annahmen wurden in einer kontrollierten Laboruntersuchung systematische Verhaltensbeobachtungen von N = 20 Gruppen bei der gemeinsamen Bearbeitung einer Konstruktionsaufgabe durchgeführt. Keine Unterschiede zwischen erfolgreichen und weniger erfolgreichen Gruppen wurden hinsichtlich des zeitlichen Anteils und der Häufigkeit der Zielklärung beobachtet. Erfolgreiche Gruppen zeichnen sich jedoch durch häufigere Handlungsplanungen sowie durch systematische rückkoppelnde Wechsel zwischen Zielklärungs- und Handlungsplanungsphasen aus. Die Ergebnisse werden kritisch diskutiert und Schlussfolgerungen für die Unterstützung des Gruppenhandels bei der Bearbeitung komplexer Probleme abgeleitet.

## Schlüsselwörter

Gruppe – Handlungsregulation – Problemlösen

## ABSTRACT

Goal-clarification and action-planning during problem-solving in groups

On the basis of action-regulation theory promising factors for collective problem-solving strategies will be derived and empirically tested on the example of engineering design. From perspective of action-regulation the main quality of engineering design tasks is the lack of knowledge about task-goals and routine strategies. Therefore, it is argued that successful group-action needs a clarification of task-goals and of the cooperative strategies to achieve these goals (action-planning). In a controlled study N = 20 groups performed an engineering design task. Systematic observations of the video-taped problem-solving process were carried out. Against the hypotheses, no differences between successful and less successful groups regarding the time and frequency of goal-clarification were observed. Successful groups can be characterized by systematic feedback loops between phases of goal-clarification and action-planning. Results will be critically discussed and implications for the support of group-action during complex problem-solving will be derived.

## Keywords

group – action regulation – problem solving

---

1 Die Untersuchung entstand im Rahmen des Forschungsprojektes „Implicit knowledge in the product innovation process“, welches durch die Initiative „Innovationsprozesse in Wirtschaft und Gesellschaft“ der VolkswagenStiftung finanziell gefördert wurde. Der Dank des Autors gilt den Projektkolleginnen und Projektkollegen Frau Dr. Britta Herbig, Frau Dipl.-Psych. Kostanija Petrovic, Frau Dipl.-Ing. Nadja Pecquet, Herrn Dr.-Ing. Christoph Jung und Herrn Dipl.-Ing. Martin Graebisch.

## 1 Einleitung

Ziel dieses Beitrags ist es, erfolgskritische Faktoren für die Bearbeitung komplexer Probleme in Gruppen handlungstheoretisch abzuleiten und empirisch zu prüfen. Mit anderen Worten, es soll die Frage behandelt werden, *was* Gruppen, *wann* und *wie* bei der gemeinsamen Problemlösung tun sollten.

Beispielhaft soll dies anhand der Tätigkeit des Konstruierens geschehen. Konstruieren ist „denkendes Entwerfen des noch nicht Gegebenen“ (Hacker, 1996, S. 111) bei der Entwicklung eines technischen Gegenstandes, seiner Bauteile und Baugruppen (Pahl, Beitz, Feldhusen & Grote, 2005). Ein wesentliches Merkmal von Konstruktionsaufgaben ist es, dass Anforderungen zu Beginn entweder nur „*lückenhaft umschrieben*“ (Hacker, 2005) oder nur *abstrakt bzw. gar nicht definiert* sind (Dörner, 1999). Konstruieren erfordert damit die Orientierung innerhalb eines komplexen und zunächst noch unklar definierten Problemraums (vgl. Dörner, 1999) und ist demzufolge prototypisches Beispiel für ein durch Denktätigkeit zu lösendes komplexes Problem.

Konstruieren kann durch die Zusammenarbeit in Gruppen erleichtert werden, da Gruppen über eine potentiell größere Informationsverarbeitungskapazität sowie potentiell umfangreicheres Wissen als Individuen verfügen (Hinsz, 1990; Hinsz, Tindale & Vollrath, 1997) und demzufolge komplexe Probleme besser handhaben sollten. Jedoch liegt die Gruppenleistung bei der Bearbeitung unterschiedlichster Aufgaben, insbesondere aufgrund von Motivations- und Koordinationsverlusten, fast immer unter der – aufgrund der Einzelfähigkeiten der Gruppenmitglieder zu erwartenden – optimal möglichen Leistung (Steiner, 1972). Eine Leistungssteigerung bei der Bearbeitung komplexer Probleme (bzw. von Konstruktionsaufgaben) durch Gruppenarbeit kann also nicht ohne weiteres vorausgesetzt werden. Erfolgskritische Faktoren sind demnach zu überprüfen.

## 2 Theorie

„Was“ sollten Gruppen „wann“ bei der Problemlösung tun?

Da Problemlösen Denken erfordert, erscheint es für die Beantwortung unserer Fragestellung hilfreich, Gruppen als durch Informationsverarbeitung gesteuerte, handelnde Systeme aufzufassen (vgl. Arrow, McGrath & Berdahl, 2000; Hinsz et al., 1997; Tschan, 2000; von Cranach, Ochsenbein & Valach, 1986). Gruppenhandeln ist nach dieser Auffassung zweistufiges Handeln mit individuellen Prozessen auf der unteren Stufe und Gruppenprozessen auf der oberen Stufe (Larson

& Christensen, 1993; Tschan, 2000): Gruppenhandeln setzt demnach individuelle Prozesse des Verstehens und Bewertens voraus. Zusätzlich müssen die Resultate dieser individuellen Informationsverarbeitung durch Kommunikation in den Gruppenkontext eingebracht werden, um ein gemeinsames Gruppenhandeln zu ermöglichen. Wesentlicher Erfolgsfaktor für ein effizientes Gruppenhandeln ist somit ein gemeinsames geteiltes mentales Modell der Gruppenmitglieder über die Arbeitsaufgabe (Cannon-Bowers, Salas, Converse & Castellan, 1993; Klimoski & Mohammed, 1994; Mohammed & Dumville, 2001). Es erlaubt eine gemeinsame Orientierung der Gruppenmitglieder, unterstützt die Integration von Individualwissen in den Gruppenprozess und trägt gleichzeitig zur Herausbildung von Wissen auf Gruppenebene bei. Gruppenhandeln ist demzufolge komplexer als individuelles Handeln, da es die Koordination zwischen Gruppenmitgliedern für die Ausführung der gemeinsamen Handlung voraussetzt (vgl. Tschan, 2000).

Aus Sicht der Handlungstheorie kann Gruppenhandeln als zielgerichtetes Handeln aufgefasst werden (Tschan, 2000; von Cranach, Ochsenbein & Tschan, 1987; von Cranach et al., 1986). Ziele werden dabei definiert als die bewusste gedankliche Vorwegnahme eines angestrebten Handlungsergebnisses einschließlich der Vornahme zu dessen Verwirklichung (Hacker, 2005). Ziele mobilisieren die Anstrengungsbereitschaft der Gruppenmitglieder, richten deren Aufmerksamkeit und steigern die Hartnäckigkeit sowie die Entwicklung angemessener Handlungsstrategien (vgl. Wegge, 2004). Dieser positive Einfluss von Gruppenzielen auf die Gruppenleistung ist empirisch vielfach belegt (für einen Überblick vgl. O’Leary-Kelly, Martocchio & Frink, 1994).

Da für Problemlösungen kaum Handlungsrouinen zur Verfügung stehen (vgl. Duncker, 1955/1974) sind gemeinsame Ziele bei der Bearbeitung komplexer Probleme in Gruppen nicht ausreichend. Gruppen sollten also bei der Problembearbeitung zusätzlich zur Zielklärung geeignete *Maßnahmen* für die Zielerreichung festlegen, anpassen oder neu entwickeln. Der Vorgang der Verknüpfung einer Zielstruktur mit bedingungsangemessenen Maßnahmen zur Realisierung dieser Ziele sowie die Festlegung einer zweckmäßigen Abfolge dieser Maßnahmen wird als *Handlungsplanung* bezeichnet (Dörner, 2001).

Zielklärung und Handlungsplanung haben demnach bei der gemeinsamen Problemlösung in Gruppen eine wichtige handlungsregulative Funktion. Es stellt sich nun die Frage, *wann* eine Zielklärung und Handlungsplanung erfolgen sollte: Gruppenhandeln kann als Abfolge qualitativ unterschiedlicher Regulationsvorgänge beschrieben werden (Tschan, 2000). Die Kontrolle der Handlungsausführung erfordert rück-

koppelnde Vergleiche mit Zielen und Handlungsplänen (Miller, Galanter & Pribram, 1960). Dies ermöglicht die Einordnung aktueller Handlungen in ein Gesamtschema der Aufgabe, die Erkennung von unzulässigen Abweichungen des Handlungsvollzugs und eine Abschätzung des angemessenen Zeitaufwandes für Handlungen zur Realisierung eines Teilziels (vgl. Hacker, 2005). Zielklärungen und Handlungsplanungen sollten damit also möglichst *frühzeitig* stattfinden, da dies die Ausführung aller weiteren Handlungsschritte erleichtert.

Es ist nun zu klären, *wie* Gruppen bei der gemeinsamen Bearbeitung von Problemen gemeinsame Ziele und Handlungspläne erarbeiten sollten.

*„Wie“ sollten Gruppen bei der Problemlösung Ziele und Handlungspläne entwickeln?*

„Klassische“ handlungstheoretische Annahmen gehen davon aus, dass Handlungen durch die *systematische* Zergliederung eines Oberziels in „behandelbare“ untergeordnete Teilziele organisiert und dass diese Teilziele dann sequenziell „abgearbeitet“ werden (hierarchisch-sequenzielle Handlungsorganisation; Hacker, 2005; Miller et al., 1960; Volpert, 1982). Ähnliche Auffassungen finden sich auch in handlungstheoretischen Konzeptionen des Gruppenhandelns (Tschan, 2000; von Cranach et al., 1987; von Cranach et al., 1986).

Insbesondere bei der „Behandlung“ komplexer Probleme muss die Annahme der vollständigen systematischen Planbarkeit des menschlichen Handelns jedoch in Frage gestellt werden (vgl. hierzu z.B. Greif, 1994). Ein durchweg systematisches Vorgehen würde neben der Kenntnis der genauen Zielstruktur der Aufgabe eine vollständige Durchmusterung aller als aufgabenrelevant erscheinenden Maßnahmen, ein Behalten dieser Maßnahmen im Gedächtnis, eine gedankliche Aneinanderreihung aller denkbaren Maßnahmenfolgen, eine Auswahl aus diesen Maßnahmenfolgen aufgrund ausgewählter Kriterien und eine Prüfung der Vereinbarkeit der ausgewählten Maßnahmenabfolgen erfordern (vgl. Hacker, 2002). Diesen Anforderungen steht jedoch zum einen die begrenzte Verarbeitungskapazität unseres Arbeitsgedächtnisses entgegen (vgl. z.B. Cowan, 2000), die eine angemessene Verarbeitung einer solchen Menge an Informationen nicht gestattet. Bedenkt man weiterhin, dass in Gruppen zusätzliche Zeit für eine kommunikative Verständigung hinsichtlich dieser Ziele und Maßnahmen sowie die günstigste Form der kooperativen Ausführung dieser Maßnahmen vonnöten ist, werden die Grenzen der hierarchisch-sequenziellen Organisation des Gruppenhandelns offenbar. Entsprechend wurde in den wenigen vorliegenden Studien zum kollektiven Konstruktionshandeln beobachtet, dass Gruppen nur wenig systematisch – also nicht teilaufgaben- und arbeitsschritt-

rein – vorgingen (Cross & Cross, 1995; Olson, Olson, Carter & Storrøsten, 1992; Stempfle & Badke-Schaub, 2002). Es wurde vielmehr ein unsystematischer iterativer Wechsel zwischen Phasen, in denen entwickelte Lösungen diskutiert und bewertet wurden, mit Phasen, die sich auf die Planung des Gruppenprozesses bezogen, beobachtet.

Der wesentliche „Antreiber“ eines solchen Vorgehens ist das „Prinzip der kognitiven Ökonomie“ (Hacker, 2002): Es wird vermutet, dass vorhandene mentale Modelle „vermutungsgeleitet“ (Hacker, 2005, S. 560) auf neue Probleme übertragen werden, ohne dass der gesamte Problemraum systematisch zerlegt wird. Diese Nutzung vorhandener mentaler Modelle gestattet zunächst die Wahrnehmung und sinnvolle Gruppierung problemrelevanter Informationen sowie die Identifizierung von ersten Handlungsgelegenheiten (opportunities; vgl. Hacker, 2002, 2003; Visser, 1994). Beispielsweise kann am Beginn der Neuentwicklung einer schwenkbaren Schreibtischlampe die Idee entstehen, ein bereits bekanntes „Drehkran-Prinzip“ aufzugreifen. Eine solche Idee wird die weitere Informationssuche und -erkennung leiten, das heißt, es erscheinen zunächst nur diejenigen Informationen relevant, die der Verwirklichung dieses Prinzips dienen. Hiervon ausgehend werden provisorische (Teil-) Lösungen für das Prinzip erzeugt und diese dann in wiederkehrenden Schritten verändert und beurteilt (vgl. z.B. Ullmann, Dittrich & Stauffer, 1988). Hierbei kann deutlich werden, dass durch Verwirklichung einer technischen Funktion eine andere wichtige technische Funktion nicht mehr fehlerfrei ausgeführt werden kann. Es sind also neue Lösungen zu entwickeln, die diesen Widerspruch beseitigen. Problemlösen kann demnach als fortwährende Erzeugung und nachfolgende Beseitigung von Widersprüchen beschrieben werden (Dörner, 1987). Mit anderen Worten, es entstehen beim Konstruieren (Problemlösen) fortwährend neue Informationen, die auf ihre Handlungsrelevanz eingeschätzt werden müssen. Erst im Verlaufe der Problemlösung werden damit Ziele sowie Wechselwirkungen zwischen Zielen einer Aufgabe offensichtlich. Eine „iterative Zielentwicklung unterwegs“ (Hacker, 2002, S. 15) also auch eine iterative Handlungsplanung sind damit wesentliche Inhalte des Konstruierens selbst. Dies macht eine ständige Abstimmung der Gruppenmitglieder erforderlich.

Hinsichtlich der Beantwortung der Frage, *wie* Gruppen bei Konstruktionsaufgaben eine gemeinsame Ziel- und Maßnahmenstruktur erarbeiten sollten, ist also eine unterstützende Wirkung *wiederholter* (iterativer) Zielklärungs- und Handlungsplanungsphasen zu erwarten. Da nach handlungstheoretischem Verständnis Maßnahmen der Verwirklichung der Handlungsziele dienen, sollte *eine fortwährende Einbezie-*

hung der Aufgabenziele in die Handlungsplanung den Gruppenerfolg positiv beeinflussen.

## Hypothesen

Aus den theoretischen Überlegungen sind folgende Hypothesen abzuleiten:

- H1: Erfolgreiche Gruppen unterscheiden sich von weniger erfolgreichen Gruppen durch
- einen höheren zeitlichen Anteil und
  - eine größere Anzahl von Zielklärungs-Phasen zu Beginn des Problemlöseprozesses.
- H2: Erfolgreiche Gruppen unterscheiden sich von weniger erfolgreichen Gruppen durch
- einen höheren zeitlichen Anteil und
  - eine größere Anzahl von Handlungsplanungs-Phasen zu Beginn des Problemlöseprozesses.
- H3: Erfolgreiche Gruppen unterscheiden sich von weniger erfolgreichen Gruppen durch ein häufigeres und absichtvolles rückkoppelndes Hin- und Herwechseln zwischen Zielklärungsphasen und Handlungsplanungsphasen.

## 3 Methode

### 3.1 Stichprobe

Die Untersuchung wurde in Veranstaltungen des Lehrstuhls für Produktentwicklung der TU München vorgestellt. Interessenten konnten sich in eine Terminliste eintragen. An der Untersuchung nahmen 60 Studierende der Ingenieurwissenschaften der TU München teil (weiblich = 5; Alter:  $M = 25.6$  Jahre;  $SD = 1.5$ ), die zum Zeitpunkt der Untersuchung im Durchschnitt 6.8 Fachsemester ( $SD = 1.85$ ) besucht hatten. Die Teilnehmer wurden zu 20 Gruppen à drei Personen zusammengefasst.

### 3.2 Untersuchungsablauf

Das Experiment fand in einem Raum des Lehrstuhls für Produktentwicklung der TU München statt. Die Gruppenaufgabe bestand in der gemeinsamen Erstellung einer Konstruktions-skizze einer Schreibtischlampe. Die Konstruktionsaufgabe enthielt sowohl konkrete als auch abstrakte Anforderungen: Die Lichtquelle der Lampe sollte zum einen durch zwei rotatorische

Achsen am Lampenkopf und drei translatorische Achsen am Lampenarm (vgl. Abbildung 1) mit einer Hand einstellbar sein. Der Schwenkbereich um den Lampenbefestigungspunkt war ebenfalls vorgegeben. Zum anderen sollte das Konzept innovativ sowie einfach und kostengünstig zu fertigen sein. Es wurden Zeichenpapier, Stifte, Lineal, Präsentationsmaterialien und eine Metaplanwand bereitgestellt. Weitere Hilfsmittel waren nicht gestattet. Es standen 90 Minuten für die Aufgabenbearbeitung zur Verfügung. Die Aufgabenbearbeitung wurde mit Video aufgezeichnet.<sup>2</sup> Vor der Aufgabenbearbeitung wurde jedes Gruppenmitglied gebeten, einen Test über aufgabenrelevantes technisches Wissen und Methoden der Produktentwicklung zu bearbeiten. Die Teilnehmer verpflichteten sich für den Zeitraum der Gesamt-Untersuchung zum Stillschweigen über die Untersuchungsinhalte, damit nachfolgende Experimentalgruppen keine Vorteile durch Kenntnis der Aufgabenstellung erlangen konnten. Zur Minimierung von Untersuchungsleitereffekten waren sämtliche Instruktionen standardisiert. Die Teilnahme wurde mit 50 Euro pro Gruppenmitglied vergütet. Die Experimentalgruppe, welche die beste Lösung entwickelte, erhielt eine zusätzliche Prämie von 25 Euro pro Mitglied.

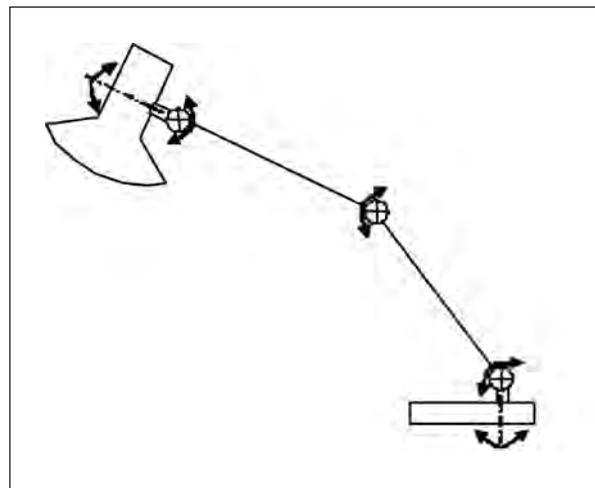


Abbildung 1: Schemazeichnung der geforderten Einstellmöglichkeiten der Schreibtischlampe

### 3.3 Erfassung des Gruppenhandelns

Zur Beurteilung des Gruppenhandelns erfolgten Analysen der Videoaufnahmen der Aufgabenbearbeitung. In Anlehnung an die Phasen der Handlungsregulation

<sup>2</sup> Die Hälfte der Gruppen wurde gebeten, die Aufgabenbearbeitung nach 45 Minuten zu unterbrechen und über ihre Aufgabe zu reflektieren. Die Ergebnisse dieser Intervention wurden bereits anderorts berichtet (Herbig et al., 2006; Müller, Herbig & Petrovic, 2006) und sind kein Gegenstand dieses Beitrages. Der Einfluss der Intervention auf die hier berichteten Ergebnisse wurde kontrolliert.

wurden hierbei folgende Beobachtungskategorien unterschieden: *Zielklärung (ZK)* umfasst die verbalen Kommunikationen bezüglich der Erfassung und Strukturierung, der in der Aufgabenstellung vorgegebenen Anforderungen, da das wesentliche Ziel des Konstruierens die Realisierung der Anforderungen eines technischen Konzepts ist. *Handlungsplanung (HP)* umfasst die verbale Kommunikation, mit der die Maßnahmen zur Verwirklichung der Ziele unter vorhandenen Bedingungen sowie deren kooperative Ausführung festgelegt werden. Weiterhin wurde unterschieden zwischen *Lösungssuche (LS)*, *Lösungsbewertung (LB)*, *Entscheidungen treffen (ET)*, *Anfertigung der Lösungsskizze (SKZ)* sowie zwischen zwei Phasen, in denen keine direkte handlungsbezogene Kommunikation erfolgte: *Sozio-emotionale Regulation (SOZ)*; z.B. gegenseitige Vorstellung der Gruppenmitglieder, Scherze, etc.) und *Fragen der Teilnehmer an Untersuchungsleiter (F TN)*. Eine ausführliche Beschreibung der Beobachtungskodes einschließlich verbaler und Verhaltensanker sowie Ausschlusskriterien findet sich in Müller (2007). Die Beobachtungseinheiten wurden begrenzt durch einen Wechsel der handlungsregulatorischen Funktion der Kommunikation (ereignisbasierte Kodierung).

Zur Einschätzung der Zuverlässigkeit der Beobachtungskategorien wurden acht der 20 Video-Aufnahmen unabhängig von jeweils zwei Beurteilern kodiert. Die so ermittelte Beurteilerübereinstimmung streut zwischen Werten von  $\kappa = .58 - .82$ . Die durchschnittliche Beurteilerübereinstimmung beträgt  $\kappa = .66$ . Damit kann von einer insgesamt guten Zuverlässigkeit des entwickelten Kategoriensystems ausgegangen werden (vgl. Fleiss & Cohen, 1973).

### 3.4 Erfassung des Gruppenerfolgs

Es erfolgten unabhängige Bewertungen sämtlicher Konzeptskizzen durch jeweils zwei Experten (Diplom-Ingenieure) anhand von drei Dimensionen: *Anforderungserfüllung* ist die invertierte Summe der nicht erfüllten Anforderungen, die in der Aufgabenstellung vorgegeben wurden. *Fehlerfreiheit* ist die invertierte Summe der den Lösungsskizzen zu entnehmenden Konstruktionsfehler. *Originalität* wurde in Anlehnung an eine von Lindemann und Baumberger (2004) vorgestellte Methode erfasst. Die Dimension wird auf einer fünfstufigen Skala eingeschätzt (1 = langweilig; 5 = originell). Aus den drei Dimensionen wurde durch Mittelwertbildung ein Gesamtscore „Gruppenerfolg“ berechnet.

Die Beurteilerübereinstimmung bei der Einschätzung der Dimension Anforderungserfüllung beträgt  $\rho = .89$ , bei Fehlerfreiheit  $\rho = .96$  und bei Originalität  $\rho = .78$ . Damit kann eine gute Zuverlässigkeit der Maße festgestellt werden. Für die weiteren Analysen wurden die Bewertungen des erfahreneren der beiden Beurteiler verwendet.

## 4 Ergebnisse

Eine Überprüfung ergab keine Hinweise auf Einflüsse zweier unterschiedlicher Untersuchungsleiter, des Erhebungszeitpunktes, der experimentellen Manipulation, der soziodemografischen Gruppenzusammensetzung (Diversität von Geschlecht, Alter, Anzahl der Fachsemester) und des expliziten aufgabenbezogenen Wissens auf den Zeitanteil und die Dauer und Häufigkeit der erfassten Handlungsphasen sowie auf den Gruppenerfolg. Weiterhin unterschieden sich erfolgreiche und weniger erfolgreiche Gruppen (Median-Split der Variable „Gruppenerfolg“) nicht signifikant bezüglich des Zeitanteils und der Häufigkeit der zusätzlich zu Zielklärung und Handlungsplanung erfassten Handlungsphasen.

### 4.1 Beschreibung des Gruppenhandelns

Tabelle 1 fasst die deskriptive Statistik des zeitlichen Anteils sowie der Anzahl der beobachteten Handlungsphasen zusammen. Etwas mehr als die Hälfte der Zeit verbrachten die Gruppen mit der Suche nach und Entwicklung von Lösungskonzepten, gefolgt von der Kommunikation über Aufgabenziele und der Bewertung der entwickelten Konzepte. Für die Zielklärung wurde etwa 14% für die Handlungsplanung etwas weniger als 10% der Zeit aufgewendet. Zwischen den Gruppen sind jedoch auch große Unterschiede zu registrieren. So beschäftigte sich eine Gruppe in nur 2% der zur Verfügung stehenden Zeit mit der Zielklärung während für die Gruppe mit dem höchsten Zeitanteil an Zielklärung 32.02% zu verzeichnen sind.

Durchschnittlich wurden 64.6 Phasen beobachtet (vgl. Tabelle 1). Das bedeutet, dass – bei einer Aufgabenbearbeitungszeit von 90 Minuten – etwa alle 1.4 Minuten ein Wechsel in der Handlungsregulation der Gruppen stattfand. Jedoch sind hier auch starke Schwankungen zwischen den Gruppen zu beobachten.

Abbildung 2 verdeutlicht beispielhaft den Verlauf des Handelns einer Gruppe bei der Aufgabenbearbeitung.

Tabelle 1: Deskriptive Statistik des zeitlichen Anteils sowie der Anzahl der beobachteten Handlungsphasen

Handlungsphasen	Zeitlicher Anteil (in %)				Anzahl			
	M	SD	MIN	MAX	M	SD	MIN	MAX
ZK	13.71	7.75	2.00	32.02	8.85	3.69	3	16
HP	9.05	4.99	.62	21.07	14.60	5.58	2	26
LS	51.85	8.88	35.92	72.65	18.70	5.39	9	28
LB	13.29	8.66	1.92	35.86	9.30	5.50	2	22
ET	1.37	1.34	.00	4.87	3.50	2.70	0	9
SKZ	9.52	6.17	.00	19.54	6.00	3.60	0	14
SOZ	.65	.88	.00	3.96	2.30	2.96	0	14
F TN	.56	.69	.00	2.51	1.35	1.46	0	5
<b>Gesamt</b>	<b>100</b>	<b>-</b>	<b>-</b>	<b>-</b>	<b>64.60</b>	<b>16.72</b>	<b>35</b>	<b>92</b>

Anmerkung:

N = 20. ZK = Zielklärung, HP = Handlungsplanung, LS = Lösungssuche, LB = Lösungsbewertung, ET = Entscheidungen treffen, SKZ = Skizzieren, SOZ = Sozio-Emotionale Regulation, F TN = Frage der Teilnehmer an die Versuchsleiter.

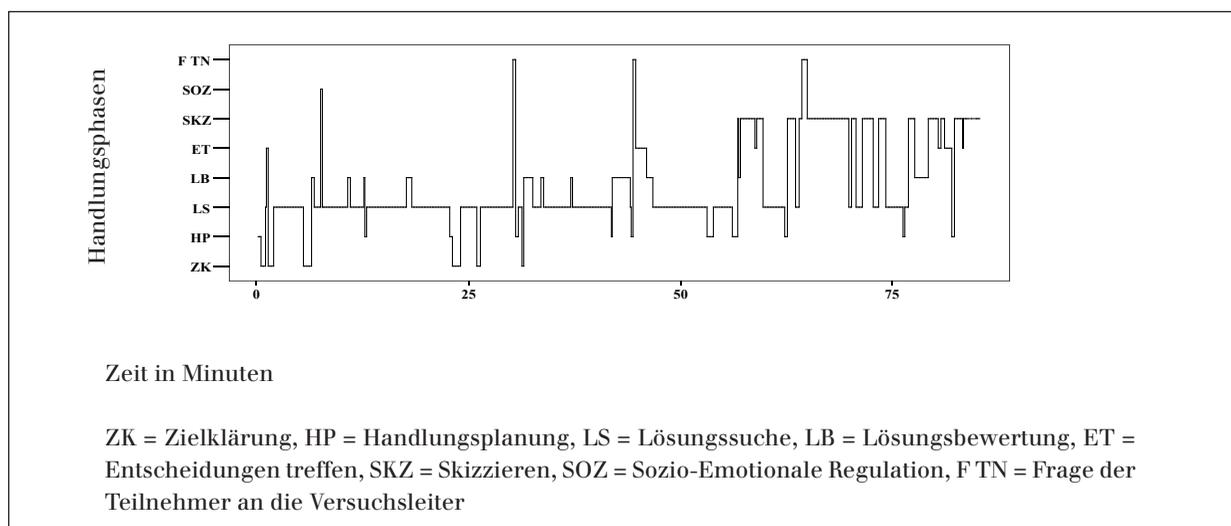


Abbildung 2: Beispiel des Handlungsverlaufs einer Gruppe

Nach einer kurzen Phase der Handlungsplanung und Zielklärung beginnt die Gruppe schnell mit der Lösungssuche. Diese wird jedoch immer wieder durch kurze Phasen der Zielklärung, Handlungsplanung aber auch Lösungsbewertung unterbrochen. In der zweiten Hälfte der Aufgabenbearbeitung sind überhaupt keine Zielklärungsphasen mehr zu beobachten. Etwa zu Anfang des letzten Drittels der Aufgabenbearbeitung beginnt die Gruppe mit der Anfertigung der Konzeptskizze. Die Fertigung der Skizze wird häufiger unterbrochen durch erneutes Lösungssuchen zum Teil auch Handlungsplanungs-, Lösungsbewertungs- und Entscheidungsphasen.

Das Beispiel verdeutlicht, dass die Gruppe – in Übereinstimmung mit den theoretischen Überlegungen – nicht „arbeitsschrittrein“ vorgeht, sondern stattdessen einzelne Handlungsphasen iterativ durchläuft.

### Statistische Hypothesentestung

Zur Überprüfung von *Hypothese 1* und *Hypothese 2* wurden erfolgreiche und weniger erfolgreiche Gruppen (Mediansplit der Variable Gruppenerfolg) hinsichtlich des aufgewendeten Zeitanteils und der Anzahl der Zielklärungs- (ZK) und Handlungsplanungsphasen (HP) in der ersten Hälfte der Aufgabenbearbeitung verglichen. Die Ergebnisse sind in Abbildung 3 grafisch dargestellt.

Entgegen den Annahmen wurde bei erfolgreichen im Vergleich zu weniger erfolgreichen Gruppen in der Tendenz ein – nicht signifikant – geringerer Zeitanteil sowie eine geringere Häufigkeit an Zielklärung beobachtet. Hingegen wenden erfolgreiche Gruppen in der ersten Hälfte der Aufgabenbearbeitung erwartungsgemäß mehr Zeit für Handlungsplanung auf als wenig erfolgreiche Gruppen (Mann-Whitney-U = 31;  $p \leq .10$ ;  $d = .66$ ) und teilten diese Zeit auf mehr Phasen

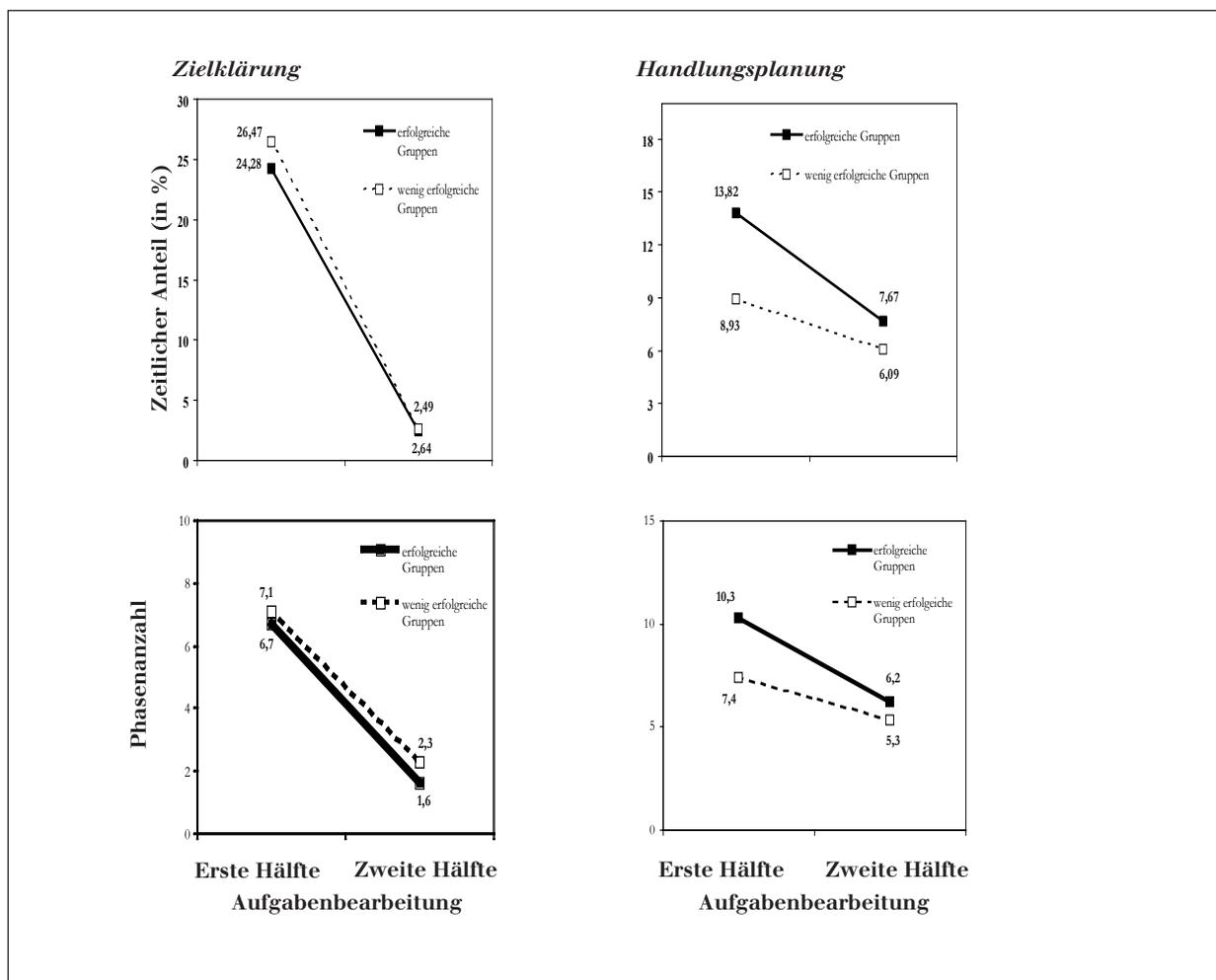


Abbildung 3: Gruppenerfolg und zeitlicher Anteil der Zielklärung und Handlungsplanung im Vergleich zwischen erster und zweiter Hälfte der Aufgabenbearbeitung

auf (Mann-Whitney-U =19.5;  $p \leq .05$ ;  $d = .89$ ). Die Effektstärke  $d$  deutet auf eine mittlere bzw. hohe Bedeutsamkeit dieser Gruppenunterschiede hin.

Damit ist Hypothese 1 zurückzuweisen während Hypothese 2 bestätigt werden kann. Aus Abbildung 3 ist weiterhin erkennbar, dass im zweiten Abschnitt der Aufgabenbearbeitung bei beiden Vergleichsgruppen eine rapide Abnahme an Zielklärung in nahezu gleichem Ausmaß zu beobachten ist. Hingegen ist bei erfolgreichen im Vergleich zu wenig erfolgreichen Gruppen auch in Teil 2 der Aufgabenbearbeitung eine tendenziell zeitlich umfassendere und häufigere Handlungsplanung zu beobachten.

Es wurde weiter angenommen, dass sich erfolgreiche von weniger erfolgreichen Gruppen durch häufigere Wechsel zwischen Zielklärungs- und Handlungsplanungs-Phasen unterscheiden (Hypothese 3). Hierzu wurden zunächst mit dem Mann-Whitney U-Test Vergleiche erfolgreicher und weniger erfolgreicher Gruppen hinsichtlich der Häufigkeit aufeinanderfolgender Handlungsphasen Zielklärung: Handlungsplanung bzw. Handlungsplanung: Zielklärung durchgeführt. Die Ergebnisse des Gruppenvergleichs sind Tabelle 2 zu entnehmen.

Bei erfolgreichen Gruppen sind signifikant mehr direkt aufeinanderfolgende Wechsel von Zielklärung zu Handlungsplanung zu beobachten als bei weniger erfolgreichen Gruppen. Die Effektstärken weisen auf eine hohe praktische Bedeutsamkeit dieses Unterschiedes hin. Ebenfalls sind bei erfolgreichen Gruppen häufigere umgekehrte Wechsel von HP auf Zielklärung zu beobachten. Die Differenz zu weniger erfolgreichen Gruppen ist jedoch nicht signifikant.

Zusätzlich wurde geprüft, ob Handlungsplanung mit größerer Wahrscheinlichkeit auftrat, wenn zuvor Zielklärung durchgeführt wurde und umgekehrt. Eine solche Prüfung ist von Bedeutung, da eine bloße Zufälligkeit des Auftretens beider Phasenpaare eine Interpretation im Sinne einer absichtsvollen Einbeziehung von Aufgabenzielen in die Handlungsplanung nur bedingt rechtfertigen würde. Es wurden getrennt für beide Vergleichsgruppen Zeitreihenanalysen mittels logit-loglinearer Modelle (Bakeman, Adamson & Strisik, 1995; Gottman & Roy, 1990) über alle aufgabenrelevanten Handlungsphasen (s. o.) durchgeführt. Sowohl im Falle der wenig erfolgreichen ( $\chi^2 = 50.59$ ,  $df = 19$ ,  $p = .00$ ) als auch der erfolgreichen Gruppen ( $\chi^2 = 104.49$ ,  $df = 19$ ,  $p = .00$ ) weicht ein sogenanntes Nullmodell, welches annimmt, dass nach dem Auftreten einer bestimmten Handlungsphase, jede andere Phase mit der gleichen Wahrscheinlichkeit auftreten kann, signifikant von den empirischen Daten ab. Daraus ist zu schließen, dass die Gruppen generell nicht völlig unsystematisch zwischen den Handlungsphasen hin und herwechselten.

Im zweiten Schritt wurde geprüft, *welche* der spezifischen Phasenpaare überzufällig auftraten. Hierfür wurde geprüft, inwieweit Residuen – also „Resthäufigkeiten“ – der einzelnen Phasenpaare signifikant vom Nullmodell abweichen (vgl. Bakeman et al., 1995).

Erfolgreiche Gruppen wechselten bei der Aufgabenbearbeitung sowohl überzufällig häufig von Zielklärung auf Handlungsplanung als auch von Handlungsplanung auf Zielklärung, während wenig erfolgreiche Gruppen nur überzufällig häufig von Handlungsplanung zu Zielklärung wechseln (vgl. Tabelle 3). Damit ist Hypothese 3 insgesamt bestätigt.

Tabelle 2: Unterschiede zwischen erfolgreichen und weniger erfolgreichen Gruppen hinsichtlich iterativer Wechsel zwischen Zielklärungs- und Handlungsplanungsphasen

	Gruppenerfolg					
	gering (n = 10)		hoch (n = 10)		P <sub>(exakt)</sub>	d
Phasenpaare	M	SD	M	SD		
ZK:HP	2.50	1.06	4.20	2.70	.04	.95
HP:ZK	2.90	1.45	4.60	2.95	.11	.75

Anmerkung:

Mann-Whitney U-Test. p(exakt): exakte einseitige Signifikanz. ZK = Zielklärung, HP = Handlungsplanung

Tabelle 3: Überprüfung systematischer iterativer Wechsel zwischen Zielklärungs- und Handlungsplanungsphasen

		$f_{obs}$	$f_{exp}$	$e$	$e_{stand}$
wenig erfolgreiche Gruppen (n = 10)	ZK:HP	23	18.17	4.83	1.13
	HP:ZK	29	19.81	9.19	2.06**
erfolgreiche Gruppen (n = 10)	ZK:HP	42	26.24	15.76	3.07**
	HP:ZK	46	27.75	18.25	3.46**

Anmerkung:

ZK = Zielklärung, HP = Handlungsplanung.  $f_{obs}$  = beobachtete Häufigkeiten;  $f_{exp}$  = erwartete Häufigkeiten;  $e$  = Residuen;  $e_{stand}$  = standardisierte Residuen. \* $p \leq .05$ ; \*\* $p \leq .01$ .

## 5 Diskussion

Ziel dieser Arbeit war die handlungstheoretische Ableitung und empirische Überprüfung erfolgskritischer Faktoren des Gruppenhandelns bei der gemeinsamen Bearbeitung komplexer Probleme. Komplexe Probleme sind durch abstrakte bzw. nur „lückenhaft umschriebene“ (Hacker, 2005) Handlungsziele und unzureichende Kenntnis von Maßnahmen-Routinen gekennzeichnet. Es wurde angenommen, dass Gruppen zu einem frühen Zeitpunkt der Problemlösung eine Klärung gemeinsamer geteilter Ziele sowie geeigneter Maßnahmen zu deren Verwirklichung vornehmen sollten (*Was* sollten Gruppen, *wann* bei der Problemlösung tun?). Weiterhin wurde vermutet, dass aufgrund der kognitiven und kooperativen Ökonomie die Zielklärung und Handlungsplanung bei der Bearbeitung komplexer Probleme nicht systematisch arbeitsschrittweise durchgeführt werden kann, sondern opportunistisch in wiederholten Schritten jedoch unter systematischer Einbeziehung von Zielen in die Handlungsplanung erfolgen sollte (*Wie* sollten Gruppen Probleme lösen?). Die empirische Überprüfung dieser Annahmen erfolgte durch Verhaltensbeobachtungen von  $N = 20$  Gruppen, die eine Konstruktionsaufgabe bearbeiteten.

Entgegen den Erwartungen führten erfolgreiche Gruppen zu Beginn der Aufgabenbearbeitung keine zeitlich umfassendere und häufigere Zielklärung durch. Auch im weiteren Verlauf der Aufgabenbearbeitung konnten diesbezüglich keine Unterschiede zwischen erfolgreichen und weniger erfolgreichen Gruppen beobachtet werden. Hingegen unterscheiden sich erfolgreiche von weniger erfolgreichen Gruppen hypothesenkonform durch zeitlich umfassendere und

häufigere Handlungsplanungen in der ersten Hälfte der Aufgabenbearbeitung. Tendenziell wurden entsprechende Unterschiede auch in der zweiten Hälfte der Aufgabenbearbeitung beobachtet. Besonders ausgeprägt ist der Unterschied zwischen den Vergleichsgruppen hinsichtlich der *Häufigkeit* Handlungsplanung. Das Ergebnis stützt die Annahme, das Problemlösen in Gruppen eine *fortwährende* Modifikation des Gruppenhandelns während des gesamten Aufgabenverlaufs und weniger eine langfristige Vorabplanung des Handelns erfordert. Ebenfalls bezogen erfolgreiche Gruppen erwartungsgemäß Ziele systematischer in die Handlungsplanungen ein. Aus den vorliegenden Ergebnissen ist daher nicht zu schließen, dass Zielklärung keine handlungsunterstützende Funktion für die erfolgreiche Problemlösung in Gruppen besitzt. Zielklärung scheint hierfür vielmehr eine notwendige, jedoch nicht hinreichende Bedingung zu sein. In komplexen Problemsituationen, in denen Routine-Maßnahmen nicht unmittelbar abrufbar sind, sind Maßnahmen unter Einbeziehung von Handlungszielen wiederholt kommunikativ zu klären, zu adaptieren oder neu zu entwickeln. Eine Zielklärung allein ist nicht ausreichend.

Für die gemeinsame Problemlösung in Gruppen ist daher zu empfehlen, reflektiert vorzugehen. West (1996) definiert ein solches Vorgehen als „the extent to which group members overtly reflect upon the group’s objectives, strategies and processes, and adapt them to current or anticipated endogenous or environmental circumstances“ (S. 559). Dieses Prinzip der Reflexion und Adaption entspricht der in dieser Arbeit bei der Lösung komplexer Probleme als erfolgskritisch angesehenen iterativen Zielklärung und Handlungsplanung. Gruppen sollten sich im Verlauf des Gruppenhandelns

wiederkehrend folgende oder ähnliche Fragen stellen und beantworten:

- Sind unsere Aufgabenziele klar definiert?
- Wenn ja, was sind unsere Aufgabenziele? Wenn nein, welche Ziele sind noch unklar und wie sind sie zu klären?
- Inwieweit ist unser gegenwärtiges Handeln geeignet, die Aufgabenziele zu realisieren? Müssen wir unser Vorgehen ändern, um unsere Aufgabenziele zu erreichen?

Es kann vermutet werden, dass durch eine fortwährende reflektierende Vergegenwärtigung von Aufgabenzielen und – wenn nötig – planerische Adaption von Maßnahmen, die Entwicklung einer umfassenderen, intellektuellen, gemeinsamen geteilten Repräsentation der Aufgabe gefördert, ein vorausschauendes Agieren statt Reagieren ermöglicht und damit die Kontrolle über das Gruppenhandeln gesteigert wird (vgl. Hacker, 2005; West, 1996). Dadurch sollten ebenfalls typische Fehlhandlungen bei der Bearbeitung komplexer Probleme – wie beispielsweise die Fixierung auf Teilziele und deren Behandlung als Endziele (vgl. Dörner, 2006) – vermieden werden.

In künftigen Forschungsarbeiten sollten bei der Betrachtung der Zielklärung und Handlungsplanung beim Problemlösen in Gruppen insbesondere folgende Aspekte Beachtung finden:

1. Es ist eine differenzierte Erfassung der *Qualität der Zielklärung* vorzunehmen, um zu prüfen, inwieweit qualitative Unterschiede zum Gruppenerfolg beitragen. So könnte zwischen dem Erkennen, dem Verstehen und der Vornahme der Erfüllung von Zielen unterschieden werden sowie die Klärung widersprüchlicher Beziehungen zwischen Zielen (vgl. hierzu Dörner, 2006) und die Vollständigkeit der Zielklärung – im Sinne einer vollständigen Anforderungskklärung (vgl. Hacker et al., 2002) – erfasst werden.
2. Es ist die *Übertragbarkeit der vorliegenden Ergebnisse auf Experten-Gruppen* zu prüfen. Bis dato existierende Befunde über das Vorgehen von Experten-Gruppen (vgl. Olson et al., 1992) weisen zwar auf keine grundsätzlichen Unterschiede zu dem in dieser Arbeit beobachteten Vorgehen von Studenten-Gruppen hin, jedoch besitzen Experten im Vergleich zu Novizen ein umfangreicheres und zugleich anders organisiertes Wissen über den jeweiligen Bereich der Expertise (z.B. Larkin, McDermott, Simon & Simon, 1980; vgl. auch Rouse & Morris, 1986). Unklar ist, wie sich diese Unterschiede auf Gruppenebene manifestieren.

3. Wünschenswert sind auch Analysen, die *längere Zeiträume der Kooperation von Gruppen bei der Bearbeitung komplexer Probleme erfassen*. So berichten Levesque, Wilson & Wholey (2001), dass der Grad der gemeinsamen und geteilten Wahrnehmung einer Gruppenaufgabe durch die Gruppenmitglieder im Zeitverlauf (nach einem, zwei und dreieinhalb Monaten) abnahm. Es ist zu vermuten, dass nach einer hinreichenden Erkundung des Problemraums das Gruppenhandeln nicht mehr fortwährend koordiniert werden muss, sondern eine längerfristige stabile arbeitsteilige Organisation des Gruppenhandelns erfolgen kann.
4. Es ist die *Rolle der Zielklärung und Handlungsplanung im organisationalen Alltag* – mit seinen vielfältigen zusätzlichen Einflussfaktoren für erfolgreiches Gruppenhandeln zu prüfen (siehe hierzu für den Bereich der Produktentwicklung die ausführliche Studie von Badke-Schaub und Frankenberger, 2004).

Mit der vorliegenden Arbeit wird kein Anspruch auf vollständige Darstellung erfolgskritischer Vorgehensmerkmale des gemeinsamen Problemlösens in Gruppen erhoben. Aus Darstellungen zum individuellen Konstruktionshandeln (z.B. Hacker, Wetzstein, & Römer, 2002) können weitere potentielle erfolgskritische Einflüsse abgeleitet werden: Vor allem zu nennen sind die Entwicklung verschiedener Lösungsalternativen im Vergleich zur ständigen Modifizierung einer Lösung (generierende vs. korrigierende Variation; Dylla, 1991); der Wechsel zwischen den Modalitäten der Lösungsentwicklung (sprachlich vs. symbolisch) sowie zwischen mentaler und externaler Repräsentation durch Skizzieren (vgl. z.B. Sachse, 2002; Sachse & Leinert, 2002) und das wiederholte rückkoppelnde Bewerten von Lösungen. Eine systematische Untersuchung des erfolgskritischen Einflusses solcher Faktoren auf das Problemlösen in Gruppen steht noch aus. Erste empirische Befunde zeigen jedoch, dass einige dieser Faktoren auch beim Gruppenhandeln von Bedeutung sind: So finden sich in einer Untersuchung von Müller (2007) erste empirische Hinweise für die handlungsunterstützende Funktion des gemeinsamen Skizzierens beim Konstruieren in Gruppen. Stempfle und Badke-Schaub (2002) beobachteten, dass rückkoppelnde Wechsel zwischen Lösungssuche und Lösungsbewertung – im Sinne der Erweiterung und der Einengung des Problemraums – auch in Gruppen bei der Problemlösung eine zentrale Rolle spielen.

Abschließend ist zu bemerken, dass komplexes Problemlösen im Arbeitskontext sich nicht nur auf so genannte „white-collar jobs“ beschränkt, zu denen der hier betrachtete Bereich der Produktentwicklung si-

cher gerechnet werden kann. In der sich wandelnden Arbeitswelt nehmen die Komplexität und die Neuartigkeit arbeitsseitiger Anforderungen in nahezu allen Bereichen zu. Die Übertragbarkeit der im Zusammenhang mit der Entwicklung technischer Produkte diskutierten erfolgskritischen Faktoren des Gruppenhandelns bei der Lösung komplexer Probleme auf andere Aufgabentypen ist daher ebenfalls in zukünftigen Studien zu prüfen.

## Literatur

- Arrow, H., McGrath, J. E. & Berdahl, J. L. (2000). *Small groups as complex systems. Formation, coordination, development, and adaptation*. Thousand Oaks: Sage.
- Badke-Schaub, P. & Frankenberger, E. (2004). *Management kritischer Situationen. Produktentwicklung erfolgreich gestalten*. Berlin: Springer.
- Bakeman, R., Adamson, L. B. & Strisik, P. (1995). Lags and Logs: Statistical Approches to Interaction (SPSS version). In J. M. Gottman (Eds.), *The Analysis of Change*, 279-308. Mahwah, NJ: Lawrence Erlbaum.
- Cannon-Bowers, J. A., Salas, E., Converse, S. & Castellan, N. J. (1993). Shared mental models in expert team decision making. In N. J. Castellan (Ed.), *Individual and group decision making: Current issues*, 221-246. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Cowan, N. (2000). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87-185.
- Cross, N. & Cross, A. C. (1995). Observations of teamwork and social processes in design. *Design Studies*, 16, 143-170.
- Dörner, D. (1987). *Problemlösen als Informationsverarbeitung*. Kohlhammer: Stuttgart.
- Dörner, D. (1999). Approaching design thinking research. *Design Studies*, 20, 407-415.
- Dörner, D. (2001). *Bauplan für eine Seele*. Reinbek bei Hamburg: Rowohlt Taschenbuch Verlag.
- Dörner, D. (2006). *Die Logik des Misslingens. Strategisches Denken in komplexen Situationen*. Reinbek bei Hamburg: Rowohlt Taschenbuch Verlag.
- Duncker, K. (1935/1974). *Zur Psychologie des produktiven Denkens*. Berlin: Springer.
- Dylla, N. (1991). *Denk- und Handlungsabläufe beim Konstruieren*. München: Hanser.
- Fleiss, J. L. & Cohen, J. (1973). The equivalence of weighted Kappa and the intraclass correlation coefficient as measures of reliability. *Educational and Psychological Measurement*, 33, 613-619.
- Gottman, J. M. & Roy, A. K. (1990). *Sequential Analysis. A Guide for Behavioral Researchers*. Cambridge: Cambridge University Press.
- Greif, S. (1994). Handlungstheorie und Selbsttheorie und Selbstorganisationstheorien - Kontroversen und Gemeinsamkeiten. In B. Bergmann & P. Richter (Hrsg.), *Die Handlungsregulationstheorie. Von der Praxis einer Theorie*, 89-114. Göttingen: Hogrefe.
- Hacker, W. (1996). Entwickeln und Konstruieren als Denktätigkeit - zu einer Arbeitswissenschaft geistiger Erwerbstätigkeit. *Zeitschrift für Arbeitswissenschaft*, 50, 111-116.
- Hacker, W. (2002). Konstruktives Entwickeln. Psychologische Grundlagen. In W. Hacker (Hrsg.), *Denken in der Produktentwicklung. Psychologische Unterstützung der frühen Phasen*, 9-25. Zürich: Rainer Hampp.
- Hacker, W. (2005). Action Regulation Theory: A practical tool for the design of modern work processes? *European Journal of Work and Organizational Psychology*, 12, 105-150.
- Hacker, W. (2005). *Allgemeine Arbeitspsychologie. Psychische Regulation von Wissens-, Denk- und körperlicher Arbeit*. Bern: Huber.
- Hacker, W., Wetzstein, A. & Römer, A. (2002). Gibt es Vorgehensmerkmale erfolgreichen Entwerfens von Produkten? *Zeitschrift für Arbeitswissenschaft*, 56, 305-317.
- Hinsz, V. B. (1990). Cognitive and consensus processes in group recognition memory performance. *Journal of Personality and Social Psychology*, 59, 705-718.
- Hinsz, V. B., Tindale, R. S. & Vollrath, D. A. (1997). The emerging conceptualization of groups as information processors. *Psychological Bulletin*, 121, 43-64.
- Klimoski, R. & Mohammed, S. (1994). Team mental model: Construct or metaphor? *Journal of Management*, 20, 405-437.
- Larkin, J., McDermott, J., Simon, D. P. & Simon, H. A. (1980). Expert and novice performance in solving physics problems. *Science*, 208, 1355-1342.
- Larson, J. R. & Christensen, C. (1995). Groups as problem solving units: Toward a new meaning of social cognition. *British Journal of Social Psychology*, 32, 5-30.
- Levesque, L. L., Wilson, J. M. & Wholey, D. R. (2001). Cognitive divergence and shared mental models in software development project teams. *Journal of Organizational Behavior*, 22, 135-144.

- Lindemann, U. & Baumberger, C. (2004). Bewertung von Innovationsleistungen in der Unternehmung – Ein Konzept für das Benchmarking von Innovationsprozessen. *ZWF – Zeitschrift für wirtschaftlichen Fabrikbetrieb*, 99, 368-375.
- Miller, G. A., Galanter, E. & Pribram, K.-H. (1960). *Plans and the structure of behavior*. New York: Holt.
- Mohammed, S. & Dumville, B. C. (2001). Team mental models in a team knowledge framework: Expanding theory and measurement across disciplinary boundaries. *Journal of Organizational Behavior*, 22, 89-106.
- Müller, A. (2007). *Iterative Zielklärung und Handlungsplanung als Faktoren erfolgreichen Gruppenhandelns bei der Lösung komplexer Probleme. Eine handlungstheoretische Betrachtung des Konstruierens in Gruppen*. Berlin: Mensch & Buch.
- O'Leary-Kelly, A. M., Martocchio, J. J. & Frink, D. D. (1994). A review of the influence of group goals on group performance. *Academy of Management Journal*, 37, 1285-1501.
- Olson, G. M., Olson, J. S., Carter, M. R. & Storrøsten, M. (1992). Small Group Design Meetings: An Analysis of Collaboration. *Human-Computer Interaction*, 7, 547-574.
- Pahl, G., Beitz, W., Feldhusen, J. & Grote, K.-H. (2005). *Pahl/Beitz Konstruktionslehre. Grundlagen erfolgreicher Produktentwicklung. Methoden und Anwendung (6. Auflage)*. Berlin: Springer.
- Rouse, W. B. & Morris, N. M. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin*, 100, 549-563.
- Sachse, P. (2002). *Idea materialis: Entwurfsdenken und Darstellungshandeln: Über die allmähliche Verfertigung der Gedanken beim Skizzieren und Modellieren*. Berlin: Logos.
- Sachse, P. & Leinert, S. (2002). Skizzen und Modelle - Wieso Hilfsmittel des Denkens und Handelns beim Konstruieren? In W. Hacker (Hrsg.), *Denken in der Produktentwicklung. Psychologische Unterstützung der frühen Phasen*, 63-81. Zürich: Rainer Hampp.
- Steiner, I. D. (1972). *Group Processes and Productivity*. New York: Academic Press.
- Stempfle, J. & Badke-Schaub, P. (2002). Thinking in design teams - an analysis of team communication. *Design Studies*, 23, 473-496.
- Tschan, F. (2000). *Produktivität in Kleingruppen. Was machen produktive Gruppen anders und besser?* Bern: Huber.
- Ullmann, D. G., Dittrich, T. G. & Stauffer, L. A. (1988). A model of the mechanical design process based on empirical data. *AI EDAM*, 2, 53-52.
- Visser, W. (1994). Organisation of design activities: opportunistic, with hierarchical episodes. *Interacting with Computers*, 6, 239-274.
- Volpert, W. (1982). Das Modell der hierarchisch-sequenziellen Handlungsorganisation. In W. Hacker (Hrsg.), *Kognitive und motivationale Aspekte der Handlung*, 38-58. Bern: Huber.
- von Cranach, M., Ochsenbein, G. & Tschan, F. (1987). Actions of social systems: Theoretical and empirical investigations. In G. R. Semin & B. Krahe (Eds.), *Issues in contemporary German social psychology: History, theories and application*, (119-155): Sage Publications, Inc.
- von Cranach, M., Ochsenbein, G. & Valach, L. (1986). The group as a self-active system: Outline of a theory of group action. *European Journal of Social Psychology*, 16, 193-229.
- Wegge, J. (2004). *Führung von Arbeitsgruppen*. Göttingen: Hogrefe.
- West, M. A. (1996). Reflexivity and work group effectiveness: a conceptual integration. In M. A. West (Eds.), *Handbook of Work Group Psychology*, (555-579). Chichester: Wiley.

#### Korrespondenz-Adresse:

Dr. phil. Andreas Müller  
 Ludwig-Maximilians-Universität  
 Klinikum der Universität München  
 Institut und Poliklinik für Arbeits-,  
 Sozial- und Umweltmedizin  
 Ziemssenstraße 1  
 D-80336 München  
 andreas.mueller@med.uni-muenchen.de

# The influence of a problem solving training on shared mental models of spatial planners

Rinat Saifoulline\*, Rüdiger von der Weth\*,  
Walter Schönwandt\*\*, Christoph Hemberger\*\*, Jens-Peter Grunau\*\*

\* HTW Dresden / University of Applied Sciences Dresden / Human Resource Management and Ergonomics /  
Department of Business Administration

\*\* University of Stuttgart / Institute for the Foundations of Planning

## ABSTRACT

### Shared mental models of spatial planners

A methodology for spatial planners was developed to support very complex cooperative problem solving processes in this field. Learning and training this methodology should improve shared mental models of best practise and shared awareness of work processes. This should lead to more efficient cooperation and better results for planning teams. We compared trained teams of students with novices in an experimental study. The activities of the planning teams, interview protocols, and the work results were observed and analysed. The methodological training improved the degree of matching between mental models within the teams and the work results. New theoretical assumptions about the role of shared mental models for complex work processes have been developed.

## Keywords

Shared mental model – complex problem solving – spatial planning

## Acknowledgements

This study was supported by the DFG „German Research Foundation“ Number WE 2444/3-1.

## 1 Introduction

Problem solving processes in spatial planning are very often complex and deficient. Spatial planners must take into account a large number of heterogeneous aspects (social, ecological economical, political, and administrative). The results of planning cannot be seen immediately after the decisions and activities have been concluded. Often, positive and negative outcomes as well as side effects cannot be seen until much later. Hence, spatial planners get no immediate and only poor feedback on the consequences of their activities (von der Weth, 2002). Many stakeholders with different interests are involved in spatial planning projects. Therefore, the actors have to find compromises without firm knowledge about the outcome of possible actions. Spatial planning is not just a design process, but also a search for consensus between the decision makers involved and the people concerned.

Because all these people often have a very heterogeneous background, a lot of discrepancies can be caused by different planning approaches (Schönwandt & Voigt, 2005). These are paradigmatic assumptions about the background and the basic mechanisms of the problem situation. The planning approaches have a great influence on the course of discussions, decisions, and activities. A planning methodology was developed for coping with complex tasks in the field of spatial planning in a better way. The content of a course based on this methodology consisted of a curriculum of 20 units with topics like participation of stakeholders, forecasting methods, assessment of measures and projects. According to this methodology, at least six core activities have to be performed:

1. Planners must find problem definitions which can be communicated to most stakeholders in the planning process. These problem definitions

- should be as concrete as possible („Discussing problems“).
2. To enlarge the search space for solutions, the problems have to be „shifted“. The idea is to find new formulations, which allow more solutions to be found which are convenient and useful for most stakeholders („Problem shifting“).
  3. Because spatial planning is very complex, the causes of the problem should be analysed very carefully („Discussing causes“).
  4. This analysis should allow a stringent derivation of measures from these causes. The resulting measures should be designed to solve the problem effectively in the long run („Generating measures“).
  5. The definition of keywords should generate common understanding about the problem. The problem space, the goals, and the number of possible solutions change according to these definitions („Defining concepts“).
  6. Every planner uses a so called „planning approach“. Problems can be solved by different approaches. Users of the planning methodology should learn to explore and to switch between different approaches. This helps to criticise their own approach, to broaden their own horizon, and to find more possible solutions and compromises which are convenient for all groups involved (Schönwandt & Voigt, 2005) („Discussing planning approaches“).

These core activities of the planning methodology are called „key six“; for more information see Hemberger and colleagues (2008a). The spatial planning methodology described above does not prescribe a fixed sequence for these activities. Usually, they are conducted several times in an iterative loop.

This methodology improved the performance of spatial planning teams who took part in a 10 days training course (von der Weth et al., 2008). The next step was to find out in which way the cooperation in spatial planner teams was influenced by the training and, on the other hand, which psychological processes at team level influenced the process and the result.

According to empirical research in cognitive psychology, an important reason for a good team performance is the degree of matching of cognitive representations in teams. Klimoski and Mohammed (1994) describe this as „team mental models“, and Cannon-Bowers, Salas, and Converse (1995) discuss „shared mental models“. According to these theories, shared mental models allow a correct anticipation of future developments of a situation, and facilitate the team coordination (Cannon-Bowers, Salas, & Converse, 1995; Klimoski & Mohammed, 1994). The authors have dif-

ferent ideas about the content of these representations. According to Klimoski and Mohammed, the declarative part of team mental models consists of task aspects (e.g., goal definitions, success factors, important stimuli in the environment) and team aspects (e.g., role allocation, responsibilities, attitudes). The concept of „shared mental models“ is more differentiated (Cooke, Salas, Cannon-Bowers, & Stout, 2000). It is related to attributes of the tasks and distinguishes four knowledge domains which can be more or less concordant within a team. The so called „*equipment model*“ is about infrastructure and equipment (e.g., tools, machines, and their function). The possible actions and their outcomes are limited by these factors. Strategies, tactics, and heuristics for problem solving and knowledge of the demands are called the „*task model*“. The „*team interaction model*“ is related to communication channels and patterns, responsibilities, and role allocation, whereas the „*team model*“ is about abilities, skills, and other attributes of the team members. Most contents of the spatial planning methodology are related to the task model, which is also according to Cannon-Bowers and colleagues (1995) an important aspect for successful acting. First studies about the correlation between shared mental models and performance were done with professional teams under time pressure such as war plane crews (Rouse, Cannon-Bowers, & Salas, 1992). Later, the shared mental model approach was expanded to tasks in less dynamic environments like traditional organisations (e.g., Levine & Moreland, 1999). According to Badke-Schaub (2002), the development of shared mental models is an important step towards better cooperation and performance in mechanical design. These models should be flexible to enable design teams to implement flexible project management strategies.

The training course examined in this study should help develop a higher degree of matching of the task models within teams. According to different models about developing expertise (e.g., Anderson, 2005; Schön, 1985; Hacker, 1992), learning action strategies start with declarative knowledge about procedures, heuristics, and strategies. In professional contexts such knowledge is taught in schools, universities, or vocational training. With more and more experience, this knowledge will be adjusted to the requirements of the specific actor (e.g., tuning; c.f. Anderson, 2005). This is not a process of mere conditioning or model learning, but a process in which acting and reflection are interacting in a highly differentiated way (Schön, 1985). According to Hacker (1992), experts develop a more and more differentiated task model. Information in this model enables the expert to take into account all „w-questions“ in his/her activities (who? when? where? what? why?). A methodological teaching course

which gives (a) a theoretical foundation and description of the process, (b) the chance to develop shared experience in applying this methodology, and (c) opportunities for teams to reflect these activities should help to develop a common approach towards how to solve complex problems. If this works, the same team should improve its cooperation in other complex tasks as well. Trained teams should have more concordant task models as well as improved cooperation and better results.

## 2 Method and sample

A control group design was chosen for this study (Figure 1). The participants were students of the Faculty of Architecture and Urban Planning at the University of Stuttgart. They were rewarded with credit points. The best team in the experiment got the chance to participate in an excursion to a high-tech civil engineering project (tunnelling in St. Gotthard, Switzerland). First of all, the abilities and qualifications of the 58 participants of the study were observed (a questionnaire about work experience and study results, questionnaire on planning styles [Hacker, 1994]). There were no differences between the experimental group and the control group.

The experimental group was then given provided a training course in planning methodology. The control group had to work on example tasks of the course, but without learning the rules of the methodology. In this way the authors ensured that the effects of the training derived from the methodology rather than from learning by doing. Moreover a process of team building was possible in both groups. Measures of team cohesion were very high in both groups and without significant differences. After that, the participants had to work on a realistic spatial planning task: to make proposals of

how to solve a complex traffic problem.

The instructions for the first task started as follows: „The B27 motorway is one of the busiest roads in Stuttgart where there are often traffic jams. Yesterday, exponents of different interest groups had a meeting with the mayor of the city for a round table discussion. Political parties, business representatives, residents, environmentalists, and spatial planners took part. After long discussions about expansion of the road network and the resulting particulate matter emission, the discussion was adjourned ... “

The result should be a conceptual study of about 12 pages. The participants worked in teams of three students. They had to be ready in three days. Before and after the course the experimental variables were observed. The students of the control group took part in the training course several weeks later, and after that both groups again worked on a second complex spatial planning task concerning housing shortage in the city and region of Stuttgart (Germany). Before and after that second task the experimental variables were observed again.

## 3 Stimuli and Design

The *treatment* of this experiment was a ten days course in the planning methodology described above. The course consisted of teaching elements and training sequences. For more information see Hemberger and colleagues (2008a). The registration of *knowledge* was based on a questionnaire with „teaching back“ instruction (Strohschneider, 1990). The subjects had to describe the best practise strategy for solving this planning problem. They had to explain the problem solving strategy adequately. For the answers to these open questions a so-called „*matching index*“ was computed in the following way: the answers in the open question-

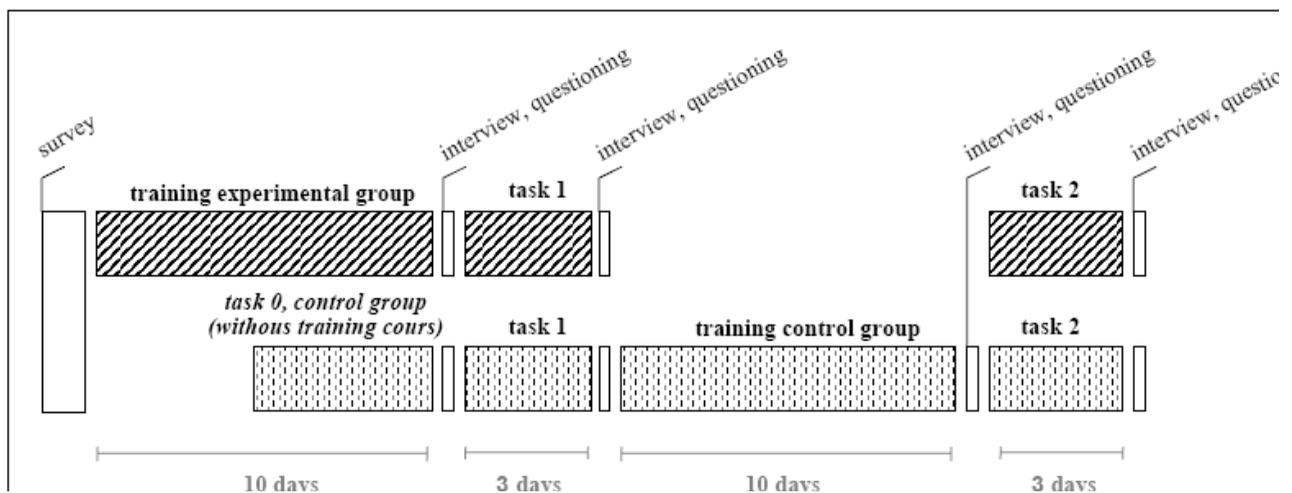


Figure 1: Experimental design

naire were compared to a list of 21 possible activities (Table 1).

Table 1: Methodological steps analysed in this study

1	<b>Discussing planning approaches</b>
2	Discussing problems and aims
2.1	Discussing initial state of planning
2.2	<b>Discussing problems</b>
2.3	Discussing advantages of problems
2.4	Discussing aims
2.5	<b>Problem shifting</b>
2.6	<b>Defining concepts (terms)</b>
2.7	Examining propositions
3	Compiling forecasts
4	<b>Discussing causes</b>
5	Discussing measures
5.1	<b>Generating measures</b>
5.2	Discussing constraints
5.3	Examining and assessing measures
5.4	Proposing measures for implementation
6	Discussing general frameworks
6.1	Discussing political arena and political agenda
6.2	Proposing participation of stakeholders
7	Gathering information
8	Organizing working process

Note: The most important steps (key six) are marked in bold.

19 of them were related to methodological steps and six of these were central methodological categories (key six). Additionally, we defined two more activities not defined in methodology, „gathering information“ and „organisation of work processes“. The rest was categorised as „behaviour not related to methodology“ and „behaviour not related to the task“. From these data we derived a methodological profile of each subject. Finally, we compared the degree of matching of this profile to the key six categories. The matching index increased by 1 point when two team members had a match for a methodological category and by 2 points when this category appeared in all profiles of the same team.

The quality of the course was assessed by a standardised questionnaire for teaching evaluation after the courses. It was about the satisfaction of the participants with the content, teachers' performance, and the organisational and technical prerequisites for their work.

To analyse the *planning activities*, the work process of all groups was continually observed by specially trained monitors through all three days. Every five minutes the predominant activity of this interval was recorded in a log sheet. The sum of 21 categories for activities which could be assigned to certain methodo-

logical steps was summarised as „methodological activity“. Six of these categories can be described as core activities of methodological planning; the sums were the key six activities. As mentioned above, these activities are unique and typical of the methodological approach, which underlies the training course evaluated in this study. The remaining four categories were summarised as „non-methodological categories“.

To compare the observed activities with the cognitive representation of their own activities, the subjects had to *describe their own problem solving process*: after finishing the work on the planning tasks, they had to answer an open question („How did you proceed in your work process?“) and fill in a questionnaire about the intensity and importance of all 21 methodological activities. The open description was analysed in the same way as the question concerning best practise, calculating a „completeness index“. A higher index indicates a more methodological reflection of the problem solving process. The additional questions should help to assess the general importance of the methodology for good results and the importance of specific steps of the methodological process.

After these open and closed questionnaires a discussion within the teams took place. A description of the planning process and an analysis of their own strengths and weaknesses in the work processes were asked for. The questions were based on an expert interview by Hacker (1992). The question of whether a group was able to discuss its own work objectively after methodological training was analysed. This is indicated by a lower level of blaming each other in the case of failure and generally a lower portion of group dynamic topics in the discussion, compared to the analysis of the problem solving process and contents related to spatial planning.

The *quality of the resulting conceptual studies* was assessed by an expert rating using 77 criteria. It was conducted independently by 3 experts from the Institute for the Foundations of Planning. These experts did not know whether the studies belonged to the experimental group teams or to the control group. There was a 5-step-ranking for each criterion; thus it was possible to make three rankings for the quality of the studies: one for the first task, one for the second task, and also one for both tasks together. Moreover, we analysed the number of solution proposals which were made in the respective conceptual study. A broader variability of solutions improves the quality of a conceptual study because it generates a greater selection of possible other solutions which could be useful in the case of conflicts or changes in the demands of the task.

*Control variables* consisted of a questionnaire concerning the individual planning style (Hacker, 1994) in the beginning, and several scales from a ques-

Table 2: Overview of research methods used (research plan)

	Start	After first training / before first task	During first task	After first task / before second training	After second training / before second task	During second task	After second task
Course evaluation		x			x		
Planning style	x						
Team work questionnaire				x			x
Knowledge	x	x		x	x		x
Behaviour recording			x			x	
Self reflection				x			x
Team reflection				x			x

tionnaire on team quality (TeamPuls by Wiedemann et al., 2001). Two processes were observed by video recordings of all activities. This allows the performance of detailed single case analyses.

The following table 2 shows when these methods were applied during the course of the study.

#### 4 Research questions and hypotheses

The effects of the training course were analysed for individuals and teams. On an individual level, strong effects of the teaching course could be shown (von der Weth et al., 2008): the usefulness of the course was rated highly. After the training course the individual cognitive representations were based more strictly on methodological concepts. This influenced the representation of the individual activities and the ideas about best practise. The effects of this course on individual behaviour and an improvement of results could be found. According to our assumptions, these findings should be connected with the development of shared mental task models.

**Higher degree of matching of mental models:** we began with the assumption that the new planning methodology includes important elements of a task model (heuristic rules, strategies, plans). Therefore, the teaching course should improve the shared men-

tal models of the teams. There should be a better consensus about the best way to perform in the planning tasks. Moreover, we believed that there is a higher degree of matching of the cognitive representation of the own problem solving process. Both variables were measured by a specific matching index. There should be a difference between the experimental and the control group after the first task and a significant improvement in the control group between the first and second tasks.

**Correlation between matching indices and behaviour measures (strategy):** then we tested the assumption that shared mental models based on the teaching course are correlated to a behaviour more organised according to methodological principles. We analysed the correlation between the two matching indices described above (best practise, individual strategy) on the one hand and the amount of time used for key six activities on the other hand. According to our hypotheses, such correlations between mental models and performed strategies do exist.

**Correlation between degree of matching and results:** finally the connection between the matching indices and the outcome of the planning process was tested. Correlations should exist between the matching indices (best practise, own strategy) and the ranking of the results.

5 Results

**Improved shared mental models:** we tested the effects of the training course by a multivariate analysis of variance (MANOVA). First we analysed the matching of the best practise models (Figure 2, left). It could be shown that there is a significant difference between the two tasks ( $F(1,0) = 11.21, p < .01, \text{partial } \eta^2 = .38$ ). There is also a significant interaction effect ( $F(1,0) = 23.40, p < .01, \text{partial } \eta^2 = .56$ ). Testing the effects on the representation of the teams' problem solving process again, an effect could be found. Although there was no significant difference between the two tasks ( $F(1,0) = 2.08, \text{n.s.}$ ), an interaction effect could be found ( $F(1,0) = 7.92, p < .01, \text{partial } \eta^2 = .50$ , Figure 2, right). The teaching course influenced the shared awareness of the individual activities significantly.

**Correlation between matching indices and behaviour measures (strategy):** it could be shown that there is a strong correlation between the matching indices and the general amount of methodologically oriented activities. The matching indices for best practise (Spearman rank correlation coefficient,  $r_s = .383, p < .01$ ) and for individual problem strategy ( $r_s = .419, p < .01$ ) both correlated to this variable. Analysing the shared mental models on a single item level, we found the following correlations between the degree of matching best practise models and key six activities (Table 3). According to the detailed results, this index was connected with the amount of time needed to discuss about measures and their causes.

**Correlation between matching indices and results:** it could be shown that there is a strong correlation between the matching indices and the rankings of the results. Both the matching indices for best practise

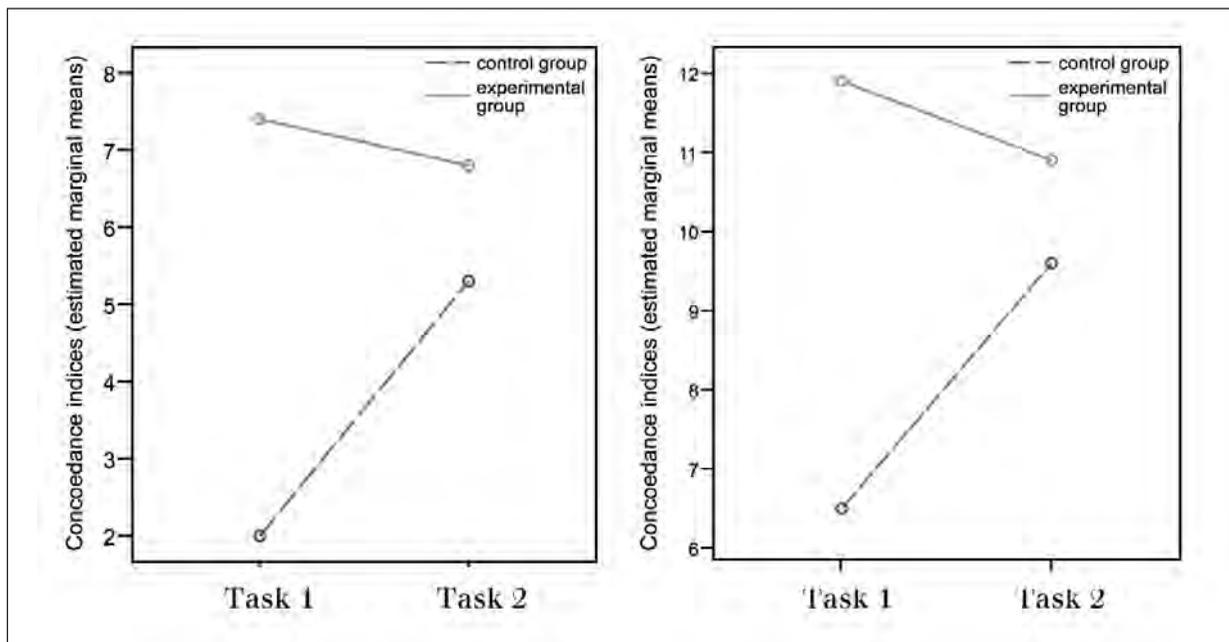


Figure 2: Matching indices of best practise (left) and of the representation of the individual activities (right).

Table 3: Correlations between single items of shared mental models about best practise and key six activities

Matching index for best practise aspects (Single Items)	Key six activities (Behaviour)
<b>Discussing planning approaches</b>	$r_s = .148, \text{n.s.}$
<b>Discussing problems</b>	$r_s = .292, \text{n.s.}$
<b>Problem shifting</b>	$r_s = .246, \text{n.s.}$
<b>Defining concepts (terms)</b>	$r_s = .170, \text{n.s.}$
<b>Discussing causes</b>	$r_s = .327, p < .05$
<b>Generating measures</b>	$r_s = .533, p < .01$

Table 4: Correlations between single items of shared mental models about best practise and quality of conceptual studies

Matching index for best practise aspects (Single items)	Quality of conceptual studies
<b>Discussing planning approaches</b>	$r_s = .249$ , n.s.
<b>Discussing problems</b>	$r_s = .513$ , $p < .01$
<b>Problem shifting</b>	$r_s = .013$ , n.s.
<b>Defining concepts (terms)</b>	$r_s = .169$ , n.s.
<b>Discussing causes</b>	$r_s = .516$ , $p < .01$
<b>Generating measures</b>	$r_s = .450$ , $p < .01$

( $r_s = .477$ ,  $p < .01$ ) and for individual problem strategy ( $r_s = .646$ ,  $p < .01$ ) correlated to this variable. Analysing shared mental models on a single item level, we found the following correlations between results and degree of matching about best practise (Table 4).

## 6 Discussion

The results show that the methodological course influenced the shared mental models of team members, the cooperative work process, and its results in a positive way. The methodological knowledge learned in this context was an important part of the task model of the participants. The training course changed the perceived competence for complex spatial planning tasks. The subjects found the methodological strategies useful. After training the methodology and working on a complex planning task, they developed a higher degree of matching of the mental models within the teams. The ideas about what makes a good problem solving strategy and the perception of their own work process became more similar. Apparently, this depended on certain behaviour patterns induced by methodological training. Subjects spending more time on discussion of the causal network of the problem and on the reasons for proposed measures developed shared mental models in a better way. The improved quality of results of teams with a higher degree of matching indicated that shared mental models, improved by methodological training, led to more efficient cooperation within the teams.

We assume that methodological training generates a basic structure to assess the quality of possible activities in the work process (shared mental models about best practise) and harmonises the perception of opportunities and threats (shared perception of the individual work process). The common methodological basic structure regulates the process of problem solv-

ing and makes it more effective and successful. Additionally, plausible methodological training and better cooperation in following work processes leads to more subjective control in that work process. According to empirical results (Dörner, 1996; Jansson & Smith, 1989), this lowers the probability of making serious mistakes in complex problem solving.

Whether the improvement in cooperation works in work constellations outside the university has to be analysed in further studies. There are additional aspects which have to be included in our research, e.g., in which way do conflicts and power struggles coming from diverging interests and a more interdisciplinary context influence the planning process and the results? A field study is planned to improve our knowledge about shared mental models in complex work processes in the domain of spatial planning.

## References

- Akademie für Raumforschung und Landesplanung (ARL) (Ed.) (2005). *Handwörterbuch der Raumplanung. [Concise dictionary of land use planning.]*. Hannover: Verlag der ARL (4., neu bearbeitete Auflage).
- Anderson, J. R. (2005). *Cognitive Psychology and Its Implications*. (6<sup>th</sup> edition). New York, NY: Worth Publishers.
- Cannon-Bowers, J. A., Salas, E., & Converse, S. (1995). Shared mental models in expert team decision making. In J. J. Castellan (Ed.), *Individual and group decision making: Current issues* (pp. 221-246). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Cooke, N. J., Salas, E., Cannon-Bowers, J. A., & Stout, R. (2000). Measuring team knowledge. *Human Factors*, 42, 151-175.

- De Jong, T. & Ferguson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist*, 31 (2), 105-115.
- Dörner, D. (1996). *The Logic of Failure*. New York: Metropolitan Books.
- Frey, D. & Schulz-Hardt, S. (2000). Entscheidungen und Fehlentscheidungen in der Politik: Die Gruppe als Problem. In H. Mandl & J. Gerstenmaier (Hrsg.), *Die Kluft zwischen Wissen und Handeln. Empirische und theoretische Lösungsansätze* (S. 75-95). Göttingen: Hogrefe.
- Hacker, W. (1992). *Expertenkönnen*. Göttingen: Verlag für Angewandte Psychologie.
- Hacker, W., Krassek, U., Heisig, B., & Ullrich, G. (1994). *Handlungsfragebogen. Manuskriptdruck, Institut für Allgemeine Psychologie und Methodenlehre*. Dresden: Technische Universität.
- Hemberger, C., Schönwandt, W. L., Grunau, J. Voermanek, K., von der Weth, R., & Saifoulline, R. (2008a): *Messbar bessere Ergebnisse beim Lösen komplexer Planungsprobleme – Empirische Evaluation einer Planungsmethodik* [Measurable better results on complex problem solving – An empiric evaluation of an planning methodology] (submitted for publication).
- Ickes, W. & Gonzalez, R. (1996). „Social“ cognition and social cognition. In J. L. Nyle & A. M. Brower (Eds.), *What's social about social cognition? Research on socially shared cognition in small groups* (pp. 285-308). Thousand Oaks: Sage.
- Janis, I. (1982). *Groupthink*. Boston: Houghton-Mifflin.
- Jansson, D. G. & Smith, S. M. (1989). Design fixation. *Proceedings on NSF Engineering Design Research Conference* (pp. 232-245).
- Klimoski, R. & Mohammed, S. (1994). Team mental model: construct or metaphor. *Journal of Management*, 20 (2), 403-437.
- Levine, J. M. & Moreland, R. L. (1999). Knowledge transmission in work groups: Helping newcomers to succeed. In L. L. Thompson, J. M. Levine, & D. M. Messick (Eds.), *Shared cognitions in organizations: The management of knowledge* (pp. 267-296). Mahwah, NJ: LEA.
- Mohammed, S. & Dumville, B. C. (2001). Team mental models in team knowledge framework: Expanding theory and measurement across disciplinary boundaries. *Journal of Organizational Behavior*, 22, 89-106.
- Rouse, W. B., Cannon-Bowers, J. A., & Salas, E. (1992). The role of mental models in team performance in complex systems. *IEEE Transactions On Systems, Man, And Cybernetics*, 22(6), 1296-1308.
- Schön, D. (1983). *Reflective practitioner: How professionals think in action*. New York: Basic Books.
- Schönwandt, W. L. (2008). *Planning in Crisis? Theoretical Orientations for Architecture and Planning*. Aldershot: Ashgate.
- Schönwandt, W. L. & Voigt, A. (2005). *Planungsansätze. Akademie für Raumforschung und Landesplanung (ARL)*, 769-776.
- Stout, R. J., Cannon-Bowers, J. A., Salas, E., & Milanovich, D. M. (1999). Planning, shared mental models, and coordinated performance: An empirical link is established. *Human Factors*, 41(1), 61-71.
- Strohschneider, S. (1990). *Wissenserwerb und Handlungsregulation*. Wiesbaden: Deutscher Universitätsverlag.
- Strohschneider, S. & von der Weth, R. (2002). „Ja, mach nur einen Plan...“. *Pannen und Fehlschläge. Ursachen, Beispiele, Lösungen. 2. vollständig überarbeitete, erweiterte und aktualisierte Auflage*. Bern: Huber.
- Thalemann, S. (2003). *Die Rolle geteilten Wissens beim netzbasierten kollaborativen Problemlösen. (The role of shared knowledge in net-based collaborative problem solving)* (Doctoral dissertation, Albert-Ludwigs-Universität Freiburg i. Brsg., 2003)
- von der Weth, R., Saifoulline, R., Schönwandt, W., Hemberger, C., & Grunau, J. (2008). *Evaluation of a complex problem solving methodology in spatial planning*. (submitted for publication)
- Wiedemann, J., Watzdorf, E. V., & Richter, P. (2001). *TeamPuls – Internetgestützte Teamdiagnose. (2. Auflage)* (Methodensammlung Band 15). Dresden: Technische Universität, Institut für Arbeits- und Organisationspsychologie.

Correspondence to:

Dipl.-Psych. Rinat Saifoulline

University of Applied Sciences Dresden

Human Resource Management and Ergonomics

Department of Business Administration

Friedrich List-Platz 1

D-01069 Dresden

r\_saifoulline@wiwi.htw-dresden.de

# Employment Prospects of Temporary and Permanent Workers: Associations with Well-being and Work Related Attitudes<sup>1</sup>

Thomas Rigotti\*, Nele De Cuyper\*\*, Hans De Witte\*\*, Sabine Korek\* and Gisela Mohr\*

\* University of Leipzig, Institute of Psychology II, Work and Organisational Psychology (Germany)

\*\* Katholieke Universiteit Leuven, Research Group Stress, Health and Well-being (Belgium)

## ABSTRACT

This study investigates the relation between employment prospects and well-being (irritation, anxiety, and depressive moods) and job and organisational attitudes (job satisfaction, affective organisational commitment) in temporary versus permanent workers. The perception of employment prospects are considered to be part of the employer-employee-relationship and express the employees' feeling of employment security and career progress within the company. We argue that for temporary workers reduced employment prospects are anticipated and predictable. They adapt their expectations regarding ongoing employment accordingly, whereas permanent workers may still expect lifelong employment. Results based on a German sample of 643 employees show that temporary workers experience more job insecurity than permanent workers, and they perceive fewer employment prospects. However, job insecurity is more strongly related to poor well-being among permanent workers than among temporary workers as expected, but the relationship of job insecurity with work attitudes was not different for the two employment groups. For both groups no difference was found concerning the relationship of psychological contracts on health outcomes, but the relationship to employees' attitudes was stronger among permanent workers than among temporary workers. It has to be taken into account, that precarious short term arrangements are underrepresented in the sample.

## Keywords

employment prospects – job insecurity – psychological contract – temporary employment – well-being – work attitudes

## 1 Introduction

Temporary employment has become one of the primary human resource instruments to promote organisational flexibility. For example, over 40% of new job offers in Germany are temporary in nature (Bellmann, Dahms, & Wahse, 2004). In March 2004, more than 2.4 million employees in Germany reported having a fixed term contract (Statistisches Bundesamt, 2004); i.e. the employee is hired directly by the employer for a short-term period (EU Directive 99/71/EC concerning the Framework Agreement on Fixed Term Work). This evolution towards increased temporary employ-

ment has contributed to a decline in employees' future employment prospects, especially in times of high unemployment rates.

The aim of this study is to investigate this development towards reduced employment prospects in relation to employees' well-being and attitudes. More specifically, we compare the reactions of those directly confronted with this development, as is the case for temporary workers, to those only observing the trend, as is the case for permanent workers. Future employment prospects in this study refer to job insecurity (i.e., the employees' concerns about the continuity of the job

1 This research is part of the Psycones-project (PSYchological CONtracts across Employment Situations) supported by a grant from the EU, 5th framework programme (HPSE-CT-2002-00121). Further information about the project is available on the web-page [www.uv.es/~psycon](http://www.uv.es/~psycon).

in the future; Klein Hesselink & Van Vuuren, 1999), or to employers' promises to provide future employment. These promises are part of the psychological contract, defined by Rousseau (1989, p. 121) as „an individual's perceptions about reciprocal promises between that individual and the organisation, and of what each party is entitled to receive as a function of those promises“. While job insecurity has been used in earlier studies in the realm of temporary work research, the specific focus upon psychological contracts is innovative to this study. We consider multiple outcomes: employees' attitudes, such as job satisfaction and affective organisational commitment, are used to measure employees' short-term reactions. Work-related well-being (depression, anxiety and irritation) is used to reflect long term reactions (Sverke, Hellgren & Näswall, 2002). This fairly large set of outcome variables presents yet another contribution to the literature.

### 1.1 Job Insecurity Among Temporaries and Permanents

Research consistently shows higher job insecurity among temporary workers (for a review, see De Cuyper et al., 2008). This high correlation has resulted in a near to interchangeable use of temporary employment and job insecurity, as illustrated in suggestions to use temporary employment as an objective indicator of job insecurity (e.g., Büssing, 1999; De Witte & Näswall, 2003; Pearce, 1998). However, a broad set of variables may inflate the relationship between job insecurity and temporary employment. For example, slightly more women than men work on temporary arrangements, temporary workers are on average younger than permanent workers, they are somewhat less educated, they are less likely to be a union member, and they work less hours per week on average (De Cuyper, Isaksson & De Witte, 2005; OECD, 2002). At the same time, research has shown gender (Kinnunen, Mauno, Nätti & Happonen, 1999; Näswall & De Witte, 2005; Sverke et al., 2004), age (Burchell et al., 1999; Kinnunen et al., 1999; Mohr, 2000; Näswall & De Witte, 2005; Roskies & Louis-Guerin, 1990; Van Vuuren, Klanderman, Jacobson, & Hartely, 1991), education (Näswall & De Witte, 2003; Schaufeli, 1992) and union membership (De Witte, 2005; Sverke et al., 2004) to be correlates of job insecurity. Family status appears also to affect feelings of job insecurity (e.g., Mauno, Kinnunen, Makingangas & Nätti, 2005). Even though these demographic variables show inconsistent relationships with job insecurity, it may nevertheless be important to control for them when studying job insecurity in the context of temporary employment. Our first hypothesis consequently reads as follows:

H1: After controlling for relevant demographic variables, employees on temporary contracts report higher job insecurity than employees on permanent contracts.

The harmful effects of job insecurity on job and organisational attitudes and well-being are well-documented (see e.g., Cheng & Chan, 2007; De Witte, 1999; Sverke et al., 2002). As respects job attitudes, a negative relationship between job insecurity and job satisfaction has been demonstrated repeatedly. Similarly, job insecurity relates negatively to organisational attitudes, such as affective organisational commitment (Benz, 2002; Borg, 1992; De Witte & Näswall, 2003; Lord & Hartley, 1998). In addition, job insecurity has been found to be positively related to anxiety and depression (Roskies & Louis-Guerin, 1990; Roskies, Louis-Guerin & Fournier, 1995; Orpen, 1995; Van Vuuren et al., 1991), burnout (Dekker & Schaufeli, 1995; Landsbergis, 1988) and irritation (De Cuyper & De Witte, 2005).

A tempting conclusion could be that job insecurity might be a severe stressor for temporary workers: temporary workers report higher job insecurity, and job insecurity associates with overall unfavourable outcomes. For example, Beard and Edwards (1995) predict worse effects of job insecurity among temporary workers: temporary workers are confronted with the threat of total job loss; the probability that this threat will be realized is close to maximal; and, because temporaries may experience difficulties in finding a permanent job, the threat of job loss is perceived to be important.

Alternatively, Jacobson and Hartley (1991, see also Pearce, 1998) argue that job insecurity might be less harmful for temporary workers than for permanent workers: job insecurity is part of the expectations and everyday experience of temporary workers. Unlike permanent workers, temporary workers may not perceive job insecurity as unpredictable or uncontrollable (Cavanough & Noe, 1999). Unpredictability and uncontrollability are however at the core of job insecurity theories. For example, Mohr (1997; 2000) argues that job insecurity can be broken down into four phases. The main difference between these four phases is the level of predictability. If predictability is low, individuals do not have enough information to decide on how or when they should act in order to prevent unemployment. Only in the final fourth phase, where dismissals have already been arranged, are the options for action evident.

Recently, authors have started to investigate the hypothesis on potential interactions between job insecurity and contract type. Evidence suggests that job insecurity is less problematic for temporary than for permanent workers. For example, the bulk of studies (De Cuyper & De Witte, 2006; 2007; De Witte & Näswall,

2005; Guest & Conway, 2000; Mauno et al., 2005) show that job insecurity associates with a reduction of job satisfaction and affective organisational commitment among permanent workers, but not among temporary workers. These findings have been replicated for other work-related outcomes, such as involvement, positive work-home interference, trust and turnover intention (De Cuyper & De Witte, 2005), as well as for general well-being and health (Bernhard-Oettel, Sverke & De Witte, 2005; Virtanen, Vahtera, Kivimäki, Pentti & Ferrie, 2002; Sverke, Gallagher & Hellgren, 2000). In line with this evidence, we hypothesize the following:

H2: The negative relationships between job insecurity and job satisfaction (H2a) and affective organisational commitment (H2b), and the positive relationships between job insecurity and irritation (H2c), anxiety (H2d) and depression (H2e) are stronger in permanent as compared to temporary workers.

## 1.2 Psychological Contracts of Temporaries and Permanents

An interesting question is whether such interaction effects hold for other aspects of employment prospects; for example, employees' perceptions on what the employer has promised to provide as part of the psychological contract that exists between employer and employee. In this respect, various authors have speculated about the emergence of a new psychological contract (Guest & Conway, 2000; Herriot & Pemberton, 1995; Robinson, Kraatz & Rousseau, 1994). The new psychological contract aligns with current labour market evolutions: it highlights the importance of organisational flexibility to cope with increased worldwide competition, and it stresses employability to guarantee continuous employment for employees. This new psychological contract has often created, intentionally or not, a more transactional relationship for employees (Guest & Clinton, 2005), in which economic exchanges and short-term benefits prevail. In contrast, the old psychological contract aims at establishing a long-term employment relationship by broadening the range of promises to include socio-emotional aspects, most notably job security. The presence or absence of promises concerning employment prospects might be the most critical indicator for differentiating between the old and the new psychological contract. Leading authors have suggested that the new psychological contract will become increasingly important for all employees (Anderson & Schalk, 1998; Cooper, 1999; Coyle-Shapiro & Kessler, 2002; Hiltrop, 1995; Koh & Yer, 2000; Millward & Brewerton, 2000; Rousseau, 1995). Today, the new psychological contract might already be relevant for temporary as compared

to permanent workers (Millward & Brewerton, 2000; Millward & Hopkins, 1998): This suggests that temporary compared with permanent workers may perceive fewer promises regarding reasonable job security or future career prospects in the organisation, which was demonstrated in the study by Claes et al. (2002). This study did not, however, control for demographics, and this may have influenced the results. Accordingly, our third hypothesis reads as follows:

H3: Controlling for demographics, temporaries perceive fewer employers' promises regarding future employment prospects than permanents.

Previous studies (e.g., McDonald & Makin, 2000; Millward & Hopkins, 1998; Van Dyne & Ang, 1998) have hypothesized that the effects of this new psychological contract are likely to be negative for employees. For example, Beard and Edwards (1995) suggest that it may prevent the development of a trust relationship with the employer, and this, in turn, would result in detrimental outcomes. We challenge this assumption of overall negative effects: We suggest differential effects for temporary and permanent workers based on two reasons: We argue that perceived psychological contract promises should be evaluated in light of employees' general expectations on what constitutes contemporary employment. If employees expect ongoing employment, as in the case of permanent employment, poor well-being and unfavourable attitudes may result when these expectations are not matched with employer's promises on employment prospects. In contrast, no such harmful effects are expected in cases where both expectations and promises are absent, as might be the case for temporary workers. Hence, we hypothesize the following:

H4: The positive relationships between psychological contract expectations regarding future employment prospects and job satisfaction (H4a) and affective organisational commitment (H4b), and the negative relationships between such psychological contract expectations and irritation (H4c), anxiety (H4d) and depression (H4e) are strong in permanent compared with temporary workers.

## 2 Method

### 2.1 Procedure

Data were gathered in 34 organisations across Germany in 2004. Fourteen organisations ( $N = 226$ ) were selected from the educational sector (primarily organisations providing vocational education), 11 organisations ( $N = 202$ ) were selected from the retail sector (shops, banks) and 9 organisations ( $N = 215$ ) were selected from the food sector (production plants). All

organisations willing to participate were included. Response rates within organisations varied between 13% and 100%, but were, on average, above 60%. A special effort was made to sample temporary workers. Most questionnaires were distributed by HR-managers. Participation was voluntary and anonymity was guaranteed.

## 2.2 Respondents

Overall, we gathered data from 643 employees, 45% of which ( $N = 290$ ) had a temporary employment contract. Temporary workers were evenly represented in all three sectors, with 42% in the food sector, 41% in the retail sector, and 51% in the educational sector ( $\chi^2(2, N = 643) = 5.52, p > .05$ ). All temporary workers were directly hired by the organisation they worked for (excluding temporary agency workers). The contract duration of their current contract was on average  $M = 17.90$  months ( $SD = 24.25$ ) and they had been employed with their current employer for  $M = 27.87$  months ( $SD = 40.27$ ) on average. Thus for many it was not the first temporary assignment with the same employer.

Overall, the sample consisted of 51% females. The average age was 37 years ( $SD = 12$  years), with a range from 16 to 72 years. The temporary and the permanent sample differed in line with population trends. Temporary workers ( $M = 32.77, SD = 12.55$ ) were on average eight years younger than permanent workers ( $M = 40.78, 10.16$ ),  $t(629) = -8.93, p < .001$ . 38% of the temporary employees, and only 16% of the permanent employees were employed as blue-collar workers,  $\chi^2(1, N = 643) = 40.48, p < .001$ . The temporary sample had a slightly lower educational background,  $t(619) = -2.68, p < .01$ , mainly because permanent workers in our sample were more likely to have a university degree (48% as compared to 38%). No differences regarding the distribution of gender across the two subsamples of temporary and permanent workers could be found,  $\chi^2(1, N = 638) = 1.82, p = .18$ . Permanent workers (27%) were twice as likely to belong to a union than temporary workers (13%),  $\chi^2(1, N = 640) = 20.43, p < .01$ . Furthermore, permanent workers ( $M = 37.21, SD = 8.48$ ) worked about 5 hours more a week than temporary workers ( $M = 32.23, SD = 11.85$ ),  $t(630) = -6.14, p < .001$ . Permanent workers (79%) were more likely to live with a partner than temporary workers (65%),  $\chi^2(1, N = 629) = 15.14, p < .01$ .

## 2.3 Measures

### *Type of contract*

Respondents were asked to state whether they hold a permanent (1) or a temporary (0) employment contract.

### *Control Variables*

Age (years) and weekly working hours (average hours per week) were continuous variables. Gender (0 = female; 1 = male), occupational position (1 = blue collar worker; 0 = others), family status (0 = single; 1 = living with partner), and union membership (0 = no; 1 = yes) were dichotomous variables. In line with the annual Microsensus analyses by the German federal statistics office (Statistisches Bundesamt Deutschland, 2005), educational level was assessed by using a combination of school and vocational education, which corresponds with the International Standard Classification of Education (ISCED, OECD, 1999). This measure is generally considered to be metric. Finally, type of sector was dummy-coded, with educational sector being the reference category.

### *Job insecurity*

A four-item measure, developed by De Witte (2000), was used to assess perceived job insecurity (e.g., „I feel insecure about the future of my job“). Responses ranged from 1 (strongly disagree) to 5 (strongly agree). Confirmatory factor analysis ( $\chi^2 = 1.07, df = 2, p > .05$   $GFI = .99, AGFI = .99, RMSEA = .00$ ) supported one dimensionality. The scale yielded an alpha of .74.

### *Psychological contract*

To measure perceived employer's promises regarding future employment prospects as part of the psychological contract, we used a set of four items: „Has your organisation promised or committed itself: (1) to provide you with a reasonably secure job?, (2) to provide you with a career?, (3) to provide you with opportunities to advance and grow?, and (4) to improve your future employment prospects?“ Respondents could indicate whether they perceived the promise to be given (1) or not (0), resulting in a scale ranging from 0 to 4. A confirmatory factor analysis supported a one-dimensional construct ( $\chi^2 = 1.94, df = 2, p > .05$   $GFI = .99, AGFI = .99, RMSEA = .00$ ). Cronbach's Alpha for the four items was .72.

### *Dependent variables*

*Job Satisfaction* was measured using four items developed by Price (1997; e.g., „I find enjoyment in my job“). Respondents could answer on a five-point response format (1 = strongly disagree; 5 = strongly agree). One-dimensionality was supported by a confirmatory factor

Table 1: Pearson correlations, means and standard deviations, and reliability for all variables

	1		2		3		4		5		6		7		Permanents (n=353)		Temporaries (n=290)		Overall sample (n=643)		
	M	SD	M	SD	M	SD	M	SD													
1 Job insecurity																					
2 Job Satisfaction																					
3 Commitment																					
4 Cognitive Irritation																					
5 Depressive mood																					
6 anxiety																					
7 Psychological Contract																					

\*\*p<.01, \*p<.05., Cronbach Alphas in Parentheses

analysis ( $\chi^2 = 1.18$ ,  $df = 2$ ,  $p > .05$   $GFI = .99$ ,  $AGFI = .99$ ,  $RMSEA = .00$ ). The scale yielded an Alpha of .75.

*Affective Organisational Commitment* was measured with four items developed by Cook and Wall (1980; e.g., „To know that my own work has made a contribution to the good of the organisation would please me“), with responses varying from 1 (strongly disagree) to 5 (strongly agree). The items reflected one latent construct ( $\chi^2 = 7.93$ ,  $df = 5$ ,  $p > .05$   $GFI = .99$ ,  $AGFI = .98$ ,  $RMSEA = .05$ ). Cronbach's Alpha of the scale in our sample was .68.

*Cognitive Irritation (Rumination)* was measured using a three-item subset of the Irritation scale (Mohr, Rigotti & Müller, 2005). Irritation is defined as an early state of psychological impairment caused by perceived goal discrepancy, and includes rumination about problems at work (cognitive irritation) and emotional irritation (Müller, Mohr & Rigotti, 2004, p. 223). Cognitive Irritation was assessed by items such as the following: „I often think of my problems at work even at home“ (1 = strongly disagree; 7 = strongly agree). As CFA is not applicable to three items, we performed a principal component analysis that clearly indicated the presence of a single factor, explaining 79% of the variance. The scale yielded a Cronbach's Alpha of .87.

Finally, anxiety-versus-contentment (anxiety) and depression-versus-enthusiasm (depressive mood) were measured with six items each, using the scale developed by Warr (1990). Respondents had to indicate how often (1 = rarely or never; 5 = very often or always) they felt each of the 12 moods (e.g., tense, contented) with respect to their work during the past few weeks. A higher score reflected a more negative affect. The anxiety scale yielded an alpha of .80, and the depression scale an alpha of .84. Confirmatory factor analyses yielded satisfactory results for both anxiety ( $\chi^2 = 75.68$ ,  $df = 9$ ,  $p < .00$ ;  $GFI = .96$ ,  $AGFI = .91$ ,  $RMSEA = .11$ ), and depression ( $\chi^2 = 86.88$ ,  $df = 9$ ,  $p < .00$ ;  $GFI = .94$ ,  $AGFI = .85$ ,  $RMSEA = .12$ ). Even though the RMSEA points towards a redundancy in the item-pool, the alpha could not be altered by deletion of items. We therefore decided to include all items.

Table 1 gives an overview of correlations between variables, as well as means and standard deviations of the instruments for the sample of permanent and temporary workers separately.

## 2.4 Analyses

Differences between permanent and temporary employees on job insecurity (H1) and psychological contract

content (H5) were tested using multiple regression analyses with the control variables in the first step (age, gender, weekly working hours, occupational position, educational level, family situation, union membership, and sectors) and type of contract in the second step.

Hypotheses 2 and 4 were tested using multiple regression analyses separately for all dependent variables, and applying list wise deletion. In the first step, we entered the set of control variables, and type of contract. In the second step, we added job insecurity, and psychological contract content (cf. Aiken & West, 1991; Baron & Kenny, 1986). In a last step, we included the interaction term (job insecurity \* type of contract for hypothesis 2, and psychological contract \* type of contract for hypothesis 4). In line with Aguinis (2004), interactions are to be taken seriously when explaining 1% of the variance. Because of the categorical nature of the moderator variable, we additionally calculated the odds ratio of error variance: these should be below 1.50 to support the assumption of homogeneity of error variances (Aguinis, Peterson & Pierce, 1999) across categories of the moderator variable. When interaction terms proved significant, we plotted the regressions below and above one standard deviation of the mean (Aiken & West, 1991), and performed simple slope regression analyses separately for the samples of temporary and permanent workers (including socio-demographics as controls).

## 3 Results

### 3.1 Job Insecurity and psychological contracts on employment prospects in temporary and permanent workers

Temporary workers experienced higher job insecurity than permanent workers ( $\beta = -.30$ ,  $p < .001$ ), even when controlling for background variables. This was in line with hypothesis 1. Control variables predicted 13% of the variance in job insecurity. Type of employment contract explained an additional 6% of the variance in job insecurity.

In the following, we only present information on significant predictors. Blue collar workers reported higher levels of job insecurity ( $\beta = .21$ ,  $p < .001$ ); working hours were positively related to job insecurity ( $\beta = .13$ ,  $p < .001$ ); and employees from different sectors differed in their perception of job insecurity. Compared to the educational sector, employees from the food industry and from retail and service gave lower ratings on job insecurity ( $\beta = -.27$ ,  $p < .001$ , and  $\beta = -.10$ ,  $p < .05$ , respectively).

Permanent workers perceived more promises on future employment prospects than temporary workers,

Table 2: Results of multiple regression analyses

Step	Job satisfaction					Commitment					Cognitive Irritation					Depressive Mood					Anxiety						
	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
1. Age	.11*	.11*	.10*	.14**	.15**	.14**	.06	.07	.06	.01	.01	.06	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01
2. Male	-.11**	-.10*	-.10*	.00	-.01	-.01	-.07	-.09*	-.08	-.05	-.05	-.07	-.09*	-.08	-.05	-.05	-.05	-.05	-.05	-.05	-.05	-.05	-.05	-.05	-.05	-.05	-.05
3. weekly working hours	.02	.05	.05	.02	.03	.03	-.01	-.03	-.01	.07	.03	-.01	-.03	.07	.03	.04	.04	.04	.17**	.13**	.15**	.13**	.15**	.13**	.15**	.13**	.15**
4. Blue collar	-.20**	-.15**	-.14**	-.27**	-.21**	-.21**	-.12*	-.14**	-.11*	.08	-.01	-.01	-.01	-.04	-.05	-.05	-.05	-.02	-.02	-.02	-.02	-.02	-.02	-.02	-.02	-.02	-.02
5. Educational Level	.04	.04	.02	.08	.09	.07	.01	.01	-.01	-.04	-.04	-.01	-.01	-.04	-.05	-.05	-.05	-.02	-.02	-.02	-.02	-.02	-.02	-.02	-.02	-.02	-.02
6. Living with partner	.03	.03	.02	-.03	-.03	-.03	.06	.06	.06	.00	.00	.06	.06	.00	.00	.00	.00	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01
7. union member	-.03	-.03	-.01	.03	.02	.03	-.02	-.02	-.03	.02	.03	-.02	-.02	.02	.02	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01	.01
8. Food industry	.11	.04	.00	.28**	.25**	.25**	-.18**	-.14	-.18*	-.12*	-.12*	-.14	-.18*	-.12*	-.04	-.05	-.05	-.23**	-.14*	-.16**	-.16**	-.16**	-.16**	-.16**	-.16**	-.16**	-.16**
9. Retail	.07	.04	.05	.17**	.15**	.14**	-.04	-.02	-.05	.05	.05	-.02	-.05	.05	.08	.07	.07	-.05	-.02	-.05	-.05	-.05	-.05	-.05	-.05	-.05	-.05
10. Permanent contract	-.07	-.15**	-.56**	-.06	-.15**	-.27*	.09	.11*	-.41**	.10*	.10*	.11*	-.41**	.10*	.25**	-.01	-.01	.06	.16**	-.11	-.11	-.11	-.11	-.11	-.11	-.11	-.11
11. Job insecurity		-.27**	-.28**		-.22**	-.21**		.12**	-.04			.12**	-.04		.56**	.25**		.51**	.20								
12. Psychological contract	.01	.01	-.15*	.12**	.12**	-.01	.06	.06	-.03	-.03	-.03	.06	-.03	-.03	-.06	.00	.00	-.01	-.01	.05	.05	.05	.05	.05	.05	.05	.05
13. Job insecurity x Type of contract			.06		.00	.00		.44**				.44**			.29*	.29*		.50**									
14. Psychological contract x Type of contract			.28**		.22**	.22**		.17*				.17*			-.10	-.10		-.06									
$\Delta R^2$	.07**	.06**	.02**	.10**	.06**	.01*	.10**	.01*	.03**	.05*	.11**	.01*	.03**	.05*	.11**	.01*	.06**	.08**	.01*	.01*	.01*	.01*	.01*	.01*	.01*	.01*	.01*
Heterogeneous error variance?	1:1.26			1:1.22			1:1.04			1:1.01			1:1.07														

\*p &lt; .05, \*\*p &lt; .01. Figures shown in the table are standardised beta-weights.

even when controlling for background variables ( $\beta = .22, p < .001$ ). This aligned with hypothesis 3. Control variables accounted for 9% of the variance. Contract type added 3% of the variance in the dependent variable. Regarding the control variables, men reported more promises than women ( $\beta = .12, p < .01$ ), and working hours were positively related to promises ( $\beta = .12, p < .01$ ). Employees from the food industry reported fewer promises than those from the educational sector ( $\beta = -.24, p < .001$ ).

### 3.2 Interactions between Job Insecurity, Employment prospects and Contract Type

The results of multiple regression analyses are presented in table 2. Control variables (including type of contract) accounted for between 3% (depressive mood) and 10% (affective organisational commitment, cognitive irritation) of the variance in the dependent variables.

Differences in dependent variables between temporary and permanent workers revealed to be non-significant (except for depressive mood – with higher values for permanent workers) until including job insecurity, and employment prospects within the psychological contract.

Main effects of job insecurity (Step 2, under control of sociodemographics and type of contract) could be found for all dependent variables, while for the employment prospects within the psychological contract only a significant relationship to affective commitment could be found.

The interaction term of job insecurity with type of employment contract added significantly in predicting variance in cognitive irritation, depression, and anxiety, as hypothesized (Table 2). No such interactions were found for job satisfaction and organisational commitment. In contrast the interaction term of employment prospects (psychological contract) significantly explained variance in job satisfaction, and affective commitment, as well as cognitive irritation (but with a lower effect size as compared to job insecurity), but not of depressive mood, and anxiety. The assumption of homogeneous error variance is met for all variables. Figures 1 to 5 show the regression-lines for the five significant interactions. As the interaction between psychological contract and type of employment contract had a very low effect size, and only marginally surpassed the significance level of 5% ( $\beta = .17, p = .046$ ), we did not plot this interaction. As predicted the association between job insecurity and all indicators of well-being was stronger for permanent than for temporary workers. H2 was supported partly, that is only for the health outcomes, not concerning the work

attitudes. Also H4 was partly supported, but with an opposite pattern.

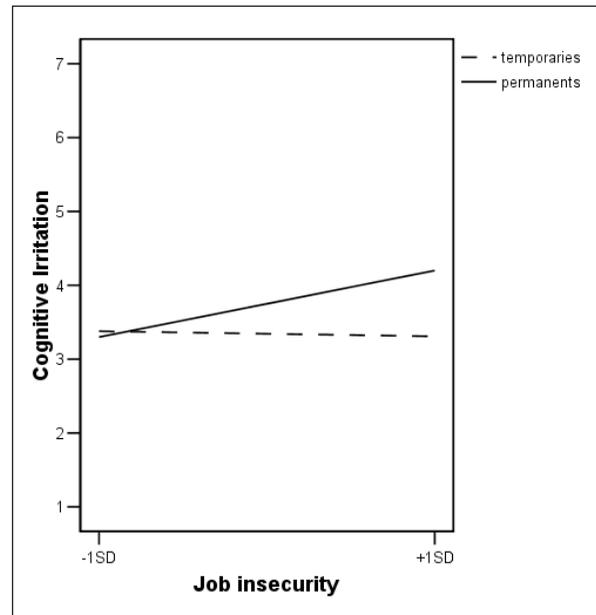


Figure 1: Significant interaction effect between perceived job insecurity and type of employment contract on cognitive irritation

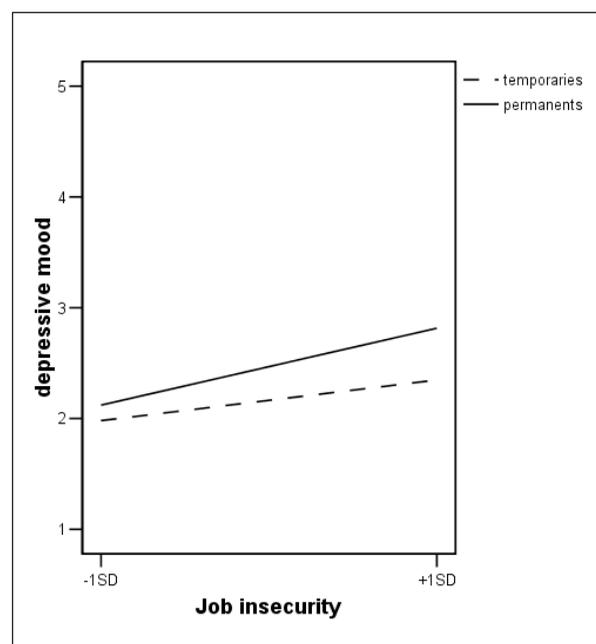


Figure 2: Significant interaction effect between perceived job insecurity and type of employment contract on depressive mood

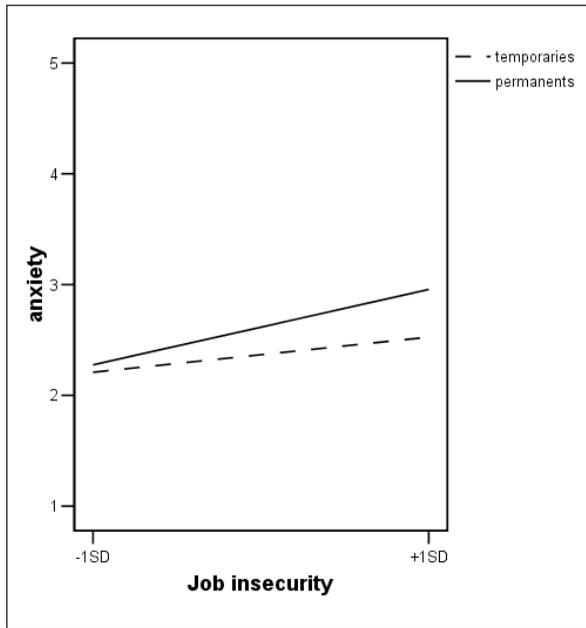


Figure 3:  
Significant interaction effect between perceived job insecurity and type of employment contract on anxiety

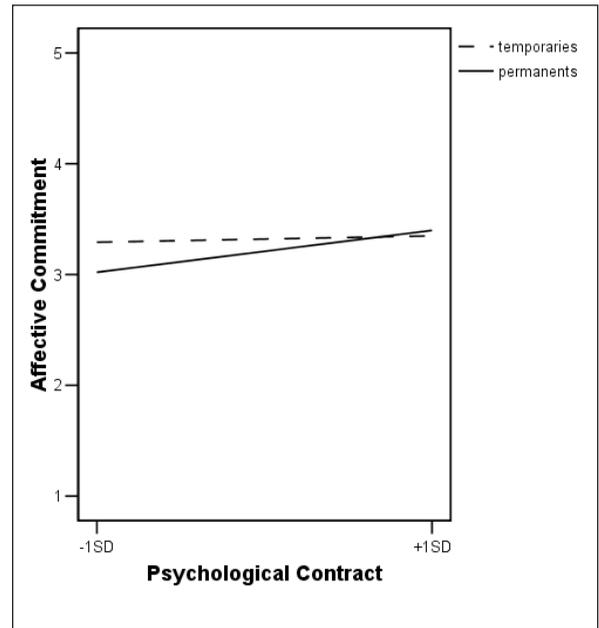


Figure 5:  
Significant interaction effect between psychological contract and type of employment contract on commitment

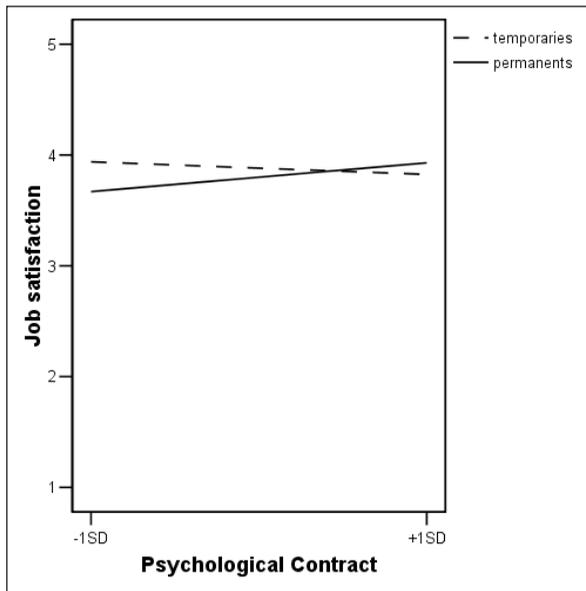


Figure 4:  
Significant interaction effect between psychological contract and type of employment contract on job satisfaction

When testing the samples of temporary and permanent workers separately, the following results were found with respect to job insecurity: For permanent workers, job insecurity could explain variance in all dependent variables above the impact of controls (job satisfaction:  $\beta = -.24, p < .001, \Delta R^2 = .06$ ; commitment:  $\beta = -.23, p < .001, \Delta R^2 = .05$ ; cognitive irritation:  $\beta = .22, p < .001, \Delta R^2 = .05$ ; depressive mood:  $\beta = .41, p < .001, \Delta R^2 = .16$ ; anxiety:  $\beta = .36, p < .001, \Delta R^2 = .15$ ). But this was also primarily the case for temporary workers, where job insecurity significantly added to the explanation of variance in the dependent variables, with the exception of cognitive irritation (job satisfaction:  $\beta = -.21, p < .001, \Delta R^2 = .05$ ; commitment:  $\beta = -.17, p < .01, \Delta R^2 = .02$ ; cognitive irritation:  $\beta = .00, p > .05, \Delta R^2 = .00$ ; depressive mood:  $\beta = .26, p < .001, \Delta R^2 = .05$ ; anxiety:  $\beta = .22, p < .001, \Delta R^2 = .04$ ).

The psychological contract could significantly explain variance in dependent variables for permanent employees, except in the case of cognitive irritation (job satisfaction:  $\beta = .18, p < .001, \Delta R^2 = .05$ ; commitment:  $\beta = .26, p < .001, \Delta R^2 = .06$ ; cognitive irritation:  $\beta = .06, p > .05, \Delta R^2 = .00$ ; depressive mood:  $\beta = -.18, p < .01, \Delta R^2 = .05$ ; anxiety:  $\beta = -.12, p < .05, \Delta R^2 = .01$ ). No significant relationship between psychological contract and outcomes could be found for temporary employees.

## 4 Discussion

This study aimed to address the overall growing concern about reduced employment prospects in the current labour market and the impact that such a reduction might have on employees' attitudes (job satisfaction, organisational commitment) and well-being (depressive mood, anxiety, irritation). We investigated this development from the perspective of temporary and permanent workers, for they may differ in the extent to which they experience this change. Indeed, as expected temporary workers in our study were more insecure about their jobs (H1) and they perceived fewer employers' promises with respect to securing long-term employment (H3).

Nevertheless we assumed, that these two circumstances – though being evidently more prone for temporary workers – have stronger negative effects on well-being and positive work attitudes like job satisfaction or affective organisational commitment for permanent workers than temporary workers (H2, H4). We argued that temporary workers may anticipate reduced employment prospects and adapt their expectations accordingly, i.e. in line with their everyday experience. For them the options on how to deal with reduced employment prospects are more clear. In contrast, the discrepancy between expectations and reality might be more distressing for permanent workers (also perceiving unfavourable employment prospects and job insecurity though to a lower degree), because they may feel betrayed when confronted with reduced employment prospects. Additionally the options what to do are less clear than for a temporary worker who knows anyhow that he or she will soon have to search for a new job. As long as no term of notice is given it remains unclear for permanents when and if he or she should take action and leave the company. Because of this aspect of less predictability we assumed, that job insecurity and perception of reduced employment prospects, though less distinct in permanent workers, may show a stronger relation to health outcomes and work attitudes than in the case of temporary workers.

Our results partly supported this argument: Job insecurity was more strongly related to poor well-being (irritation, depression and anxiety) among permanent workers than among temporary workers, but this was not the case for the two work attitudes. Similarly, a limited number of employers' promises concerning long-term employment related more strongly to unfavourable work attitudes (job satisfaction and organisational commitment) among permanent than among temporary workers, but this was not the case for the three health outcomes. The question then is why the interaction term between contract type and either job insecurity or psychological contract promises were dif-

ferently related to the outcomes; i.e., to well-being and organisational outcomes, respectively.

A possible explanation could be that employment prospects is organisation-specific and tied up with the current job. Hence, employees' reactions may primarily reflect discontent with the current job (e.g., job satisfaction) and with the organisation (e.g., organisational commitment). It does not necessarily imply a threat, whilst job insecurity may additionally represent a fairly general threat to future employment prospects. A threat may be linked more to emotional responses like depressive mood and anxiety. Employees' reactions in terms of reduced well-being may mirror this generalized perception. By being specific about effects on health outcomes we do not claim, that feelings of job insecurity are unrelated to one's current job and organisation. Job insecurity had a main effect on all outcome variables in our study, pointing to its overall harmful effect.

Similarly, for job insecurity and psychological employment prospects different control variables proved to be relevant. For example, employment in the sector of education as compared to employment in the food industry was associated with higher job insecurity, but also with more perceived employers' promises regarding future employment. This puzzling finding may, at least partly, be explained by the labour market situation: in the educational sector, we mainly sampled private organisations offering vocational training, and public grants for individuals following such trainings were cut shortly before the study. As a result, the educational sector was going through a severe crisis when data were gathered. This may have caused higher job insecurity levels, although employees' perceptions of promises may not yet have changed. This may provide additional evidence for different processes underlying the perception of promises, and perhaps also for the impact of job insecurity and the development of psychological contract content.

### 4.1 Limitations

This study shares the limitations of many studies: A cross-sectional design and reliance on self-reports. Considering the short-term nature of temporary employment contracts and hence, the high level of organisational turnover among temporary workers, it is however difficult especially for that topic to establish a longitudinal design. The reliance on self reports may lead to inflated relationships (cf. Percept-percept-inflation paradigm; Crampton & Wagner, 1994). This threat of common variance effects may be considerably reduced by the objective nature of the type of contract. Also, common method effects are likely to attenuate

rather than to strengthen interaction effects (Conway & Briner, 2002). As these interaction effects cover the core question of this study, we feel confident that this methodological limitation does not substantially limit the meaning and validity of our findings. Also it should be considered that employment prospects are defined as part of the psychological contract, which is – in contrast to the formal contract – subjective per definition.

Furthermore, the sample used in this study was not representative for the total working population, if only because of the overrepresentation of temporary workers. The subsample of temporary workers included a large majority of fixed term contract workers. Even though fixed term contracts are the most common temporary contract type both in Germany and in Europe (De Cuyper et al., 2005), it may hamper possibilities for generalizing findings to all types of temporary workers, for example temporary agency workers (c.f., Silla, Garcia & Péiro, 2005) or more precarious arrangements like on call workers and those with very short contracts. Similarly, due to the specific German labour market situation, and most notably its high unemployment rate, results may be different in other countries. Nevertheless, the varied organisational contexts represent a major advantage of this study. And, as mentioned earlier, our results do at least partially replicate findings of various earlier European studies: Bernhard-Oettel et al. (2005) with a Swedish sample of  $N = 954$  nurses and De Cuyper and De Witte (2007) with a sample of  $N = 477$  of temporary and permanent workers from different occupations in Belgium also found an interaction of type of employment contract and perception of job insecurity in line with our results (see also De Cuyper & De Witte, 2006). It has to be mentioned, that there are great differences concerning facts related to temporary employment in these countries. To name only a few: In Sweden participation in unions is much larger. In Belgium Human Resource Managers anticipate workforce expansion whereas in Germany they anticipate workforce reduction, based on a survey in 2004 (Claes, 2008).

#### 4.2 Conclusion

The results of our study should not be misinterpreted as a plea for the use of temporary employment contracts. Many of desired employees' attitudes and behaviors can only be fostered through a long-term relationship, which include employment prospects and mutual trust. The reported results rather point towards negative side effects of increased numerical flexibility even for the core employees.

The study added to the line of research about job insecurity another moderating variable: perceived em-

ployers' promises concerning ongoing employment. These employment prospects are more predictive of the attitudes of permanent workers than of those from the temporary workers. Additionally, our study showed that job insecurity and employment prospects are complementary in their interaction with contract type. This highlights the importance of using various measurements of employee prospects to understand processes underlying the responses of employees; a finding that is innovative to this study.

We argued for an explanation of our findings in terms of employees' overall expectations of what constitutes work in today's labour market. More specifically, we proposed that, while permanent workers may expect ongoing employment based primarily on the open-ended nature of their employment contract, temporary workers may anticipate reduced employment prospects and align their expectations with this everyday experience. This process may neutralize the negative effects of limited employment prospects among temporaries and may even represent a functional coping mechanism. The long-term effects of such a decrease in expectations have yet to be evaluated. This may be one of the most acute challenges for further research in this field.

#### References

- Aguinis, H. (2004). *Regression Analysis for Categorical Moderators*. New York: Guilford Press.
- Aguinis, H., Petersen, S. A., & Pierce, C. A. (1999). Appraisal of the homogeneity of error variance assumption and alternatives to multiple regression for estimating moderating effects of categorical variables. *Organizational Research Methods*, 2, 315-339.
- Aiken, L. S. & West, S. G. (1991). *Multiple Regression: Testing and Interpreting Interactions*. Newbury Park, CA: Sage.
- Anderson, N., & Schalk, R. (1998). The psychological contract in retrospect and prospect. *Journal of Organizational Behavior*, 19(SI), 637-647.
- Baron, R. M., & Kenny, D. A. (1986). The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. *Journal of Personality and Social Psychology*, 51(6), 1175-1182.
- Beard, K. M., & Edwards, J. R. (1995). Employees at risk: Contingent work and the psychological experience of contingent workers. *Trends in Organizational Behaviour*, 2, 109-126.

- Bellmann, L., Dahms, V., & Wahse, J. (2004). IAB Betriebspanel Ost – Ergebnisse der achten Welle 2003 – Teil II [IAB organizational panel East – Results of the eighth wave 2003 – part II]. *IAB Forschungsbericht 3*, 2004, Bundesagentur für Arbeit: Nürnberg.
- Benz, D. (2002). *Motivation und Befinden bei betrieblichen Veränderungen. Zur Bedeutung unterschiedlicher Facetten von Arbeitsplatzunsicherheit* [Motivation and mental condition in organizational change processes. The meaning of different facets of job insecurity]. Weinheim: Beltz Psychologie Verlags Union.
- Bernhard-Oettel, C., Sverke, M., & De Witte, H. (2005). Comparing three alternative types of employment with permanent full-time work: How do employment contract and perceived job conditions relate to health complaints? *Work and Stress*, 19(4), 501-518.
- Borg, I. (1992). Überlegungen und Untersuchungen zur Messung der subjektiven Unsicherheit der Arbeitsstelle. *Zeitschrift für Arbeits- und Organisationspsychologie*, 36(3), 107-116.
- Burchell, B. J., Day, D., Hudson, M., Lapido, D., Makelew, R., Nolan, J. P., Reed, H., Wichert, C. I. & Wilkinson, F. (1999). *Job Insecurity and Work Intensification: Flexibility and the Changing Boundaries of Work*. York Publishing Services: York.
- Büssing, A. (1999). Can Control at Work and Social Support Moderate Psychological Consequences of Job Insecurity? Results from a Quasi-experimental Study in the Steel Industry. *European Journal of Work and Organizational Psychology*, 8, 219-242.
- Cavanaugh, M. A., & Noe, R. A. (1999). Antecedents and consequences of relational components of the new psychological contract. *Journal of Organizational Behavior*, 20, 525-540.
- Cheng, G. H. L., & Chang, D. K.-S. (2008) Who suffers more from job insecurity? A meta-analytic review. *Applied Psychology: An International Review*, 57(2), 272-303.
- Claes, R. (2008) HR Practice and Workforce Prospects. *Zeitschrift für Personalpsychologie*, 7(4), 170-177.
- Claes, R., De Witte, H., Schalk, R., Guest, D., Isaksson, K., Krausz, M., et al. (2002). Het psychologisch contract van vaste en tijdelijke werknemers (The psychological contract of permanent and temporary workers). *Gedrag en Organisatie*, 15, 436-455.
- Conway, N., & Briner, R. B. (2002). Full-time versus part-time employees: Understanding the links between work status, the psychological contract, and attitudes. *Journal of Vocational Behavior*, 61, 279-301.
- Cook, J., & Wall, T. (1980). New work attitude measures of trust, organizational commitment and personal need non-fulfilment. *Journal of Occupational Psychology*, 53, 39-52.
- Cooper, C. L. (1999). The changing psychological contract at work. *European Business Journal*, 11, 115-119.
- Coyle-Shapiro, J. A.-M., & Kessler, I. (2002). Contingent and non-contingent working in local government: Contrasting psychological contracts. *Public Administration*, 80, 77-101.
- Crampton, S. M., & Wagner, J. A. (1994). Percept-percept inflation in micro-organizational research: an investigation of prevalence and effect. *Journal of Applied Psychology*, 79, 67-76.
- De Cuyper, N., De Jong, J., De Witte, H., Isaksson, K., Rigotti, T. & Schalk, R. (2008). Literature review of theory and research on the psychological impact of temporary employment: Towards a conceptual model. *International Journal of Management Reviews*, 10(1), 25-51.
- De Cuyper, N. & De Witte, H. (2005). Job Insecurity: Mediator or moderator of the relationship between contract type and well-being. *South African Journal of Industrial Psychology*, 31(4), 79-86.
- De Cuyper, N. & De Witte, H. (2006). The impact of job insecurity and contract type on attitudes, well-being and behavioural reports: A psychological contract perspective. *Journal of Occupational and Organizational Psychology*, 79, 395-409.
- De Cuyper, N. & De Witte, H. (2007). Job insecurity in temporary versus permanent workers: Associations with attitudes, well-being, and behaviour. *Work and Stress*, 21(1) 65-84.
- De Cuyper, N., Isaksson, K., & De Witte, H. (2005). *Employment contracts and Well-being among European Workers*. Hampshire, UK: Ashgate Publishing.
- De Witte, H. (Ed.) (2005). *Job Insecurity, union involvement and union activism*. Aldershot: Ashgate.
- De Witte, H. & Näswall, K. (2003). 'Objective' vs. 'Subjective' Job Insecurity: Consequences of Temporary Work for Job Satisfaction and Organizational Commitment in Four European Countries. *Economic and Industrial Democracy*, 24, 149-188.
- De Witte, H. (1999). Job insecurity and psychological well-being: Review of the literature and exploration of some unresolved issues. *European Journal of Work & Organizational Psychology, Special Issue: Job insecurity*, 8, 155-177.
- De Witte, H. (2000). Arbeidsethos en jobonzekerheid: meting en gevolgen voor welzijn, tevredenheid en inzet op het werk. In R. Bouwen, K. De Witte, H. De Witte & T. Taillieu (Eds.), *Van groep naar gemeenschap. Liber Amicorum Prof. Dr. Leo Lagrou* (pp. 525-550). Leuven: Garant.

- De Witte, H., Näswall, K., Chirumbolo, A., Goslinga, S., Hellgren, J., & Sverke, M. (2004). Consequences of temporary work and job insecurity in four European countries / Consequenties van tijdelijk werk en baanonzekerheid in vier Europese landen. *Gedrag en Organisatie*, *17*, 165-186.
- Dekker, S., & Schaufeli, W. (1995). The effects of job insecurity on psychological health and withdrawal: A longitudinal study. *Australian Psychologist*, *30*, 57-65.
- Guest, D., & Clinton, M. (2005). Contracting in the UK : Current Research Evidence on the Impact of Flexible Employment and the Nature of Psychological Contracts. In N. De Cuyper, K. Isaksson & H. De Witte (Hrsg.), *Employment contracts and Well-being among European Workers* (pp. 201-224). Hampshire, UK: Ashgate Publishing.
- Guest, D. & Conway, N. (2000). *The psychological contract in the public sector*. London: CIPD.
- Herriot, P., & Pemberton, C. (1995). *New Deals*. Chichester: John Wiley & Sons.
- Hiltrop, J.-M. (1995). The Changing Psychological Contract: The Human Resource Challenge of the 1990s. *European Management Journal*, *13*, 286-294.
- Jacobson, D., & Hartley, J. (1991). Mapping the context. In J. Hartley, D. Jacobson, B. Klandermans, & T. Van Vuuren. *Job insecurity: Coping with jobs at risk* (pp. 1-22). London: Sage Publications.
- Kinnunen, U., Mauno, S., Nätti, J., & Happonen, M. (1999). Perceived job insecurity: A longitudinal study among Finnish employees. *European Journal of Work and Organizational Psychology*, *8*, 243-260.
- Klein Hesselink, D. J. & Van Vuuren, T. (1999). Job Flexibility and Job Insecurity: The Dutch Case. *European Journal of Work and Organizational Psychology*, *8*, 275-295.
- Koh, W. L., & Yer, L. K. (2000). The impact of the employee-organisation relationship on temporary employees' performance and attitude: testing a Singaporean sample. *International Journal of Human Resource Management*, *11*, 566-587.
- Landsbergis, P. A. (1988). Occupational stress among health care workers: A test of the job demands-control model. *Journal of Organizational Behavior*, *9*, 217-239.
- Lord, A., & Hartley, J. (1998). Organizational commitment and job insecurity in a changing public service organization. *European Journal of Work and Organizational Psychology*, *7*, 541-554.
- Mauno, S., Kinnunen, U., Makikangas, A., & Natti, J. (2005). Psychological consequences of fixed-term employment and perceived job insecurity among health care staff. *European Journal of Work and Organizational Psychology*, *14*, 209-237.
- McDonald, D. J., & Makin, P. J. (2000). The psychological contract, organisational commitment and job satisfaction of temporary staff. *Leadership & Organisation Development Journal*, *21*, 84.
- Millward, L. J., & Brewerton, P. M. (2000). Psychological contracts: Employee relations for the twenty-first century? *International Review of Industrial and Organizational Psychology*, *15*, 1-61.
- Millward, L. J., & Hopkins, L. (1998). Psychological contracts, Organizational and Job commitment. *Journal of Applied Social Psychology*, *28*, 1530-1556.
- Mohr, G. (1997). *Erwerbslosigkeit, Arbeitsplatzunsicherheit und psychische Befindlichkeit*. (Unemployment, job insecurity, and psychological well-being). Frankfurt/Main: Lang
- Mohr, G. (2000). The changing significance of different stressors after the announcement of bankruptcy: A longitudinal investigation with special emphasis on job insecurity. *Journal of Organizational Behavior*, *21*, 337-359.
- Mohr, G., Rigotti, T., & Müller, A. (2005). Irritation - ein Instrument zur Erfassung psychischer Beanspruchung im Arbeitskontext. Skalen- und Itemparameter aus 15 Studien [Irritation - an instrument assessing mental strain in working contexts. Scale and item parameters from 15 studies]. *Zeitschrift für Arbeits- und Organisationspsychologie*, *49*(1), 44-48.
- Müller, G., Mohr, G. & Rigotti, T. (2004). Differentielle Aspekte psychischer Beanspruchung aus Sicht der Zielorientierung. Die Faktorstruktur der Irritations-Skala [Differential aspects of psychological strain with respect to goal orientation. The factor-structure of the Irritation-Scale.]. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *25*, 213-225.
- Näswall, K., & De Witte, H. (2005). Who feels insecure in Europe? Predicting Job Insecurity from Background Variables. *Economic and Industrial Democracy*, *24*, 189-215.
- OECD. (1999). *Classifying Educational Programmes. Manual for ISCED-97 Implementation in OECD Countries, 1999 Edition*.
- OECD. (2002). *Employment outlook*. Paris: Organization for Economic Co-operation and Development.

- Orpen, C. (1993). Correlations between job insecurity and psychological well-being among white and black employees in south africa. *Perceptual and Motor Skills*, 76, 885-886.
- Pearce, J. L. (1998). Job Insecurity is Important, but Not for the Reason you Might Think: The Example of Contingent Workers (pp. 31-46). In C.L. Cooper & D.M. Rousseau (Eds.). *Trends in Organizational Behavior*. New York: John Wiley.
- Price, J. L. (1997). Handbook of organisational measurement. *International Journal of Manpower*, 18, 505-558.
- Robinson, S. L., Kraatz, M. S., & Rousseau, D. M. (1994). Changing obligations and the psychological contract: A longitudinal study. *Academy of Management Journal*, 37, 157-152.
- Roskies, E., & Louis-Guerin, C. (1990). Job insecurity in managers: Antecedents and consequences. *Journal of Organizational Behavior*, 11, 345-359.
- Roskies, E., Louis-Guerin, C. & Fournier, C. (1993). Coping with job insecurity: How does personality make a difference? *Journal of Organizational Behavior*, 14, 617-630.
- Rousseau, D. M. (1989). Psychological and Implied Contracts in Organizations. *Employee Responsibilities and Rights Journal*, 2, 121-139.
- Rousseau, D. M. (1995). *Psychological contracts in organizations. Understanding written and unwritten agreements*. Thousand Oaks: SAGE Publications.
- Schaufeli, W. (1992). Unemployment and mental health in well and poorly educated school leavers. In C. Verhaar & L. Jansma (eds.), *On the Mysteries of Unemployment: Causes, consequences and policies*. (pp. 253-271). Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Silla, I., Garcia, F. J., & Peiró, J. M. (2005). Job Insecurity and Health-Related Outcomes among Different Types of Temporary Workers. *Economic and Industrial Democracy*, 26, 89-117.
- Statistisches Bundesamt Deutschland (2004). *40% der Erwerbstätigen unter 20 Jahren haben einen Zeitvertrag* [40% of employees younger than 20 years have temporary contracts]. [Online]. Retrieved November 16, 2005, from [http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Presse/pm/2005/04/PD05\\_\\_195\\_\\_135.psm1](http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Presse/pm/2005/04/PD05__195__135.psm1)
- Statistisches Bundesamt Deutschland (2005). *Erwerbstätigkeit* [Gainful employment]. [Online]. Retrieved April 13, 2005, from [http://www.destatis.de/themen/d/thm\\_erwerbs.php](http://www.destatis.de/themen/d/thm_erwerbs.php)
- Sverke, M., Hellgren, J. & Näswall, K. (2002). No security: A meta-analysis and review of job insecurity and its consequences. *Journal of Occupational Health Psychology* 7, 242-264.
- Sverke, M., Gallagher, D. G., & Hellgren, J. (2000). Alternative Work Arrangements: Job Stress, Well-being, and Work Attitudes among Employees with Different Employment Contracts. In K. Isaksson, L. Hogstedt, C. Eriksson & T. Theorell (Eds.), *Health effects of the new labour market*. New York: Plenum.
- Sverke, M., Hellgren, J., Näswall, K., Chirumbolo, A., De Witte, H. & Goslinga, G. (2004). *Job Insecurity and Union Membership. European Unions in the Wake of Flexible Production*. Peter Lang: Bruxelles.
- Van Dyne, L. & Ang, S. (1998). Organizational citizenship behavior of contingent workers in Singapore. *Academy of Management Journal*, 41, 692-703.
- Van Vuuren, T., Klanderman, B., Jacobson, D., & Hartley, J. (1991). Predicting Employees' Perceptions of Job Insecurity. In J. Heartley, D. Jacobson, B. Klandermans, & T. van Vuuren (Eds.). *Job Insecurity, Coping with Jobs at Risk* (pp. 65-78). Sage Publications: London.
- Virtanen, P., Vahtera, J., Kivimäki, M., Pentti, J., Ferrie, J. (2002). Employment security and health. *Journal of Epidemiology and Community Health*, 56, 569-574.
- Warr, P. (1990). The measurement of well-being and other aspects of mental health. *Journal of Occupational Psychology*, 63, 193-210.

Correspondence to:  
 Thomas Rigotti  
 University of Leipzig  
 Institute of Psychology II  
 Work- and Organisational Psychology  
 Seeburgstraße 14-20  
 D-04105 Leipzig  
 rigotti@uni-leipzig.de

# Kritik am Dogma „Simple Structure“

Suibert Ertel

Georg-August-Universität Göttingen / Georg-Elias-Müller-Institut für Psychologie

## ZUSAMMENFASSUNG

Die Studie hat zum Ziel, einen methodologischen Irrtum aufzudecken, der die faktorenanalytische Forschung seit ihren Anfängen belastet. Das Modell der *Simple Structure* (Thurstone, 1935/1947), an dem sich die Forschung ausschließlich orientiert hat, wird infrage gestellt. Es wird aufgezeigt, dass durch die Rotation der Faktoren zur Einfachstruktur die Komplexität der Erlebens- und Verhaltensbedingungen, die in der Regel vorliegt, systematisch verschleiert wird, womit schwerwiegende Probleme entstehen. Circumplex- und Strukturgleichungs-Verfahren lösen die Probleme nicht, da auf sie die Fehler der Einfachstruktur übertragen werden. Die Kritik an Simple Structure wird am empirischen Beispiel erläutert, theoretisch begründet und mit zahlreichen wörtlichen Zitaten einer Minorität kritischer Autoren gestützt. Insgesamt wird die Kalamität der konventionellen Faktorenanalyse einem einseitig mathematischen Denken verantwortlicher Methodiker zur Last gelegt, welches andere legitime Erkenntnismittel, nicht zuletzt den von Vorurteilen oft weniger verstellten Common sense, vernachlässigt. Ein Paradigmenwechsel wird für längst fällig gehalten. Als Alternative zur konventionellen Faktoren-Transformation wird eine Transformation zur Komplexstruktur gefordert, die in einem Folge-Artikel vorzustellen ist.

## Schlüsselwörter

Faktorenanalyse – Einfachstruktur – Komplexität – Komplexstruktur – Rotation – Varimax – Sparsamkeit – positive manifold – Big Five – Forschungsprogramme – mathematische Erkenntnismittel

## ABSTRACT

Critique of the dogma „Simple Structure“

The present study aims at disclosing a methodological error which has prejudiced factor analysis since its beginning. Simple structure, the guiding principle for factor rotation (Thurstone, 1935/1947) is unveiled as a questionable dogma it generally distorts the representation of latent sources of variance of manifest empirical variables. This criticism is elucidated by an empirical example, it is based on theoretical considerations and supported by verbatim quotations of critical authors. The present calamity of factorial research is deemed due to flaws of methodical reasoning. One-sided mathematical formalization in the discipline has lost its objectives by unjustifiably ignoring ordinary sources of gaining knowledge, including common sense. The problem of simple structure cannot be solved by circumplex- and structural equation procedures which suffer no less from the flaw of simple structure. A paradigm change is overdue. An alternative factor transformation leading to complex structures is demanded and will be delineated in a succeeding paper.

## Keywords:

Factor analysis – simple structure – complexity – complex structure – factor rotation – varimax – parsimony – positive manifold – big five – Lakatos' research programmes – mathematical tools for discovery

**„Dinge zu bezweifeln, die ganz ohne weitere Untersuchung jetzt geglaubt werden, das ist die Hauptsache überall“**

Georg Christoph Lichtenberg (1742-1799)

## 1 Einleitung

Mit den Ergebnissen der konventionellen exploratorischen Faktorenanalyse (EFA) sind ihre Anwender bekanntlich oft nicht zufrieden. Warum leistet die EFA nicht das, was man von ihr erwartet? Die vorliegende Studie hat zum Ziel, einen methodologischen Irrtum aufzudecken, der die faktorenanalytische Forschung seit ihren Anfängen belastet. Für irrtümlich wird gehalten, dass durch Rotation der Faktoren nach dem Prinzip der *Simple Structure*, das durch Thurstone (1935/1947) eingeführt wurde, empirische Variablen domänen sachgerecht modelliert werden. Das Thurstone-Prinzip gibt der Bedingungskomplexität beobachtbarer Variablen keine Chance, sich modellhaft zu manifestieren. Diese Kritik rührt an ein Dogma, auf welches sich die Anwender der Faktorenanalyse beim Faktorenrotieren haben festlegen lassen. Alternative Modelle werden nicht erwogen. Im Folgenden werden nicht nur die Ansichten des Verfassers dargestellt, sondern – meist mit wörtlichen Zitaten – auch die Ansichten kritischer Autoren, die er als Wegbereiter ausfindig machen konnte, deren Stimmen aber von der Majorität der Fachvertreter ignoriert wurden. Eine Alternative zur konventionellen Faktorentransformation, eine Transformation zur Komplexstruktur, wird in einem Folge-Artikel entworfen und dort anhand von Anwendungsbeispielen erläutert. Der vorliegende Beitrag deckt Fehler des methodischen Denkens auf, die man sich hätte ersparen können, hätten die Verantwortlichen ihre formalen Techniken auf den Gegenstand, auf den sie angewendet werden, sinnvoll bezogen und dem gesunden Menschenverstand ein Mitspracherecht eingeräumt, für den die Bedingungskomplexität aller Erlebens- und Verhaltensphänomene selbstverständlich ist.

## 2 Zur Lage der faktorenanalytischen Forschung

Die Faktorenanalyse leidet unter chronischen Anomalien. Düstere Titel und Metaphern enttäuschter Autoren sind dafür ein Symptom: *Unbehagen in der Faktorenanalyse* (Kallina, 1967); *erschreckende Unverbindlichkeit, Dilemma und Aufbaukrise* (Orlik, 1967); *Subjectivity in factor analysis* (Horn, 1967); *Zwiespältigkeit faktorieller Untersuchungen* (Meili, 1969); *Destruction of generality* (Davies, 1971); *Zufallsprodukte und*

*Scheinrealisationen* (Greif, 1972); *Unsinniger Aufwand* (Revenstorf, 1978); *Rotationsambiguität von Faktorenladungen* (Buse & Pawlik, 1978); *Dubious legacy* (Schönemann, 1981); *Faktoranalytis* (Jäger & Hörmann, 1981); *Mythos Faktorenanalyse* (Lenk, 1983); *Factors are fictions*, Revelle (1983); *Morass of factor analysis* (Eysenck, 1992); *Psychopathology of factor indeterminacy* (Schönemann, 1996)<sup>1</sup>; *Pathology of psychometrics* (Borsboom, 2005). Die Mängel der Methode werden von der Mehrheit der Experten bagatelisiert und gern mit Optimismus überdeckt: Es gebe doch ‚bedeutsame Ergebnisse‘, meint Pawlik (1977) in einem umfassenden Überblicksartikel über die ersten Jahrzehnte der faktorenanalytischen Forschung.

Aus der Sicht unkonventionell eingestellter Autoren, die der Misere ins Auge sehen, ist die faktorenanalytische Forschung fehl gelaufen: *„Exploratory factor analysis has never been developed to anything approaching its full promise and potential, despite the eighty year history of its efforts...“* (Yates, 1987, p. 325). Blinkhorn (1997) behandelt in einem Überblick über *„fifty years of test theory“* die Faktorenanalyse und stellt fest, dass auch die *„considerable technical strides“* der letzten Jahrzehnte und die *„well-known contributions of Jöreskog and McDonalds“* die Misere im Grunde nicht verändert haben: *„How curious ... that we are so little further forward in our understanding of the psychology of individual differences as a result of these advances ... Can anyone identify a single publication in the last 50 years in which the use of factor analysis has led to counter-intuitive, or surprising, or genuinely enlightening outcomes? ...“* (Blinkhorn, 1997, p. 181). Schon vor 50 Jahren hätte man aufmerken können bei dem, was Schönemann (1994) retrospektiv über Louis Guttman berichtete, der 1955 eine ‚Grabrede‘ [eulogy] über die multiple Faktorenanalyse hielt: *„It was left to Louis Guttman to read the eulogy (p. 209, p. 406): The era of Multiple Factor Analysis had come to an end – for knowledgeable people at any rate. ... It was logical, then, to ask: What lies ahead for Factor Analysis? (Guttman, 1958). He answered it with a vision that challenged habits of thought that had lead nowhere. This vision he kept pursuing for the rest of his life.“* Schönemann & Borg (1996, p. 249) ziehen Bilanz: *„Heute weiß man, dass die von Thurstone eingeläutete Ära der ‚explorativen Faktorenanalyse‘ wenige Einsichten von Bestand gebracht hat“*.<sup>2</sup>

Zwei chronische Kalamitäten der EFA sind zu beklagen:

1. *EFA-Forschung hat zu einer Unzahl von Konstrukten in der Psychologie geführt und somit das Gegenteil von dem erreicht, was man mit ihr erreichen wollte.*

Die EFA hat zum Ziel, die Vielzahl miteinander korrelierter manifester Variablen durch eine geringere Zahl faktorieller Variablen, denen die Rolle *latenter Dimensionen* zugesprochen wird, sparsam und theoretisch vorteilhaft zu beschreiben.<sup>5</sup> Nun hat sich in den Jahrzehnten der EFA-Forschung herausgestellt, dass die Produktion faktorieller Variablen, die den beobachtbaren Variablen zugrunde liegen sollen, ihrerseits offenbar unerschöpflich ist. Lee Sechrest wies in einem forschungskritischen Vortrag auf der 11<sup>th</sup> European Conference on Personality (2002) in Jena auf die Unzahl von Konstrukt-Variablen in der Psychologie hin, die er nach einem Autor, der sie hochrechnete, auf 7,800 bezifferte, viele davon, die meisten vermutlich, sind Neuschöpfungen exploratorischer Faktorenanalysen.

Ist Sechrest's Schätzung unrealistisch? Nach einem Internet-Download aller Artikel mit dem Stichwort „Scale“ aus der Datenbank PubMed fanden sich immerhin 889 unterschiedliche Skalen-Bezeichnungen, von „*Abel and Becker Cognition Scale*“ bis „*Zung Self Rating Anxiety Scale*“. Die meisten Skalen sind faktorenanalytisch erarbeitet oder bearbeitet worden, so dass man bei einer geschätzten Faktorenzahl von 3 bis 4 pro Skala etwa 5,000 faktorielle Konstrukte ansetzen kann, die bisher allein im klinisch-psychologischen und medizinischen Raum entstanden sind. Hinzuzufügen sind die in der nicht-klinischen differentiellen Psychologie entwickelten Skalen und die aus ihnen gewonnenen Konstrukte, von denen allein die Zeitschrift *Personality and Individual Differences* mit jedem Heft so viel Neues bietet, dass der oben genannte Schätzwert realistisch erscheint. „*The idle practice of producing new personality scales continues unabated, making it less likely that they will ever arrive in the promised land of the paradigm which alone would endow our efforts with scientific respectability*“ (Eysenck, 1992, p. 672).<sup>4</sup>

2. *Den faktoriellen Konstrukten fehlen schon innerhalb der Analysen, die sie hervorbringen, theoretisch verwertbare Verknüpfungen.*

Ebenso unerfreulich wie die Flut faktorieller Konstrukte ist für theoretisch Denkende deren Mangel an innerer Verbundenheit. Die Produktion unverknüpfbarer Elemente führt zu bloßen Aggregaten. Erst durch Beziehungen untereinander können sich Begriffe, die bei äußerlich bleibender Betrachtung isoliert erscheinen, in eine Ordnung fügen. Die faktoriellen Konstrukte der Psychologie lassen sich jedoch kaum miteinander verknüpfen. „... *Simple structure may have resulted in an analysis into too many unrelated, and UNRELATABLE, primaries [primary factors]*“ (so Stephenson schon 1956, p. 6, Großbuchstaben durch SE). Der derzeit umfassendste kritisch-historische Überblick über die persönlichkeitspsychologische

EFA-Forschung von Andresen (1998) vermittelt den Eindruck einer chaotischen Ergebnisvielfalt.

Hat das „Big Five“ Faktoren-Modell, das in der Persönlichkeitsforschung auf faktorenanalytischem Wege entstand, stürmisch begrüßt wurde und hohes Ansehen erreichte<sup>5</sup>, den theoretischen Mangel beseitigt? Man hat sich damit begnügt, in der „chaotic plethora of personality constructs“ (Funder, 2001, p. 200) fünf leidlich invariante Konstrukte herauszustellen. Doch ist Invarianz kein hinreichendes Kriterium für die Güte von Konstrukten, die auch als fehlerhafte sehr robust sein können. Weit überzogen ist der Optimismus einiger Big-Five Proponenten, die die Rolle ihrer Faktoren in Analogie setzen zur Rolle der chemischen Elemente (begonnen hat damit offenbar Goldberg, 1981), deren Entdeckung im 19. Jahrhundert eine naturwissenschaftliche Revolution auslöste.<sup>6</sup> So wie sich aus chemischen Elementen Moleküle zusammensetzen, sollten sich komplexere differentielle Konstrukte der Psychologie aus den fünf elementareren Konstrukten zusammenfügen. Dies glaubten z. B. einige Optimisten, die die scheinbar elementaren „großen Fünf“ durch kombinatorische Verwendung in drei Superkonstrukte überführen wollten („Resilience“, „Overcontrol“ und „Undercontrol“, s. Caspi, 1998, Robins et al., 1998, Asendorpf & van Aken, 1999). Die Meinung hat sich verbreitet, dass sich die zukünftige Faktorenproduktion im Persönlichkeitsbereich durch Korrelationen mit den Big Five zu legitimieren habe.<sup>7</sup>

Doch „trägt dieser Vergleich [mit den chemischen Elementen] nicht weit“ (Lukesch & Kleiter, 1974, p. 294). Denn H, He, Li, Be, B, C, N usw. haben einen definierbaren Platz innerhalb eines Systems. Dass sich aus Atomen Moleküle bilden, ist den atomaren Bindungseigenschaften zu verdanken, die sich aus der Zahl der Protonen in den Atomkernen, aus der Elektronendichte usw. ergeben. Aufgrund gleichartiger Bestandteile ihres Aufbaus stehen chemische Elemente miteinander in Beziehung. Für die Big Five-„Dimensionen“ der Persönlichkeit aber sind keine gleichartigen Bestandteile erkennbar, die zu einem Ähnlichkeits- oder Verschiedenheitsurteil berechtigten. Nur wenige kritische Beobachter nehmen daran Anstoß, z. B. Briggs, 1989, und Block, 1995<sup>8</sup>. Die „großen Fünf“ sind wie alle anderen Konstrukt-Produkte der Faktorenanalyse Unikate, eine Analogie zu den Elementen eines Dimitri Mendelejev (1834-1907) und Lothar Meyer (1830-1895) ist verfehlt. Eher könnte man die Faktoren der EFA den vermeintlich unteilbaren Atomen Heraklits (B.C. ca. 540-480) gleich setzen, die ein ‚atomistisches‘ Denken festschrieben, bis es mit den Entdeckungen der atomaren und subatomaren Strukturen aufzugeben war.<sup>9</sup>

An dieser Stelle wäre auf Eysenck einzugehen, der auf die atheoretischen Auswüchse der EFA-For-

schung sensibel reagierte. Den Big Five-Forschern warf er vor, Faktorenanalysen mit lediglich taxonomischer Abgrenzungsabsicht zu betreiben und dabei an der psychometrischen Oberfläche zu kleben, anstatt biologisch-kausal interpretierbare Zusammenhangsmodelle zu entwickeln (Eysenck, 1992, 1997). Eysencks eigener Versuch wäre anzuführen, der eine Verbindung unter den drei Dimensionen seines PEN-Modells (Psychoticism, Extraversion, Neuroticism) zu erreichen suchte, indem er sie für eine gemeinsame Hirngrundlage postulierte und den Dimensionen unterschiedliche anatomische Strukturen und neuro-psychologische Funktionen zuwies, die zusammenwirken sollten. So begrüßenswert auch Eysencks Ziel grundsätzlich war – er hat es nicht erreicht.<sup>10</sup> Die eigentliche Ursache der von ihm beklagten „morass of factor analysis“ (1992, p. 672) hat er nicht erkannt. Sie liegt nicht dort, wo er sie vermutete.

Faktorenanalytische Datenverarbeitung findet man auch in zahlreichen nichtpsychologischen Disziplinen (vgl. Rummel, 1970, Reymert & Jöreskog, 1995, und Abbildung 1), Unbehagen wird nicht selten auch dort bekundet. So hört man z. B. aus den Reihen der Geowissenschaftler, die sich methodisch der Einfachstruktur-orientierten EFA bedienten, etwa bei Davies (1971): „*[By varimax] we may be butchering our results; cutting up the body of generality into a set of unrelated fragments without ever realizing that these fragments can ever be considered as part of a larger entity*“ (p. 113). Davies bezeichnet die Wirkung der Varimax-Rotation auf faktorisierte Datensätze mehrmals als „zerstörerisch“. Dies wird in Abschnitt 5 näher beleuchtet.

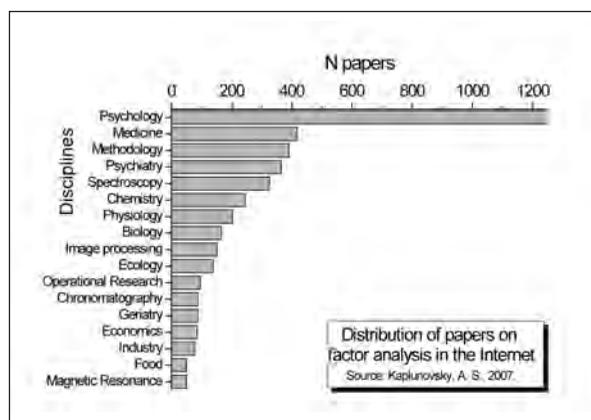


Abbildung 1: Anzahl faktorenanalytischer Arbeiten differenziert nach Disziplinen ermittelt aus Internetquellen von A. S. Kaplunovsky (2007)

### 3 Das Dogma: Prinzip der Einfachstruktur

Der Bericht über die beklagenswerte Lage der faktorenanalytischen Forschung hatte zum Ziel, ein Verständnis für die Wurzel des Übels vorzubereiten. Sie liegt nach meiner Auffassung, wie angekündigt, in der Fragwürdigkeit des Prinzips *Simple Structure*. Das Prinzip, von Thurstone 1935 eingeführt und 1947 von ihm festgeschrieben, sollte der Transformation initialer Faktoren der Faktorenanalyse als Richtmaß dienen. An ihm hat sich seit ca. 1960 die EFA-Forschung mit all ihren methodischen Verzweigungen orientiert. Auch hat man bei den später entwickelten konfirmatorischen Verfahren, mit denen man sich vom dominanten Simple Structure-Prinzip leichter hätte lösen können, an diesem wie an etwas Unverzichtbarem festgehalten. Als Ursache für das verbreitete „Unbehagen in der Faktorenanalyse“ hat man meist Subjektivität und Willkür bei der Auswahl von Verfahrensweisen angesehen, so etwa bei der Auswahl der zu korrelierenden Variablen, bei der Auswahl der zu testenden Personenstichproben, bei der Entscheidung über die Zahl der zu extrahierenden Faktoren, bei der Entscheidung über schief- oder geradwinklige Rotation, bei der dann immer noch anstehenden Auswahl aus einer Vielzahl von Verfahren (s. Finch & West 1997, p. 464 ff). Doch sind – nach meiner Auffassung – die Auswirkungen dieser Entscheidungswillkür auf die Endergebnisse der faktoriellen Analysen geringfügig gegenüber dem, was die Misere in der EFA-Forschung tatsächlich hervorruft und worüber im Folgenden zu berichten ist.<sup>11</sup>

Der Grundgedanke der Einfachstruktur ist weithin bekannt, er wird im Studiengang für Diplom-Psychologen vermittelt, für Lehrer und Schüler ist er „intuitively compelling“ (Kaiser, 1958, p. 188). Im Koordinatensystem der initialen Faktoren ergeben sich Häufungen von Variablenpunkten (s. Abbildung 2A), die den Statistiker veranlassen, durch die Punktwolken eine Regressionslinie mit kleinstem Quadrat der Punktabstände zu ziehen. Mit diesem Ziel wird das System der Koordinaten so gedreht, dass diese mit dem Regressionsvektor zusammen fallen (Abbildung 2B). Da nacheinander immer zwei Koordinaten gleichzeitig gedreht werden, wird bei einer Drehung, wenn mehrere Punktwolken in der Ebene vorhanden sind, eine gleichzeitige Annäherung beider Koordinaten optimiert. Die Konsequenteren unter den Einfachstruktur-Methodikern übertreffen dabei die rechtwinklig rotierenden Forscher an Radikalität, indem sie Schiefwinkligkeit in Kauf nehmen, um eine maximale Annäherung an die Einfachstruktur zu erreichen.<sup>12</sup>

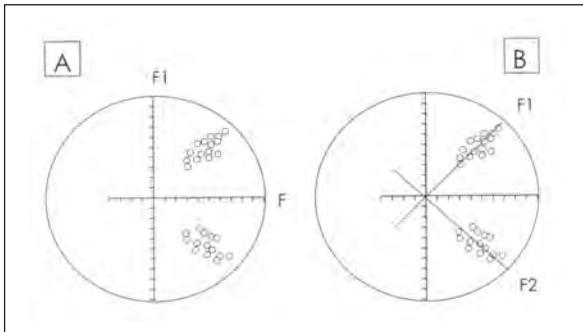


Abbildung 2: Zwei Punktwolken im initialen Faktorensystem  $F1$ ,  $F2$  (A) und nach einer orthogonalen Rotation zur Einfachstruktur (B)

Bei einem z. B. zweidimensionalen Koordinatensystem mit zwei Punktmengen im Abstand eines rechten Winkels sind nach Annäherung der Koordinaten an deren Schwerpunkte die einzelnen Variablen nur durch je einen Faktor mathematisch beschreibbar, während sie bei der initialen Lösung durch zwei Faktoren zu beschreiben wären. Mit anderen Worten: Ergeben sich zwei Faktoren, dann wird diejenige Lösung für einfacher gehalten, bei der die Variablen mehrheitlich auf nur einem der beiden Faktoren laden.

Den Vereinfachungsschritt, der bei der Zuteilung von Faktoren zu einzelnen Variablen aus ‚Sparsamkeit‘ vorzunehmen sei, hält man im übrigen für eine zwangsläufige Fortsetzung der Vereinfachungsstrategie, die zur Reduktion der Variablenvielfalt auf wenige Faktoren führte: „Im Faktorenproblem bestand das Streben nach Einfachheit darin, eine möglichst kleine Anzahl von Faktoren zu finden ... Beim Rotationsproblem handelt es sich darum, innerhalb des ... festgelegten ... Faktorensystems das Zueinander der Variablen und Faktoren so einfach als möglich zu gestalten“ (Überla, 1971, p. 176). Am Prinzip der Vereinfachung im Sinne von ‚weniger ist besser als mehr‘, dem die Faktorenextraktion folgte, wird also bei der anschließenden Faktorenrotation festgehalten, so als ob man das gleiche Ziel auch hier verfolgen müsse.

#### 4 Die Folgen des Irrtums

Indem bei der Datenverarbeitung eine Einfachheit durchgesetzt wird, die dem formalen Modelldenken entgegenkommt, handelt man sich für das Verständnis der durch das Modell abzubildenden Sachverhalte üble Folgen ein. Überla gibt dafür ungewollt ein lehrreiches Beispiel: Bei 90 Männern wurde im Laufe einer halben Stunde sechs Mal der systolische und diastolische Blutdruck gemessen. Die initialen Ergebnisse einer Faktorenanalyse der 2 mal 6 = 12 Variablen sind in Abbildung 3A dargestellt, Abbildung 3B enthält die

Ergebnisse einer Varimax-Lösung dieser Faktoren, die die Einfachstruktur orthogonal optimiert. Die Lösung 3B stellt Überla als die einzig sinnvolle hin, da, wie er meint, die Variation der Messungen „im wesentlichen von zwei Einflussgrößen bestimmt wird, eben dem systolischen und dem diastolischen Blutdruck“ (p. 265).

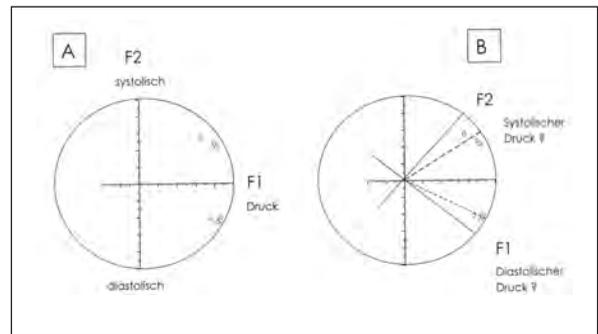


Abbildung 3: Initiale Faktorenlösung der Blutdruck-Daten (Überla) (A) und Rotation zur Einfachstruktur (B) orthogonal (A) und schiefwinklig (B)

Die Behauptung Überlas ist inakzeptabel. Tatsächlich sind doch ‚systolisch‘ und ‚diastolisch‘ keine ‚Einflussgrößen‘, sondern lediglich verschiedene Momente der Blutdruck-Verhältnisse innerhalb des pulsierenden Blutumschs. Blutdruck ist eine Einheit mit lokaler und zeitlicher Varianz. Im Moment der Systole und Diastole schlägt sich Blutdruck auf jeweils verschiedenem Niveau nieder analog dem Luftdruck in verschiedenen Höhenlagen. Zu den klinischen Abweichungen gehören die des Blutdrucks, der zu hoch oder zu niedrig sein kann. Auch ist eine therapeutische Regulation durch Medikamente nicht auf entweder systolische oder diastolische ‚Einflussgrößen‘ ausgerichtet, sondern auf eine Verbesserung des Blutdrucks als Systemmerkmal (s. Näheres im *Journal of Human Hypertension und Hypotension*).<sup>13</sup> Eine Faktorenanalyse der 12 Blutdruck-Variablen würde den physiologischen Sachverhalt dann am besten abbilden, wenn die miteinander hoch korrelierende Varianz ‚systolisch-diastolisch‘, d. h. die Einheitlichkeit des Geschehens, durch einen Generalfaktor zur Darstellung kommt. Der zu fordernde Generalfaktor ist, wie Abbildung 3A zeigt, in der Initial-Lösung durchaus vorhanden und geht durch Rotation der Koordinaten zur Einfachstruktur (3B) verloren.

Eine Differenzierung nach ‚systolisch‘ und ‚diastolisch‘ legt indessen der zweite initiale Faktor nahe. Er ist wie der Generalfaktor aus physiologischen Gründen zu erwarten, denn die beiden Blutdruckmomente zeigen zusätzliche Varianz, die auf besondere Bedingungen zurückzuführen ist. So reagieren systolische Blutdruckwerte auf äußere Einflüsse generell etwas stärker als diastolische. Die Varianz der systolischen

Werte ist bei circadianen Messwertreihen größer als die der diastolischen Werte (Halberg, 1980, Fig. 8, p. 552), beim Altershochdruck steigen die systolischen Werte stärker an als die diastolischen usw. Offenbar wird die Eigenvarianz der beiden Messmomente in der initialen Lösung durch den zweiten Faktor repräsentiert, dem ein geringerer Varianzanteil zukommt als dem ersten.

Analoges findet man auf dem Gebiet der faktoriellen Intelligenzforschung. Auch bei Faktorenanalysen zur Intelligenz wird theoretisch ein Generalfaktor  $g$  gefordert, der in aller Regel in der initialen Lösung mit großer Annäherung bereits vorhanden ist. Durch Rotation zur Einfachstruktur geht  $g$  verloren. Da aber ein  $g$ -Faktor theoretisch zu erwarten bleibt, wird dieser meist in einem umständlichen weiteren Arbeitsgang wieder hergestellt<sup>14</sup>, etwa durch Orthogonalisierung der zunächst schiefwinklig rotierten Primärfaktoren nach einem Verfahren von Schmid-Leiman (1957). So gelangt man zu Faktoren ‚zweiter Ordnung‘ und damit zu einer ‚Hierarchisierung‘ der Modellkomponenten. Damit aber werden theoretische Probleme in Kauf genommen, die man bislang ignoriert hat.<sup>15</sup>

Prägnant beschreibt A. Jensen die Wirkungen einer Rotation zur Einfachstruktur auf den initialen ersten Faktor von Intelligenztest-Analysen: „... *The tests are all positively correlated and therefore all had some factor in common – a general factor, or  $g$ . The general factor that was so prominent in the analysis depicted in Figure 3.3 [showing the initial solution] seems to have disappeared from Figure 3.4 [showing the simple structure-rotated solution] as a result of rotating the factor axes. Actually, it has simply been dispersed (or redistributed) among the rotated factors ... So if you ask where  $g$  went, the answer is that it has been divided up and lies ‚hidden‘ among all of the tests‘ smaller loadings on all of the orthogonally rotated factors. Its variance has not disappeared, it has simply been obscured by being dispersed throughout the whole factor matrix.*“ (Jensen, 1998, p. 66). Pett et al. (2005, p. 145) zitieren Autoren, die das Problem gesehen haben: „*Nunnally and Bernstein (1994) ... warn against prematurely concluding that, based on a Varimax-solution, a general factor is absent because Varimax is designed to eliminate general factors (Gorsuch, 1983). Comrey and Lee (1992) suggest that the researcher avoid including too many factors in a Varimax rotation solution because it tends to overinflate the importance of lesser factors. Although the authors do not indicate how many factors are ‚too many‘ they point out that trial and error is the only way to arrive at the appropriate number of factors.*“ Dass man den erkannten Fehler durch ein fragwürdiges Versuch-und-Irrtum-Verfahren zu korrigieren empfiehlt, kennzeichnet die verworrene Lage.

## 5 Fehleranalysen im Detail

### 5.1 Fragwürdige Abwertung initialer Faktorstrukturen

Thurstone und alle, die der ‚amerikanischen Schule‘ folgten, hielten das initiale orthogonale Faktorenmuster, das eine Zentroid-, Hauptachsen- oder Hauptkomponentenanalyse generiert, für „im allgemeinen nicht interpretierbar“ (Überla, 1971, 175), zudem für instabil. Zur vermeintlichen Instabilität der initialen Faktoren äußert sich Überla: „*[Sie, die initialen Faktoren,] ... ändern sich von Stichprobe zu Stichprobe ... da neue Variablen die Gewichte oft entscheidend verschieben und die Position der Achsen dadurch abändern*“ (p. 175). Dass faktorielle Invarianz ein besonderes Problem der initialen Faktoren sein soll, dem man durch Transformation beikommen könne, ist indessen eine kaum begründete Behauptung. In der Literatur zur Faktorenanalyse ließ sich kein Befund ausmachen, wonach sich rotierte Faktoren als invarianter erwiesen hätten als unrotierte – ein Defizit, das auch Andresen (1998) konstatiert.<sup>16</sup> Stattdessen finden sich durchaus Klagen über einen Mangel an Invarianz bei *rotierten* Faktoren (Fittkau, 1968, Butler, 1969).<sup>17</sup>

Widerspruch ist auch gegenüber der Behauptung anzumelden, initiale Faktoren seien nicht interpretierbar. Da die Variablen zumeist auf mehreren initialen Faktoren laden, so heißt es, könne man nicht wissen, was eine Variable, der mehrere Faktoren zugrunde liegen, inhaltlich darstelle. Doch darf man von zwei Interpretationen eines Sachverhalts nicht unbesehen diejenige für richtig halten, die dem Forscher verständlicher erscheint. Man müsste sonst am geozentrischen Weltbild festhalten und das kopernikanische verwerfen.

Dass initiale Faktoren nach Einfachstruktur transformiert werden sollen, wird auch damit begründet, dass unter ihnen meist ein erster Faktor herausragt. Der „Generalfaktor“ sei ein methodisches Artefakt, da er bei sehr verschiedenen Datensätzen extrahiert werde. Überla gibt eine verbreitete Auffassung wieder, wenn er sagt: „*Die [initialen] Faktoren ... weisen eine willkürliche Varianzaufteilung auf. Die Varianzaufteilung ist nicht aus den Daten abgeleitet, sondern der Methode inhärent.*“ Hier wird einer rechnerisch ermittelten Regelmäßigkeit ‚Willkür‘ zugetraut, anstatt zu vermuten, dass eine Faktorenanalyse unter den Varianzquellen, die den untersuchten Phänomenen zugrunde liegen, in der Regel tatsächlich eine dominierende zum Vorschein bringt, der man den größten Aufklärungsanteil zuzubilligen hat, während Varianzquellen ab Faktor 2 weniger Varianz aufklären.<sup>18</sup>

Aufschlussreich ist in diesem Zusammenhang, dass inhaltsorientierte Forscher trotz solcher War-

nungen der Lehrbuchautoren den ersten unrotierten Faktor einer Faktorenanalyse oft als Endergebnis verwerten. Manch ein Intelligenzforscher entdeckt im unbehandelten ersten Faktor den Niederschlag eines erwarteten Generalfaktors der Intelligenz (expliziert bei Jensen, 1998, p. 65-68). Maxwell (1972) beschränkte sich bei einer Varimaxrotation von Intelligenzfaktoren auf die Faktoren  $F_2$  bis  $F_5$  und ließ den initialen Faktor  $F_1$  unrotiert. Auch erkennen Persönlichkeitsforscher bei der Analyse von Fragebogen-Daten oft schon im ersten Faktor das Konstrukt, das sie faktoriell operationalisiert sehen wollen (angefangen mit Hamilton, 1960, Lumsden, 1961, Leventhal & Stedman, 1970). Durch eine Einfachstruktur-Rotation geht in der Regel die erwünschte initiale Lösung und damit die Repräsentanz des Konstrukts, das operationalisiert werden soll, verloren. Um den ersten initialen Faktor davor zu bewahren, mit einer Simple Structure-Rotation zu verschwinden, greifen Faktoristen gern zu einem Trick: Sie schließen bei einer Wiederholung der Analyse Fragebogen-Items mit unliebsam hohen Ladungen auf einem zweiten und dritten Faktor aus, womit der zweite oder dritte usw. Eigenwert unter das Ausschlusskriterium von Kaiser und Guttman (= 1.0) gedrückt wird. So entgeht man ohne Verletzung der Konvention dem vermeintlichen Zwang zur Faktorenrotation und seinen konstrukt-destruierenden Folgen – ein formalistisches Vorgehen, das eher der Ausnutzung juristischer Gesetzeslücken ähnelt als daß es der Ermittlung sachlicher Gegebenheiten dient.<sup>19</sup>

Methodisch sauberer gingen Gibbons & Hederer (1992) vor, die das Problem der oft sich aufdrängenden Deutbarkeit des initial auftretenden Generalfaktors mit einem mathematisch ausgearbeiteten sogenannten Zweifaktorenmodell („bi-factor model“) zu lösen suchten. Beim Zweifaktorenmodell wird ein Generalfaktor ( $F_1$ ) vorausgesetzt bzw. akzeptiert, an dem alle Variablen partizipieren, während bei jeder einzelnen Variablen einer von mehreren weiteren Faktoren ( $F_2$  ...  $F_j$ ) zugelassen wird. Es zeigte sich, dass durch das Rotationsmodell, das den Generalfaktor konserviert, bei der Verarbeitung von Daten z. B. aus einem Wissenstests und aus einem Depressionsfragebogen ein signifikant besseres Passungskriterium erreicht wurde als durch das Einfachstruktur-Modell.<sup>20</sup>

Um den Generalfaktor zu retten, wenn ein solcher gebraucht wird, sind schon sehr früh besondere Rotationstechniken entwickelt worden. Die orthogonale Quartimax-Rotation von Neuhaus & Wrigley (1954) kann dies in gewissem Maße leisten, wenn man die offensichtlichen Schwächen, die diese Methode mit sich bringt (Gorsuch, 1974, p. 191) mit in Kauf nimmt, einschließlich der weiteren Simple-Structure-Folgen, die das Quartimax-Verfahren nicht beseitigt.

Dennoch halten es Überla und alle, die seine Lehrbuch-Ansichten teilen, für unausweichlich, die vermeintlich ‚unverständlichen‘ und ‚willkürlichen‘ initialen Faktorenlösungen in vermeintlich verständlichere und stabilere Lösungen zu überführen. Simple Structure wird als die einzig vertretbare oder sogar einzig mögliche Leitlinie hingestellt, nach der sich das Koordinatensystem auszurichten habe. Alternativen werden nicht diskutiert, etwas anderes als das Prinzip der Einfachstruktur hält man offenbar für nicht denkbar.

In der Frühzeit der Faktorenanalyse fand die amerikanische Schule Thurstones noch Konkurrenz und Kritik durch die britische Schule unter der Führung von Cyril Burt, der in den Initiallösungen der Faktorenanalyse, in den „bipolaren Strukturen“, wie er sie nannte, wertvolle Informationen erkannte: „... *to the experienced factorist both the regularities and irregularities [of the pattern of signs] will yield considerable insight into the data he is analysing, even without any further rotation or analysis*“ (Burt, 1954, p. 16). Burt nahm routinemäßig zwar auch eine Umwandlung der bipolaren Initialfaktoren zugunsten einer unipolaren Lösung vor, die er ‚Gruppenfaktorlösung‘ nannte. Doch damit entfernte er sich von der Initiallösung weniger weit als die amerikanischen Simple Structure-Faktoristen. Zurückhaltung gegenüber einer verfälschenden Faktorenrotation findet man auch u. a. bei Burroughs & Miller (1961, p. 35-37): „... *the subsequent rotations are apt to obscure [the objective and dichotomous classification based directly on the data]*“ (p. 36). Ein längeres Andauern der Konkurrenz zwischen den britischen und amerikanischen Faktoristen, zu der es nicht kam, hätte die Fehlentwicklung der Thurstoneschen Position vermutlich verzögert oder auch verhindert.

## 5.2 Fehldeutung der ‚positive manifold‘

Intelligenztests korrelieren untereinander generell positiv (‚positive manifold‘). Die initialen Faktorlösungen von Test-Interkorrelationen weisen jedoch – wie die der Blutdruckdaten – außer einem unipolar positiven Generalfaktor  $F_1$  (=  $g$ ) in der Regel nur noch bipolare Faktoren  $F_2$ ,  $F_3$  usw. auf, also auch negative Ladungen. Die Intelligenzforscher der ersten Stunde, die als Faktoren ‚Fähigkeiten‘ erwarteten, hielten negative Ladungen für erwartungswidrig. Man war davon überzeugt, dass Faktoren aus Intelligenztest-Untersuchungen nur positive Ladungen, allenfalls Null-Ladungen (= keine Fähigkeit), aber keine negativen Ladungen haben könnten. Negative Ladungen hätte man als Leistungsminderung zu interpretieren, was dem Begriff der Fähigkeit widerspreche. Eine Um-

wandlung der bipolaren zur nur-positiven Ladungsverteilung erschien also notwendig. Da man glaubte, die positiven Interkorrelationen durch einheitlich positive Ladungsvorzeichen auch in der Faktorenmatrix darstellen zu müssen, hielt man die Transformation zur Einfachstruktur, die dies bewirkte, für zusätzlich gerechtfertigt (Thurstone, 1947, p. 541).<sup>21</sup>

Anhand der Blutdruck-Untersuchung wurde deutlich, dass die aufzuklärenden Sachverhalte durch das Modell der Einfachstruktur verfälscht werden. In ähnlicher Weise wird das faktorielle Endergebnis verfälscht, wenn man die positive Mannigfaltigkeit der Intelligenztest-Korrelationen in die Faktorenmatrix zu übertragen versucht. Hier führt nicht mehr die mathematische Modellierung zu fehlerhaften Sachkonzepten, vielmehr nehmen umgekehrt vorgefasste Sachkonzepte (der Fähigkeitsbegriff) Einfluss auf die mathematische Modellierung (hier: die Vorzeichenmanipulation).<sup>22</sup> Überla war bei seiner Blutdruck-FA, offenbar so wie Forscher bei einer Intelligenz-FA, davon ausgegangen, dass man nur positive Ladungen zulassen dürfe, da der niedrigste Wert einer Hg-Skala Null und nicht negativ ist. Fälschlicherweise überträgt er dabei die Eigenschaften der Metrik, die den manifesten Beobachtungen angemessen ist – beim Blutdruck wie bei Intelligenzleistungen ist eine Verhältnisskala angemessen –, auf die Metrik der latenten Bedingungen, auf der nur die Eigenschaften einer Intervall-Skala vorliegen. „... *rigorous measures, such as direct counts, latency, or duration, are excellent measures if used as descriptions of behavior but may become arbitrary metrics if they are used to infer some psychological construct*“ (Kazdin, 2006). Nach der mit fünf Beiträgen geleisteten Diskussion zur ‚arbitrary metrics‘ (American Psychologist, 61, 2006) dürfte heute Einigkeit darüber zu erzielen sein, dass auch die Metrik, die sich auf der latenten Ebene von Faktorladungen ergibt, ‚arbitrariness‘ von Nullpunkten und von Vorzeichen nicht nur erlaubt, sondern erfordert. Schon Vukovich (1967) vertrat in diesem Punkte Liberalität: „*Mit den Messwerten ist jeder Umgang erlaubt, der ihren Abbildungscharakter nicht beeinträchtigt. Wenn die empirischen Vergleichsoperationen nichts über die Nullpunktslage und die Größe der Einheiten eines Maßstabes aussagen, kann man sie frei wählen und wird sie bequemerweise so festsetzen, dass größere Bedingungsbeziehungen übersichtlich zu beschreiben sind*“ (p. 114).

Kurz: Ein initialer Generalfaktor sollte nicht als Methodenartefakt (Überla) abgewertet und beseitigt werden. Dass in einer Phänomendomäne unter jeweils differenzierbaren Varianzquellen eine mit Haupteffekten hervortritt, kann der Normalfall sein. Die verbleibende Varianz, welche weitere Faktoren aus der Residualmatrix der Korrelationen extrahierbar macht, wird durch zusätzliche Bedingungen hervor-

gerufen, die den Haupteffekt orthogonal modifizieren. Der Generalfaktor dient den nachfolgend extrahierten Varianzquellen in der Regel als Referenzniveau. Er entspricht dem Mittelwert = 0 einer standardisierten Werteverteilung, von dem ausgehend sich positive und negative Abstände ergeben.<sup>25</sup> Das Vorkommen von negativen neben positiven Ladungen bei einem zweiten, dritten usw. Faktor bedeutet also nur, dass der Einfluss zusätzlicher Varianzquellen in ihrer Größe zu den Effekten der Hauptvarianzquelle in Beziehung gesetzt werden, aus der sich, aufgrund der Intervalleigenschaft der zusätzlichen Varianzquellen, eine bipolare Verteilung ergibt. Ähnliche Überlegungen wurden bereits von Thompson (1965) angestellt, sie fanden kaum Resonanz.<sup>24</sup>

Die vorliegende Kritik an der Einschätzung der „positive manifold“ durch Thurstone und Nachfolger sollte nicht als Empfehlung missverstanden werden, die Initiallösungen von Faktorextraktionen als „non plus ultra“ einzuschätzen. Dies wird im folgenden Abschnitt 6 und unter 9 weiter ausgeführt.

## 6 Neuorientierung

### 6.1 Basismerkmal natürlichen Geschehens: Komplexität

Am Beispiel der Blutdruck-FA wurde deutlich, dass die Gesamtvarianz der erhobenen Variablen auf zwei latente Quellen aufzuteilen ist, die eine ist zuständig für den Blutdruck unabhängig von Messvarianzen der Messzeitpunkte, die andere bringt den Varianzanteil hervor, den die Messzeitpunkte zur Gesamtvarianz beitragen. Damit werden die empirischen Gründe der vorliegenden Kritik an der Einfachstruktur transparent. Man kann sie für nahezu trivial halten: Wahrscheinlich auf allen Gebieten empirischer Forschung liegen den jeweils beobachteten Variablen eine Mehrzahl zusammen wirkender Bedingungen zugrunde. Die Einfachstruktur-orientierte Faktorenanalyse operiert an dieser Trivialität vorbei. Unbestritten bleibt zwar der Sparsamkeitserfolg des ersten Schritts einer Faktorenanalyse, da den manifesten, untereinander korrelierten Variablen eine geringere Zahl latenter Varianzquellen (faktorielle ‚Bausteine‘) zugeordnet wird. Doch wird dieser Erfolg verspielt, wenn den manifesten Variablen durch Faktorentransformation möglichst jeweils nur einen latenten Baustein oder ein Minimum von diesen zugewiesen wird. Das Gegenteil wäre zu fordern (mehr Argumente hierzu in Abschnitt 9).

Symptomatisch für den Irrtum, auf dem das Prinzip der Einfachstruktur beruht, ist die Beobachtung, dass in der bisherigen faktorenanalytischen Praxis Einfachstrukturen so gut wie nie gefunden wurden.

Alles bisherige Ringen um Einfachstruktur war eine Don Quichoterie. Schon der Klassiker Harry Harman musste zugeben: „*An orthogonal uni-factor solution is practically impossible with empirical data, and not very likely even when the factors are permitted to be oblique. Nonetheless, ... it is towards that end that the simple structure principles are proposed for the multiple factor solution*“ (Harman, 1968, p. 99).<sup>25</sup>

Das Ziel der Faktorenanalyse ist dem der multivariaten Varianzanalyse (MANOVA) ähnlich. Der Hauptunterschied besteht darin, dass der Faktorenanalytiker mit *latenten unabhängigen* Variablen (UVn, Varianzquellen, ‚Faktoren‘) zu tun hat, die den manifesten, beobachteten abhängigen Variablen (AVn, z. B. Testvariablen) zugrunde liegen. Latente Variablen sind zu Beginn einer Untersuchung unbekannt oder ungesichert und sollen aufgedeckt oder gesichert werden. Ein Varianzanalytiker dagegen hat nicht nur mit manifesten abhängigen, sondern auch mit *manifesten unabhängigen* Variablen zu tun, die ihm bekannt sind und die er experimentell manipuliert. Sein Analyseverfahren dringt indessen nicht bis zur Ebene der *latent* wirksamen Variablen der ‚Black Box‘ vor, d. h. nicht bis zu den Prozessen, die durch die Manipulation der manifesten unabhängigen Variablen aktiviert werden und als *proximale* Determinanten die abhängigen Variablen hervorrufen oder verändern.

Ein weiterer Unterschied zwischen Faktorenanalyse und MANOVA ist erwähnenswert: Die MANOVA berücksichtigt mit aller Selbstverständlichkeit auch ‚strukturelle‘ oder ‚relationale‘ Bedingungen der untersuchten AVn, also Komplexität, die als *Interaktion* unter den manifesten UVn hervortritt, wenn eine solche vorliegt. Die Einfachstruktur-orientierte Faktorenanalyse dagegen versucht, auf der latenten Ebene, auf der sie Faktor-Einheiten aufzufinden sucht, Interaktion unter diesen Einheiten auszublenden.

Mit einem vorurteilslosen Rückgriff auf Elementarwissen hätte das Prinzip der Einfachstruktur vorne herein suspekt erscheinen können, denn den operationalisierten Variablen wissenschaftlicher Forschung liegt generell ein Zusammenwirken von Komponenten zugrunde. Atome bestehen aus Kernen und Elektronen, Salz aus Natrium und Chlorid, Gene kommen nur immer in komplexer Vielfalt zur Auswirkung („Polygenie“).<sup>26</sup> Die Wahrnehmung einer einzelnen Farbe ergibt sich aus der Kombination von elektrischen Erregungen, die von drei Zapfenpigmenten produziert werden. Sätze bestehen aus Wörtern, Wörter aus Morphemen, Morpheme aus Phonemen. Jedem Phonem liegt ein Komplex phonetischer Merkmale zugrunde. Auf die Faktorenanalyse angewandt: Man sollte erwarten, dass den manifesten Variablen einer Untersuchung in aller Regel eine Mehrzahl zusammen wirkender latenter Varianzquellen zugrunde liegt.<sup>27</sup>

Komplexität, das Zusammenwirken bedingender Komponenten, verträgt sich durchaus mit Occams Ideal der Einfachheit, ja sie geht aus diesem als Konsequenz unmittelbar hervor. Wenn verfügbare Einheiten in einem System kombiniert werden können und dadurch nützliche neue Leistungen hervorgebracht werden, erübrigt sich die aufwändige Herstellung neuer Einheiten. Die Physiologie der Farbwahrnehmung kommt mit drei sensorischen Empfangsrezeptoren (Pigmente) aus, sie brauchte für die Farbqualitäten lila, ocker, ultramarin usw. keine Extra-Rezeptoren zu entwickeln. Die jüngsten Untersuchungen zur Evolution der Sprache kulminieren in der Erkenntnis, dass selbst die höchsten mentalen Leistungen des Menschen durch eine erfolgreiche Kombinatorik verfügbarer prozessualer Ressourcen zustande kommen. Für neue Funktionen wie die der Sprache bei Hominiden ein von Grund auf neues Entwicklungsprogramm einzuleiten, wäre nach Bates et al. (1992)<sup>28</sup> und Gould et al. (2002) für die Evolution eine Verschwendung von Ressourcen gewesen. Wolf Singer (2003) wird nicht müde, die menschliche Großhirnrinde mit der Fähigkeit zu beschreiben, „ein kombinatorisches Spiel zu spielen“ (p. 84). Die elementaren Strukturen des Gehirns haben „gleiches Format“, sind „erstaunlich monoton aufgebaut“ (p. 44), „die Natur ist sehr konservativ...“ (p. 46). Das Hirngeschehen bringt bei fehlender Zentralisierung „seine Leistungen“ durch „konstruktivistische Bindungen“ (= Kombinatorik) hervor (p. 75). „*[In der] funktionellen Architektur des Gehirns [ist] entscheidend, wer mit wem wie stark und ob erregend oder hemmend in Kontakt tritt*“ (p. 38). Für den Kontakt einer einzelnen Hirnzelle mit anderen ist durch Tausende Dendriten gesorgt. Eine Übertragung von Thurstones Einfachstruktur auf das Hirngeschehen, die Auffassung der neuronalen Einheiten an solitäre Pflöcke, würde dem Prinzip der freien Bindung direkt widersprechen. Das Bindungsphänomen ist für Singer das „Kernproblem der Neurowissenschaften“ (p. 57). Offenbar unterliegt das Geschehen in der gesamten belebten Natur, auch in der unbelebten, einer Kombinatorik von Elementen und Bausteinen. In einer systematischen Abhandlung weist der Biologe Humberto Maturana (1998, p. 158-189) auf die Universalität der ‚strukturellen Determination‘ des Naturgeschehens hin.<sup>29</sup>

Das Problem der Komplexität verdient eine Diskussion auch aus einer umfassenderen Perspektive, entsprechende wissenschaftsphilosophische Vorarbeiten liegen vor. Malcolm Forster (1998) rollt die Problematik der Begriffe „*Parsimony and Simplicity*“ auf und folgert: „*The paucity of parameters is a limited notion ... There are no compelling ideas about why such properties should count in favor of one theory being closer to the truth than another*“. „*Whether a simple curve is preferable to some more complex alternative,*

or the reverse is true, has nothing to do with simplicity and everything to do with predictive accuracy“ (Forster & Sober, 1994, p. 28). Auch gehören hierher Zitate aus Nicholas Reschers Complexity. A Philosophical Overview „... In the development of knowledge – as elsewhere in the domain of human artifice – progress is always a matter of complexification. An inherent impetus towards greater complexity pervades the entire realm of human creative effort“ (Rescher, 1998, p. 58). „There really are no adequate grounds for supposing the „simplicity“ of the world’s make-up. Instead, the so-called „principle of simplicity“ is really a principle of complexity-management“ (p. 61). Ebenso findet man bei Bechtel & Richardson (Discovering complexity, 1993) im Grundsätzlichen positive Resonanz: „It is ultimately the pattern of connections in the system, and not the jobs performed by specific units in the system..., that is critical to the behavioral systems“. Aus wissenschaftstheoretischen Gründen wird das Postulat der Einfachheit auch von Gawronski (2000) relativiert: Einfachheit ist ein „vages Kriterium“. „Handelt es sich um Punktwolken, die durch eine mathematische Funktion beschrieben werden sollen, so mag es noch einen gewissen Konsens über die ‚einfachste‘ Kurve geben. Handelt es sich jedoch um verbal formulierte wissenschaftliche Theorien, können Urteile über die Einfachheit in Abhängigkeit von Wissensstrukturen ... sehr unterschiedlich ausfallen. In diesem Sinne muss die Forderung nach Einfachheit als relativ betrachtet werden“ (p. 10).

## 6.2 Neues Leitmotiv zur Faktoren-Transformation: Komplexstruktur

Aus alledem folgt, dass faktorenanalytische Forschung unter ein Leitmotiv zu stellen ist, das dem der Einfachstruktur Thurstones entgegen gesetzt ist, welches unter „Einfachheit“ die sparsame Ortung von Punkten im Koordinatensystem verstand. Das neue Leitmotiv ist das der Komplexstruktur. Nach ihm ergibt sich Einfachheit von Variablen durch funktionales Zusammenwirken zugrunde liegender Parameter. Das Ziel einer Nachbehandlung initialer Faktoren durch Transformation sollte sein, der Komplexität des latenten Geschehens, das den beobachteten Variablen zugrunde liegt, uneingeschränkte Repräsentanz zu ermöglichen. Initiale Faktoren einer EFA sind so zu transformieren, dass sie das Zusammenvorkommen der gefunden Bausteine (Faktoren) optimal abbilden. Dass dies mathematisch-statistisch möglich ist und sich gegenüber den konventionellen Verfahren empirisch in überraschendem Maße bewährt, wird in einem Folgebeitrag dargestellt werden. Dort wird auch darzulegen sein, dass die initiale Faktorenlösung – für einen Ausnahmefall – schon ohne Faktorenrota-

tion ein Optimum an Komplexität hervorbringt, d. h. dann, wenn dem Datensatz nur zwei Varianzquellen (Faktoren) zugrunde liegen (wie beim Blutdruck).

Was sollte eine Komplex-Transformation der initialen Faktoren bewirken: Sie sollte den Variablen der jeweiligen Analyse, soweit dies die empirischen Gegebenheiten erlauben, Anteile *aller* faktorieller Varianzquellen zuordnen, die für die betreffende Domäne ermittelt werden.<sup>50</sup> Eine Ausrichtung auf Komplexstruktur will reale Komplexität zur Abbildung führen, während die Ausrichtung auf Einfachstruktur nicht-vorhandene Einfachheit zu erzwingen sucht. Thurstone hat den durch Simple Structure hervorgebrachten Faktoren die Attribute ‚simple‘ und ‚structure‘ zu Unrecht zugesprochen. Das Ideal der solitären Variablen (jede Variable begründe sich durch nur einen Faktor oder durch möglichst wenige Faktoren) führt zur Nicht-Einfachheit und Nicht-Sparsamkeit. Dies wird im Folgeartikel empirisch demonstriert. Komplexstruktur als Leitmodell der Faktorentransformation ist keine Absage an Newton’s ‚natura est simplex‘ und ‚Occam’s razor‘. Das Rasiermesser Occams ist lediglich so zu anzusetzen, dass es die Strategie der Sparsamkeit, die die Natur leistet, hervor bringt und sie nicht verschwinden lässt.<sup>51</sup>

Das Komplexstruktur-Modell will nicht als eine grundsätzliche Neuerung verstanden werden, denn es überträgt nur eine in der Wissenschaft ansonsten schon vertretene Auffassung auf ein methodisches Operationsfeld, auf dem sie bisher nicht hat Fuß fassen können. Dem multivariaten und variablen-interaktiven Geschehen wird in der Wissenschaft ansonsten durchaus Rechnung getragen (Konnektionismus, Vernetzung, Systemtheorie etc.). Doch hat das abwegige formale Vereinfachungskonzept der Simple Structure in einer abgeschirmten Enklave der statistischen Methodologie bis heute überdauert und die substanzwissenschaftliche Forschung fehlgeleitet.<sup>52; 53</sup>

## 7 Gründe des vorherrschenden Irrtums

Was hat zur Fehlkonstruktion der *Simple Structure* geführt? Vermutlich haben alltägliche Fehleranfälligkeiten des Denkens dazu beigetragen. Gestaltpsychologen, die sich mit meist adaptiv nützlichen Formbildungstendenzen der Wahrnehmung beschäftigten (‚Gestaltgesetze‘), warnten wiederholt auch vor ihren schädlichen Nebenwirkungen: Eine „Behinderung des Denkens durch Anschauung“ (Witte 1974), „Prägnanz as an obstacle to problem solving“ (Kanizsa, 1975) scheint auch im Falle der Punktwolken im Faktorenräum vorzuliegen, die von den umliegenden Koordinaten nach den ‚Gesetzen der Nähe und Gruppierung‘ unwiderstehlich angezogen werden. Offenbar ist man

der ‚intuitiv zwingenden‘ Wahrnehmungstendenz gefolgt, den Punktwolken einen Halt zu geben, was nach ‚curve fitting‘-Manier mithilfe der frei drehbaren Koordinaten leicht möglich war.<sup>34</sup> Gorsuch (1974, p. 164 f) spricht von einer ‚visual compellingness‘, die Thurstone bei seinen frühen Veranschaulichungen extrahierter Faktoren leiteten, welche zur Festlegung seiner bekannten fünf Kriterien für die Einfachstruktur-Rotation führten.

Eine weitere Ursache für Urteilsfehler, die dem Argument des ‚leichten Interpretierens‘ von Einfachstruktur-transformierten Faktoren zugrunde liegen könnte, beschreibt Arkes (1991): „*Suppose a person adopts a quick and dirty strategy to solve a problem. Because it is so quick, it is easy to execute. This is a benefit. Because it is dirty, it results in more errors than a more meticulous strategy. This is a cost. Although the choice of this strategy results in [more errors] ..., this cost may be outweighed by the time and effort saved*“ (p. 487). Dies nennt Arkes, wohl nicht ganz treffend, einen strategischen Urteilsfehler („strategy-based judgment error“), die Bezeichnung „cost-saving judgment error“ würde besser passen.

Ähnliches findet man bei Edmonds (2002), der den universellen Hang („bias“) zur Einfachheit einer menschlichen Schwäche zuschreibt. Denn „... *the simpler theory is not more likely to be true and is not likely to be nearer the truth ... For human beings it is much easier to elaborate ... [a failing] theory, or otherwise tinker with it, than to undertake a more radical shift (for example, by scrapping the theory and starting again). This elaboration may take on many forms, including ... complicating the model with extra equations or rules ... or using more complicated functions*“. Edmonds verweist auf die vielzitierte historische Analogie (Rettung der geozentrischen Zyklen der Planeten durch den Zusatz kleiner Korrektur-Zyklen – „Epizyklen“) und fasst zusammen, „... *Model selection, for the sake of simplicity‘ is either: simply laziness ... [or] due to pragmatic reasons*“. Er plädiert dafür, auf „Einfachheit aus Furcht vor Neuheit“ grundsätzlich zu verzichten. „... *The elaboration of [an existing] theory in order to fit a known set of data should be resisted ... The lack of success of a theory should lead to a more thorough and deeper analysis than we are usually inclined to perform*“.

Zur Immunität des Prinzips der Simple Structure hat zweifellos auch die Macht der Worte beigetragen. Gegen einen Begriff, der sich mit den attraktiven Attributen ‚einfach‘ und ‚Struktur‘ sprachlich Geltung verschafft, wird kaum jemand opponieren wollen. Hätte Thurstone sein Rotationsprinzip z.B. *Principle of Minimal Factor Contribution* genannt – das wäre präziser und bescheidener gewesen –, dann hätte man dessen Berechtigung wahrscheinlich bald angezweifelt. Heute, nach weltweiter Verbreitung des Zau-

berworts ‚Simple structure‘ (*Structure simple, Structura simple, Struttura semplice, Struktura prosta, Eenvoudige structuur, Sakhtare sade, Egyszerű szerkezet, Enorazsežna struktura, Basit yapı*), wird man sich von ihm nicht gern trennen wollen.

Zum Fortbestand des harten Kerns der Einfachstruktur haben nicht zuletzt führende Methodiker des Faches beigetragen. Dem Bedarf nach mehr Komplexität bei der faktoriellen Modellierung von Untersuchungsgegenständen begegneten sie mit erhöhter Komplizierung ihrer von Anfang an unpassenden Verfahren. Die zurückhaltende Attitüde eines exploratorisch arbeitenden Forschers, der sich für überraschende Befunde offen hält, ist zunehmend der Attitüde eines Variablen-Administrators gewichen, der nach Gutdünken die Einheiten seiner Domäne festlegt und ihnen selbst-ausgedachte Funktionen zuordnet.

Die derzeitige epistemische Lage im Fach lässt sich auch von einer höheren Warte aus betrachten. Für die Naturwissenschaft generell steht *Machbarkeit* an der Spitze der Wertrangordnung. Den Modellen, die man herstellen, manipulieren und der Natur auferlegen kann, wird mehr Respekt entgegen gebracht als den Modellen, die der Natur abgewonnen werden. Die Kompliziertheit der Produkte erfolgsstolzer Modellbastler verstellt leicht den Blick dafür, dass sich Entdeckungen von größerer Tragweite in der Regel erst nach ausgiebiger Naturbetrachtung einstellen, bei der mit rezeptiver Zurückhaltung Informationen entgegen genommen werden.

Das Prinzip der Einfachstruktur kann man mit Lakatos (1977) als Beispiel eines harten Kerns (*hard core*) auffassen, der die Entwicklung konventioneller Forschungsprogramme längerfristig begleitet. Eine Grundüberzeugung wird zu einem ‚hard core‘, wenn eine Forscher-Community sie für selbstverständlich hält und nicht infrage stellt.<sup>35</sup> Für Michell (2000) ist das blinde Festhalten an fundamentalen Voraussetzungen Symptom einer „wissenschaftlichen Pathologie“: „*A hypothesis is accepted without a serious attempt being made to test it and this failure of critical inquiry is ignored*“ (p. 648). Ein solches Symptom beschrieb Schönemann, indem er die Empörung der ‚Traditionisten‘ erwähnte anlässlich der Versuche Bargmanns, das Vorliegen von *Simple Structure* statistisch zu überprüfen: „*[They] scorned it [Bargmann’s Test] as a sacreilege of their cherished belief that simple structure is a law of nature*“ (Schönemann, 1994, p. 293). Lange vor Lakatos hat Fleck (1935) die wissenssoziologischen Auswüchse der ‚Denkkollektive‘ mit Falldarstellungen beschrieben. „*Ist ein ausgebautes, geschlossenes Meinungssystem ... einmal geformt, so beharrt es beständig gegenüber allem Widersprechenden ... Was in das System nicht hineinpasst, bleibt ungesehen oder es wird verschwiegen ... oder es wird mittels großer Kraftan-*

strengung dem Systeme nicht widersprechend erklärt“ (p. 35). „Die Beharrungstendenz der Meinungssysteme beweist uns, dass sie ... als stilvolle Gebilde zu betrachten sind. Sie ... zeigen als harmonische Ganzheiten besondere Stilmerkmale, die jede einzelne Erkenntnisfunktion bestimmen ... [Sie bringen eine ‚Harmonie der Täuschungen‘ hervor], die dann im Bereiche eines bestimmten Denkstils auf keine Weise aufzulösen ist“ (p. 45).

Hard core-Voraussetzungen in wissenschaftlichen Forschungsprogrammen erzeugen nach Lakatos auf Dauer empirische ‚Anomalien‘. Mit wachsender Zahl von Anomalien geraten die Programme in eine ‚degenerative‘ Entwicklungsphase. Zur Sicherung des harten Kerns werden dann ‚Schutzgürtel‘ erstellt und bei Bedrohung verstärkt. Die hochkomplizierten Methoden zur Verarbeitung multivariater Daten, die einseitig mathematisch ausgerichteten Methodiker seit etwa zwei Jahrzehnten entwickeln, lassen sich als Schutzgürtel verstehen, die das Dogma Simple Structure immunisieren (mehr dazu in Abschnitt 9).

## 8 Unbeachtete Stimmen der Kritik

### 8.1 Hat man an *Simple Structure* nicht schon früher Anstoß genommen?

Am Prinzip der Einfachstruktur wurden, was dessen Leistung betrifft, mitunter Zweifel laut, doch seine Berechtigung selbst wurde so gut wie nie radikal hinterfragt. W. Stephenson (1956) gehört zu den frühesten Zweiflern. Er erkannte, dass der Versuch, Persönlichkeitseigenschaften auf einzelne Faktoren zu reduzieren, auf den Widerstand empirischer Komplexität stößt: „The problem is to explain complex traits in terms of relatively few primaries“ (p. 7). Mit liberaleren Rotationsmethoden wollte er den von ihm zu Recht so genannten ‚compound traits‘ einen Freiraum der Repräsentation verschaffen. Doch sein Verfahren war unhandlich, ähnlich wie die Circumplex-Verfahren von Hofstee et al. (1992) und de Raad et al. (1994) (s. Abschnitt 9) mit ihren verwirrenden Ergebnissen, die ebenso wenig wie bei Stephenson vom Rotationsideal der Einfachstruktur abließen. Stephenson favorisierte zeitlebens die faktorielle Q-Technik der Faktorenanalyse, welche ‚Personvariablen‘ zur Analyse führt, die mit dem von ihm eigens entwickelten Q-Sort-Erhebungsverfahren gewonnen wurden. Die Personvariablen werden interkorreliert und faktorisiert, während sonst fast ausschließlich die faktorielle R-Technik eingesetzt wird, die mit *Verhaltensvariablen* operiert (z. B. Testwerte, Urteilstwerte). Stephenson umgeht damit die ungelöste Frage: Welches sind die basalen Merkmale, mit denen die Komplexität des Erlebens und Verhaltens und zugrunde liegende individuelle Dispositionen („com-

pound traits“) theoretisch verständlich wären. Zur Klärung dieser Frage ist die exploratorische R-Technik unverzichtbar, denn Ergebnisse der Q-Technik müssen durch Konstrukte interpretiert werden, die diese Technik selbst nicht zur Verfügung stellt.<sup>56</sup>

Hochgradig unzufrieden mit der Einfachstruktur Thurstones ist auch Guilford (1974, pp. 498 f). „Thurstone’s principle of simple structure ... is by no means sufficient if we want logical psychological meaning ... In numerous instances [in Thurstone’s and his students’ studies] tests of very different character are often thrown together significantly on the same factors. This does not make good psychological sense ... Rarely were varimax factors clean-cut and easy to interpret psychologically ...“. Guilfords Strukturmodell der Intelligenz war eine Reaktion auf das von ihm richtig erkannte Ungenügen der Ergebnisse, das die Anwendung von Simple Structure zur Folge hat. Er sah sich zu einer zusätzlichen Ordnungsleistung herausgefordert, die die Faktorenanalyse selbst nicht liefere. Doch setzte er dabei erst am Ergebnis der Transformation zur Einfachstruktur an, den inhärenten Fehler des Thurstone-Prinzips erkannte er nicht. Er suchte ihn durch von außen herangetragene Kategorien zu korrigieren.

Ein jüngerer Kritiker, der ähnliche Argumente vorbrachte, ist Allen Yates – Stephenson und Guilford bleiben von ihm indessen unerwähnt. Wie kein anderer vor und nach ihm brachte er die Probleme des Simple Structure-Prinzips auf den Punkt: „The factors that result from cluster-oriented factor analysis [from blind application of simple structure] are simply an index of success an investigator has had in putting together groups of collinear variables (clusters) – regardless of how complex these same variables might be in terms of their latent determinants. In other words, manifest collinearity among variables is an indication only that they share the same pattern of causal determination; it does not in any way suggest that the shared pattern involves just one latent causal factor“ (Yates, 1987, p. 39). Die Monographie von Yates enthält den konzertiertesten Angriff gegen die ‚Ideologie‘ der Simple Structure, den ich ausfindig machen konnte. Sein letztes Kapitel beginnt Yates mit: „Only a radical reorientation of current perspectives will allow researchers to apply exploratory factor analysis in the manner envisioned by its originators as a powerful technique for routine discovery of the underlying bases of observed covariation“ (Yates, 1987, p. 325). Aber auch Yates rührt das Simple Structure-Prinzip nicht an, er meint nur, man habe Thurstones ursprünglich liberalere Auffassung ignoriert (ignoriert hat sie der spätere Thurstone allerdings auch selbst). Zur Analyse komplexer Datensätze führt Yates lediglich neue Rotationsalgorithmen ein. Seine Alternativen (*Direct Geomin* und *Direct Geoplane*) wurden wie die Verfahren anderer Autoren, die den unwillkommenen

Begleiterscheinungen von Simple Structure entkommen wollten, überaus kompliziert, das Gelingen einer Faktorenanalyse zunehmend von subjektiven Entscheidungen des Forschers abhängig gemacht. Yates' Innovation sollte ein Ausweg aus dem Dilemma werden. Sie blieb von der Fachwelt unbeachtet.

Die Problematik der Simple Structure erkannte auch Rozeboom (1991), der sie für ausweglos hielt: „*For diagnosing the causal grain of common-factor space, rotation to simple structure is so disquietingly fallible that we would surely prefer another criterion were any plausible alternative at hand. (Churchill's aphorism on the inferiority of democracy comes to mind here, namely that democracy is the worst form of government there is – except for all the others ... Read ‚simple structure‘ for ‚democracy‘ and ‚rotation criterion‘ for ‚form of government‘)*“ (p. 587). Rozeboom's HYBALL Rotation, bei der mehrere Koordinaten (Unterräume) anstelle von je zwei Koordinaten sukzessive rotiert werden („somewhat more holistic than a simple sequence of planar rotations“ p. 587) erweist sich als ein ähnlich fragwürdiger Kompromiss wie der von Yates.

Zu den wenigen Kritikern, die Einfachstruktur grundsätzlicher infrage stellen, gehören auch Schönemann & Borg (1996), die resigniert konstatieren, „*dass das Einfachstruktur-Kriterium (formuliert als Varimax-Kriterium) in der faktorenanalytischen Praxis routinemäßig angewendet [wird]*“. Sie ergänzen mit kritischerem Blick: „*Auf der Strecke bleibt dabei, die wichtige Frage, warum man Einfachstruktur in vielen Systemen von Variablen überhaupt erwarten sollte*“ (Steiger, 1994, p. 204). Im Grunde sei „die Einfachstrukturhypothese wenig plausibel ...“. Sie behauptete, Testaufgaben mit Nicht-Null-Ladungen auf allen Faktoren könne es „unmöglich“ geben, damit „spannt [sie] den Wagen vor das Pferd“ (nach Guttman, 1992, p. 186).

Den Typ von Irrtum, den sich die Proponenten der Einfachstruktur leisten, hat Gigerenzer (1978) treffend diagnostiziert – allerdings an einem anderen Beispiel dimensionsanalytischer Modellbildung. Unkritische Modell-Anwender in der Psychologie ignorieren, dass ihre Modelle mit dem psychologischen Gegenstandsbereich durch inhaltliche Vorannahmen verknüpft sind, die selbst nicht unmittelbar geprüft werden. Ignoriert wird dabei das, was Gigerenzer eine „Implikationsthese“ nennt: „*[Diese These] beinhaltet, dass jedes mathematische System (z. B. eine dimensionsanalytische Methode) bereits durch seine Anwendung auf einen psychologischen Gegenstandsbereich G eine psychologische Theorie über diesen Gegenstandsbereich impliziert.*“ (p. 110).<sup>37</sup> Das formale Relationssystem (FRS) ist mit dem „empirischen Relationssystem“ (ERS) untrennbar verschränkt (Gigerenzer verweist hier auf Unterstützung durch Suppes & Zinnes, 1965). Fallen die Systeme FRS und ERS in entscheidenden

Punkten auseinander, entstehen „Divergenzartefakte“ mit „theoretisch wertlosen Ergebnissen“ (p. 111).

Auf den vorliegenden Kontext angewandt: Das FRS der Einfachstruktur setzt bei seiner Anwendung ungeprüft voraus, dass den auf der ERS-Ebene zugänglichen Variablen solitäre Varianzquellen zugrunde liegen. Gigerenzer fordert: „*Um ein Untersuchungsergebnis gegen eine Interpretation als Divergenz-Artefakt abzusichern, muss der Forscher ... das durch das mathematische Modell implizierte gegenstandsbezogene psychologische Modell explizieren ...*“ (p. 116). Genau dies haben Faktorenanalytiker bislang versäumt. Sie glauben, ihre mit mathematischer Präzision auf empirische Daten angesetzten Analysen würden ihnen die Struktur der psychologischen Realität automatisch ausliefern. Zusätzliche und unabhängige Denkmittel einzusetzen, wird für überflüssig und sogar für anrühlich gehalten, da sie mit dem Makel des ‚Subjektiven‘ behaftet werden.

## 9 Führen nichtfaktorielle Verfahren weiter?

### 9.1 Was leisten Circumplex-Verfahren?

Mit vielerlei Maßnahmen haben Methodiker auf die unbefriedigenden EFA-Resultate reagiert, mathematische Überbau-Gebilde sollten die Symptome kurieren (Revenstorff, 1978, p. 12). Den Mehrfachladungen bei Variablen (z. B. Fragebogen-Items), die Abweichungen vom Simple Structure-Ideal darstellen und in der faktoriellen Persönlichkeitsforschung ein alltägliches Ärgernis sind, versuchte man durch aufwändige Zusatzberechnungen gerecht zu werden. Dafür gaben Hofstee et al. (1992) mit ihrem AB5C-Modell (Abridged Big Five Dimensional Circumplex) ein Beispiel ab, das Simple Structure-Resultat adjektivischer Person-Beschreibungen suchten sie mit der Circumplex-Methode von Wiggins (1979) aufzubessern. Die Circumplex-Methode lässt das Zusammenkommen von jeweils zwei faktoriellen Traits beim gleichen Item zur Abbildung zu, es entlastet somit von den Restriktionen der Einfachstruktur partiell.

Dies ist aber dies ein hybrider Kompromiss, das Verfahren „*does not deal adequately with those ... variables that load highly on more than two ... factors*“. Circumplex macht also nur Teil-Zugeständnisse an die Komplexität der Realität und ist im Ergebnis weder „definitive nor comprehensive“ (Hofstee et al., p. 161). Am Simple Structure-Prinzip selbst wird nicht grundsätzlich gerüttelt, das Dimensionsproblem bleibt unangetastet. Die Autoren trauen sich nur zu „*[to] propose a partial liberalization of simple structure*“ (p. 147), wo doch erst eine Abschaffung der Simple Structure als Prinzip das Komplexitätsproblem lösen würde.

Dass Circumplex-Methoden versagen, wenn „all items are best described by more than two factors“, stellen auch Acton & Revelle (2004, p. 26) fest, die die psychometrischen Kriterien für das Vorgehen mit Circumplex explizieren. Die Voraussetzungen sind im übrigen sehr anspruchsvoll – die Autoren führen davon zehn auf –, so dass Circumplex-Verfahren selbst dann kaum zu empfehlen sind, wenn nur zweifaktoriell beschreibbare Datensätze zur Analyse anstehen.

## 9.2 Sind konfirmatorische Verfahren eine Alternative?

Die konfirmatorischen Verfahren (CFA), die durch Strukturgleichungsmodellierung (SEM) auf ein flexibles Niveau gebracht wurden, scheinen mit ihren raffinierteren Algorithmen die exploratorischen Verfahren weitgehend verdrängt zu haben. Anstatt als EFA-Forscher sich mit latenten Variablen zu begnügen, von denen man sich überraschen lässt, ist es Mode geworden, diese nach Gutdünken zu erfinden, um sich anschließend zur empirischen Realität mit konfirmatorischer Technik ins Verhältnis zu setzen. Zu diesem Zweck werden, meist mit Trial-and-error-Blindheit, Massen von Passungsversuchen unternommen, der notwendige Aufwand und die Rechenzeit sind minimal.

Doch orientieren sich Strukturgleichungsmodelle ebenfalls durchweg am Prinzip von Simple Structure, somit bleibt die Komplexität latenter Varianzquellen auf diesem Wege verdeckt.<sup>58</sup> „Die neueren Strukturgleichungsmodelle – zunächst, wie gehabt, euphorisch gefeiert ... – tragen die Probleme der Faktorenanalyse in verschärfter Form fort“ (Schönemann & Borg, 1996, p. 241). Fast immer verfehlen sie die zugrunde liegenden Strukturen.<sup>59</sup>

Die Ergebnisbilanz der SEM-Forschung fällt entsprechend mager aus („poor results“, Beauducel & Wittmann, 2005, p. 42). Man ist dabei, sich mit den unzureichenden Ergebnissen auch dieser Forschung abzufinden. Es bleibt bei Feststellungen wie „the assumption of simple structure is probably a typical ... simplification bias“, aber leider ‚necessary‘ (Beauducel & Wittmann, 2005, p. 43). „... Simple structure models of personality are unlikely to meet conventional or even fairly relaxed goodness-of-fit criteria ... Overemphasis on simple structure ... may explain some of these problems“ (p. 44). In einem kritischen Review-Artikel wiesen Mac Callum & Austin (2000) auf die Probleme der konfirmatorischen Verfahren und auf den ‚confirmation bias‘ ihrer Anwender hin, die sich mit Goodness-of-fit-Werten und ihren willkürlich gewählten Kriterien begnügen, welche nur scheinbare Passungen herbeiführen. Nach Ansicht der Autoren sollten

Analyse-Ergebnisse nach sachorientierten und nicht primär nach formalen Maßstäben bewertet werden.<sup>40</sup> Wenn die überprüften Modelle das Fit-Kriterium nicht erreichen, werde dies oft leichtfertig ignoriert („working with imperfect models“, Titel eines Artikels, MacCallum, 2003). Oder es werde nicht bedacht, dass die Passung des selektierten geprüften Modells von anderen nicht-geprüften Modellen aus einer unabsehbar großen Modellmenge leicht übertroffen werden könnte. Schon Kaiser hat von der CFA nicht viel gehalten: „I cannot resist saying that, for me at least, the earlier exploratory thrashing about was much more fun – and perhaps even represented more progress – than the forth-coming confirmatory prettying-up“ (Kaiser, 1970, p. 406).

Ernüchternde Ergebnisse aus weiteren Vergleichen zwischen CFA und konventioneller EFA findet man bei Church & Burke (1994) und Ferrando & Lorenzo-Seva (2000). Die Kritik von Cliff (1985) liegt auf der gleichen Linie. Eine jüngste umfangreiche CFA-Studie mit Simulationsdaten kommt zur Schlussfolgerung: „... trait models [of personality] assuming simple structure tend to be rejected with CFA ...“. „... there will always be some small distortion of simple structure“ (Beauducel & Wittmann, 2005, p. 72). Mit Durchsetzung des Simple Structure Modells bei Datensätzen, insbesondere bei Persönlichkeitsdaten, die die CFA ermöglicht, wird eine Kluft („gap“) geschaffen „between the large body of results based on exploratory factor analysis and CFA in personality psychology“ (p. 73). Die exploratorische Analyse der konventionellen Art, die hier gemeint ist, weist trotz Ausrichtung auf die Einfachstruktur Sekundärladungen bei vielen Variablen auf, die beim blinden Vorweg-Verteilen von Faktorgewichten im konfirmatorischen Prozedere nicht erahnt werden („... there is no [prior] knowledge of secondary loadings“, p. 43). LISREL-Ergebnisse hält Vittadini (1989) in jedem Fall für indetermiert, weil die konstruierten latenten Modellvariablen von manifesten Variablen abhängig gemacht werden: „one may actually be confirming the model because the manifest variables are determined by other variables than those hypothesized, which happen to have the same pattern of relationship to the manifest variables as given by one’s hypothesis ... One can never regard structural hypothesis as true as opposed to ‚confirmed‘“ (p. 428).

Konfirmatorische Verfahren können sich ggf. dann als nützlich erweisen, wenn die latenten Variablen in der jeweiligen Domäne durch exploratorische Analysen gesichert wurden. Velicer & Jacksons (1990) Argumentation geht in diese Richtung: „Exploratory analytic approaches ... should be preferred except for those cases where a well-defined theory exists. Exploratory approaches avoid a confirmation bias, do not force a theory-oriented approach prematurely, and represent

*a conservative strategy*“ (p. 21). Ob zur Überprüfung eines exploratorisch gewonnenen und theoretisch befriedigenden Ergebnisses eine konfirmatorische Überprüfung überhaupt noch sinnvoll einzusetzen ist – es käme doch lediglich einer Replikation der exploratorischen Analyse gleich –, bliebe dann eine wahrscheinlich zu verneinende Frage.<sup>41</sup>

### 9.5 Wie sind mathematische Erkenntnismittel zu nutzen?

Überdacht werden sollte auch generell, ob die im gegenwärtigen Wissenschaftsbetrieb verbreitete Neigung zu begrüßen ist, den mathematischen Techniken zunehmend zu vertrauen und ihren Schöpfern die Verantwortung für ihre Anwendung abzugeben. In den Schriften selbstkritischer Methodologen finden sich Warnungen: *„Those who firmly believe that rigorous science must consist largely of mathematics and statistics have something to unlearn. Such a belief implies the emasculation of the basic substantive nature of science. Mathematics is content-less, and hence not – in itself – empirical science ... rigorous treatment of content or subject matter is needed before some mathematics can be thought of as a possibly useful (but limited) partner of empirical science“* (Guttman, 1971, p. 42).<sup>42</sup> Mit ähnlicher Stoßrichtung wird der Skeptiker Guttman auch von Schönemann zitiert: *„There remains the danger of seeking data merely to fit axioms“*, dazu der Kommentar: *„In hindsight, these warnings sound positively prophetic in anticipating the present malaise in mathematical psychology some 20 years before Cliff (1992) noticed it ...“*. Schönemann (1994, p. 294) nennt noch Narens & Luce (1995) als Kritiker der Malaise. Ein Pionier der Mathematischen Psychologie, William K. Estes, beklagte schon 1975 die Situation seiner Spezialisierung: *„... it is clear that many investigators in our field are not entirely happy with their current situation“* (p. 265). Er bedauert die Kluft zwischen der mathematischen und inhaltlich-theoretisch ausgerichteten Psychologie und zitiert, um eigene Revisionsvorschläge zu unterstreichen, Leont'ev & Dzafarov, 1973, p. 20): *„An analysis of the present situation shows that contemporary psychology and contemporary mathematical instruments are still not compatible enough with one another to allow mathematization to assume a central place in the development of psychological knowledge; the reason for this is not only the low level of sophistication of the latter ... What is required is a continual interaction between mathematics and psychology, an interaction that ... would lead to a revision of existing mathematical methods into forms more amenable to the proposed mathematized conceptual systems“*.

Psychologische Inhaltsebenen werden dem Forscher zugänglich nicht primär durch formale Modelle, sondern durch die Gesamtheit seiner Erfahrungen, die zur jeweiligen Domäne gehören. Das durch frühere Erfahrungen gewonnene Wissen und Vermuten wird durch Forschung präzisiert und geprüft, alltagsprachlich kommunizierbare Repräsentationen spielen anfangs dabei eine große Rolle. Vor allem an diesen sollte, wenn möglich, die Passung der formalen Modellierungen psychologischer Sachverhalte bemessen werden. Dazu haben sich kritische Autoren geäußert:

- *„... Neither algorithmic sophistication, nor axiomatic rigor alone are apt to advance our knowledge much if they are cultivated in an empirical vacuum“*. (Schönemann, 1981, p. 412)
- *„May I ... insist once again on the absurdity of divorcing the mathematical or statistical evidence from evidence procured by other means? ... The sole claim of mathematical analysis should be to verify, by appropriate calculation, the hypotheses commonly advanced on the basis of much broader and more general lines of evidence“*. (Burt, 1949, p. 107)
- *„Let us try to be free of ... a priori mathematical and statistical considerations and prescriptions – especially codes of permission. Instead, let us try to think substantively ... and focus directly on the specific universe of observations with which we wish to do our business“*. (Guttman, L., 1971, p. 346)
- *„Factorists with more mathematical training than the rest of us have been addressing themselves to problems ... on a technical rather than upon a fundamental level ... In most cases [their models are] irrelevant ... to the problems ... of those for whom factor analysis is a research tool ...“*. (Butler, J. M., 1969, 252-3)
- *„These techniques [„for rotating factors into ‚psychologically meaningful‘ positions“] ... became the stock in trade of practicing factor analysts. ... It is probable that future historians will be severely critical of them and of their users; critical of the techniques because of ... their users for extravagant claims on their behalf“*. (Maxwell, 1959, p. 228)

Angesichts der zunehmend formalistischen Ausrichtung der psychologischen Forschung wären die nicht-mathematischen Erkenntnismittel, die als ‚hermeneutische‘ und ‚phänomenologische‘ mit negativem Vorzeichen abgedrängt werden, kompensatorisch besonders zu pflegen und zu fördern. Wenn sie auf einem tiefer liegenden Niveau operieren („on a fundamental level“, Butler) und zukünftig zur kritischen Rückschau auf die „extravagant claims“ der Formalisten benötigt

werden (Maxwell), müsste man sie heute mit Nachdruck am Leben erhalten. Innerhalb der Gesamtheit unserer Erkenntnismittel erfüllen die mathematischen nur einen, wenn auch überaus wichtigen Teil unserer Wünsche.<sup>45</sup> Präzision im Denken und im Detail und die zu diesem Zweck eingesetzte ingenieure Mathematik sind nützlich, aber nicht für jede menschliche Geistesbewegung unentbehrlich. Sie können schädlich sein, wenn sie im Übermaß eingesetzt werden – das zeigt das Debakel mit Simple Structure –, während man die basaleren, umfassenderen, holistischer ausgerichteten, wenn auch zeitweise weniger scharfen Erkenntnismittel, vielleicht ungewollt, aber doch systematisch ins Abseits drängt.<sup>44</sup> Für ein holistisches Vorgehen in der Forschung, das die methodisch herausgeschnittenen Beobachtungen auf Vereinbarkeit mit dem immer vorhandene Hintergrundwissen prüft, plädiert auf wissenschaftstheoretischer detaillierter Basis auch Gawronski (2000).

## 10 Ausblick

Der vorliegende Artikel ist kein Angriff gegen die Faktorenanalyse überhaupt, sondern nur gegen einen Stil ihrer Anwendung, die den gesunden Menschenverstand ausklammert und die Voraussetzungen ignoriert, die ihre Anwendung impliziert. Gigerenzer & Strube (1978) propagierten eine „annahmenkritische“ Anwendung der verfügbaren Verfahren. Hätte man die von diesen Autoren empfohlene Überprüfung der Annahmen vorgenommen, auf welche sich Faktorenanalysen stillschweigend stützen, hätte schon früher das erreicht werden können, was sich Yates erhofft hatte. Dieser Autor hatte einen Paradigmenwechsel für notwendig gehalten,<sup>45</sup> die ‚Pathologie‘ der faktoranalytischen Forschung wäre sonst nicht zu beheben und dem ‚Morast‘ ihrer bisherigen Ergebnisse nicht zu entkommen.

Zunächst wäre die ‚Ritualhandlung‘ des ‚Little Jiffy‘ zu beenden, die von Gigerenzer & Strube (p. 81) angeprangert wird (*Little Jiffy* = rezeptartige Anwendung des Varimax-Verfahrens). Der Geist der Kritik wäre sodann wieder aufzunehmen, der im deutschsprachigen Raum mit Kallina (1967) und den übrigen Münchener

Symposiasten, mit Kalveram (1970), Gigerenzer & Strube (1978) und Revenstorf (1980) auflebte, dann aber versiegte. Das Komplexstruktur-Paradigma braucht zu seiner Bestätigung ebenso wie zu seiner Ablehnung – falls sich eine solche mit hier vielleicht übersehenen Gegengründen durchsetzen möchte – ein aktives Engagement.<sup>46 47</sup> Empfohlen wird sodann der umfassende Einsatz des hier und im Folgeartikel vorgeschlagenen alternativen Verfahrens. Strube (2000, p. 117) übte zurecht an denjenigen Kritik, die bislang mit dem „senselessly adhering to the usual methods ...“ nur unzufrieden waren, ohne dem etwas entgegen zu setzen: „*Most of [existing work] discuss negative examples without giving positive ones*“ (p. 117). Mit Anwendungen der Faktorenanalyse, die sich vom Komplexstruktur-Modell leiten lässt, sollte eine Befreiung vom Dogma Thurstones und – vielleicht nach weiteren notwendigen Verfeinerungen der Methode – eine optimale Abbildung der Struktur der latenten Varianzquellen der manifesten Observablen erreicht werden.

Henry F. Kaiser, der für das Little Jiffy die Voraussetzungen schuf, wäre wohl zu einer radikal neuen Ausrichtung der Faktorentransformation bereit gewesen – er ist zu früh verstorben (1927-1992). Denn er hatte sich am Ende eines Vortrags über die „Second generation Little Jiffy“ vor der Psychometric Society für solche Fälle ausdrücklich angeboten: „*For the future, I can assure you of one thing: if any of you folk ... come up with some Big Breakthroughs I shall be waiting in the wings ready and eager to paste them together to produce the next generation Little Jiffy*“ (Kaiser, 1970, p. 415). Gene Glass (1991) charakterisiert Kaiser in einem Nachruf als ‚respektlos‘, eine Eigenschaft, die Kaiser offenbar nicht übel anstand, denn, so Glass weiter, „*irreverence must be a necessary ingredient in the recipe for creativity. Whoever worships received wisdom too ardently will never see beyond it*“. Das komplexitätsmodellierende Verfahren könnte ‚respektlos‘ erscheinen, da es Kaisers Varimax in sein Gegenteil verkehrt. Doch hätte Kaiser diese Unbotmäßigkeit wohl begrüßt, so wie er sie selbst praktizierte und von den Studenten seiner – und sicher auch der nächsten Generation – erwartete.

## Anmerkungen

- 1 Schönemann (1981) und Steiger (Steiger & Schönemann, 1975) gehören nach Guttman (1955) zur mittleren Generation der Methodiker, die an der Faktorenanalyse Kritik übten, ihre Kritik war scharf („theoretical problems“, „users are generally uninformed about the defects of this model“, p. 175, 188). Doch lag ihr Akzent nicht auf *Simple Structure*, sondern auf der formalen Unentscheidbarkeit über die Geltung faktorieller Dimensionen (indeterminacy) und deren Nicht-Identifizierbarkeit (lack of identifiability). Schönemann & Steiger (1976) entwickelten nach Reanalysen von dreizehn publizierten faktorenanalytischen Studien, die zur vernichtenden Kritik an diesen führte (Schönemann & Wang, 1972), ein alternatives Verfahren der multivariaten Datenreduktion (*Regression Component Decomposition, RCD*), das im Vergleich zur Faktorenanalyse größere begriffliche Klarheit und rechnerische Effizienz und die Möglichkeit der Modell-Falsifizierung versprach. Doch blieb dieser Ansatz angesichts des Siegeszugs der ‚multiplen Faktorenanalyse‘ Thurstones unbeachtet. Zur Frage der Transformation der durch RCD ermittelten ‚Komponenten‘ bot der alternative Ansatz auch nichts Neues. Zudem waren die Ergebnisse der RCD von denen der Faktorenanalyse Thurstones offenbar nicht sehr verschieden.
- 2 Übersichtsreferate über die exploratorisch-faktorenanalytische Forschungsmethodik mit Anwendungsorientierung aus jüngerer Zeit – fast alle ohne grundlegende Kritik – stammen von Moosbrugger & Hartig (2002) und Fabrigar et al. (1999), denen Artikel mit ähnlichem Ziel vorausgingen: Stevenson (1995), Tinsley & Tinsley (1987), Ford et al. (1986), Glass & Taylor (1966), Cattell (1965), Peel (1955). Die am häufigsten zitierten englischsprachigen Lehrbücher zur exploratorischen Faktorenanalyse sind – mit z. T. Angabe der letzten und ersten Auflage: Comrey & Lee (1992, 1973), Child (1990,1971), Gorsuch (1985, 1974), Harman (1976, 1968), Mulaik (1972). Deutschsprachige Äquivalente sind Weber (1978), Revenstorf (1976), Überla (1977, 1971). Eine deutschsprachige Einführung als Lehrbuch-Anhang bietet Lienert (1969) und als Lehrbuch-Kapitel (Nr. 15) Bortz (2004, 1977).
- 3 Mit dem Begriff ‚Dimension‘ wird der Anspruch einer nicht weiter hinterfragten Metrik angemeldet. So wie die drei Dimensionen des euklidischen Raumes der Ortung von Objekten dienen, so meint man, die primären Variablen psychologischer Beobachtung mithilfe von faktoriell gewonnenen ‚Dimensionen‘ verorten zu können. Dabei wird die Metapher der Raumdimensionen überstrapaziert (z.B. auch durch den Ausdruck ‚Semantic Space‘ von C. E. Osgood). Der Terminus Dimension, im mathematischen Sprachgebrauch berechtigt, ist zur Bezeichnung von Varianzquellen auf der Objektebene irreführend und entbehrlich.
- 4 Die Gesamtheit typologischer Konstruktionen der differentiellen Psychologie bis 1970, nicht nur faktorenanalytischer Provenienz, hat Ruttkowski (1974) zu erfassen versucht. Dazu kommentiert Sponzel (1998): „Weltweit gibt es über 1000 Persönlichkeits-, Charakter- oder Wesenstypologien (Ruttkowski, 1974). Die Mehrzahl davon dürfte ... umstritten sein. Viele sind in der Kultur- und Wissenschaftsgeschichte untergegangen. Viele überschneiden sich. Es hat den Anschein, als ob beliebig viele Konstruktionen möglich sind – je nach unterschiedlichen Zielen und Zwecken“. Ähnlich urteilen Gigerenzer & Strube (1987, p. 85): „Es ist die *Crux* der faktorenanalytischen Forschung, so viele ‚anerkannte‘ Persönlichkeitsfaktoren aufgefunden zu haben, dass selbst bei einfacher Dichotomisierung der Dimensionen die Zahl der dadurch zur Verfügung stehenden Quadranten höherer Ordnung mit schätzungsweise  $2^{50}$  ... etwa vierhunderttausendmal so groß ist wie die derzeitige Weltbevölkerung“.
- 5 Einen Eindruck vom Elan der Big-Five-Bewegung verschafft eine umfassende Sammlung von Zitaten von C. R. Brand. Sie ist bildschirmzugänglich über die URL <http://www.cycad.com/cgi-bin/Brand/quotes/q05.html>
- 6 Blinkhorn (1997, p. 180) kritisiert überzogene Hoffnungen schon bei den Pionieren der Faktorenanalyse: „*The words they use, for example ‚primary mental abilities‘ (Thurstone) or ‚source traits‘ (R. B. Cattell), are witness to the faith and trust placed in factor analysis as revealing the psychological analogue of the periodic table of elements, or the list of subatomic particles*“.
- 7 Ozer & Reise 1994 „*characterized the big five as the ‚latitude and longitude‘ along which any new personality construct should be routinely mapped*“ (Funder, 2001, p. 200).
- 8 Briggs (1989): „... *a coherent and falsifiable explanation for the five factors has yet to be put forward*“. „*There is no theoretical reason why it should be these five rather than some other five*“ (p. 249). „*The structure of trait attributions may not correspond straightforwardly to the deep structure or neurophysiological basis of human tendencies*“. (p. 250). „*Perhaps the critical step in elucidating these concepts [interpreting the five factors] ... is*

- the specification of their exact nature: What are the elements or components of each factor? How are they interrelated?”* (p. 253). Block (1995) zitiert Briggs und bemängelt noch präziser: *„No functioning psychological ‚system‘, with its rules and bounds, is designated or implied by the ‚Big Five‘ formulation; it does not offer a sense of what goes on within the structured, motivation-processing, system-maintaining individual“* (p. 188).
- 9 Atomistisches Denken hat der Psychologie auch sonst geschadet – das war die Klage der Gestalt-Psychologen (s. Metzger, 1975), die offenbar wenig Wirkung zeigte, wenn man nicht gerade wie der Verfasser dieses Artikels vom relationalen Denkstil dieser Richtung noch beeinflusst wurde: Die vorliegende Kritik am Atomismus der Simple Structure wäre ohne Wolfgang Metzgers Grundbegriffe vom *Zusammenhang*, vom *Bezugssystem*, der *Zentrierungsdynamik* usw. (Metzger, 1941) vermutlich nicht entstanden. Kritik am Atomismus der Thurstone-Methodologie wurde schon von Richard Meili (1969) formuliert, der noch vor Ort in der „Berliner Schule“ geprägt wurde (durch Wolfgang Köhler und Kurt Lewin, Meili war Promovend in Berlin, 1924), und der sich dem Zwang der ‚SS‘ nur mit ständigem Widerspruch fügte: *„Das Ökonomieprinzip der Minimalstruktur“* habe zwar „einen gewissen heuristischen Wert“, sei aber „für die Natur nicht bindend“. Die Annahme, man könne Denken und Intelligenz theoretisch „auf die einfachste Weise erfassen“, sei fraglich. *„Wie wäre die Chemie wohl mit der analogen Hypothese gefahren, dass in jedem vorgefundenen Stoff möglichst nur ein Element vorkommen sollte“* (p. 288). Auch Jäger & Hörmann (1981, p. 418) haben einzuwenden: *„Die nach wie vor verbreitete und perpetuierte Analogie zur Chemie, nach der Faktoren quasi wie chemische Elemente ‚gesucht‘, ‚entdeckt‘ und als ‚existierend‘ bestätigt werden, ist irreführend“*.
  - 10 *„How should the Big-5-or-6 be understood in psychological terms? Sadly, despite many years of research – especially into extraversion – the picture is still very unclear (see e.g. A. Gale & M. W. Eysenck, 1992, Handbook of Individual Differences: Biological Perspectives; G. Matthews, 1993, in A. Smith & D. Jones, Factors Affecting Human Performance.) Here are some possibilities that still look viable yet falsifiable“*. (C. R. Brand, Quelle s. Anmerkung 5).
  - 11 Schon Velicer (1977) fand, dass die Extraktionsverfahren maximum likelihood, image analysis und principal components analysis bei neun Test-Datensätzen im Ergebnis ‚extremely similar‘ (p.18) ausfielen. Goldberg (1990) konnte bei einem repräsentativen Datensatz von Eigenschaftsbeschreibungen das Big-Five-Faktorenmodell fast invariant reproduzieren, egal ob er die Methode der Faktorenextraktion variierte (principal components, principal factors, alpha-factoring, image-factoring, maximum-likelihood) und egal ob er orthogonal (Varimax) oder schiefwinklig (Oblimin) rotierte. Doch diese methodischen Variationen wurden ausnahmslos mit Simple Structure-Orientierung vorgenommen. Alternativen zur Simple Structure wurden nicht diskutiert, obgleich zumindest Goldberg nach seinem Eysenck-Zitat (1981) eine Abkehr von Simple Structure als Denkmöglichkeit in Erwägung zog: *„Correlational psychology cannot in the nature of things come up with objective, universally agreed dimensions or categories; there are innumerable, mathematically equivalent ways of rotating factors, for instance, and no statistical magic key (not even simple structure) can close the door on alternative solutions“* (p. 45).
  - 12 Warburton (1963) beschreibt in einer Übersicht über die analytischen Methoden der Faktorenrotation die Logik mathematischer Sparsamkeit mit einfachen Worten: *„He [Ferguson] approached the problem of parsimony by considering a single variable, represented by a point, and asking himself what was its most parsimonious description. He suggested that, intuitively, the most parsimonious description results when one of the axes passes through the point. It is seen that when the reference frame is rotated so that one of the axes approaches the point, the product of the two co-ordinates grows smaller ...“* (p. 169).
  - 13 Beim Einsatz der exploratorischen Faktorenanalyse durch Internisten zur Ermittlung des Insulin Resistance Syndrom (ISR), zu dem der Blutdruck gehört, wurde zum Leidwesen der Forscher festgestellt, dass der systolische und diastolische Blutdruck nicht auf dem gleichen Faktor luden, obgleich *„... systolic and diastolic blood pressure are more strongly associated with each other than they are with other components of the insulin resistance syndrom, something which most clinicians would expect“* (Lawler et al. 2004, p. 6). Man half sich dadurch, dass man nur eine der beiden Blutdruckvariablen im ISR-Variablensatz beließ, um dem Methodenartefakt auszuweichen: *„Therefore, the evidence for inclusion or exclusion of hypertension in the definition oft he syndrome is based on whether one or two blood pressure measurements are included in the model rather than on any sound clinical or pathophysiological reasoning“* (Lawler et al., 2004, p. 1016).
  - 14 Schon 1958 wurde dieses Verfahren von britischen Autoren karikiert, die sich dem Thurstone-Dog-

- ma noch nicht verpflichtet fühlen: *„It has been said that what was thus thrown out of the door [the general factor] returned through the window: for correlated factors in turn give rise to a second order factor, and this is virtually the general factor of the centroid ... under another name“* (Hamilton, 1958), p. 167).
- 15 Dass mit hierarchischen Modellen schwerwiegende Folgeprobleme entstehen, kann hier nur angedeutet werden. ‚Hierarchie‘ suggeriert eine logische Gliederung nach tieferen und höheren Ebenen. Die bio-psychologische Realität aber operiert nicht nach Regeln der Begriffseinteilung. Was diese Realität tatsächlich aufweist, sind umfassendere und weniger umfassende Einflussgrößen, die *simultan* zusammenwirken. Die umfassenden Einflussgrößen, z. B. die allgemeine Intelligenz „g“, ist auch ohne Mitwirkung der weniger einflussreichen Bedingungen funktional manifest (ähnlich eine kritische Bemerkung von Revenstorf, 1976, p. 313). Es gibt auf der funktionalen Ebene nichts Über- und Untergeordnetes, sondern nur Umfassenderes (Durchschlagenderes) und Spezielleres (Modifizierendes).
  - 16 *„Für die Aufgaben der Konstrukteexplorierung im Bereich der Fragebogenskalen zur Persönlichkeit steht der entscheidende Beweis der Überlegenheit der Einfachstruktur-optimierenden Methode im Hinblick auf definitiv verbesserte variablenstichproben-übergreifende Strukturreplikationen noch aus.“* (Andresen, 1998, p. 74).
  - 17 Fittkau, 1968, p. 110: *„... die Ergebnisse jedes der analytischen Rotationsverfahren [sind] ... nicht invariant gegen Hinzunahme weiterer Variablen oder Ersetzen einiger Variablen durch andere“*. Butler, 1969, p. 13: *„the simple structure concept does not solve one of the most crucial and fundamental problems of factor analysis, the problem of the likelihood of factorial invariance“*.
  - 18 Die Behauptung, dass die Lage der Achsen einer Initiallösung ‚willkürlich‘ sei und deshalb ohne interpretativen Wert, stammt von Thurstone (1934): *„A characteristic of the multiple factor problem is that the location of the axes is arbitrary and that hence the factorial components are to that extent arbitrary and without fundamental psychological significance“*.
  - 19 Ein typisches Beispiel für das beliebte Ausschließen von Items, mit dem man dem vermeintlichen Zwang zur Simple Structure-Rotation entgehen möchte, gibt Thalbourne, 1998. Groner & Groner (1991) beschränken sich sogar nur auf den ersten von 17 (!) initialen Faktoren mit Eigenwerten größer als 1 allein aus Interpretationsgründen. Sabouria et al. (1990) informieren über eine jahrzehntelange Kontroverse, die um die Frage des Verlustes an erwünschter und behaupteter Eindimensionalität anhand eines unrotierten ersten Faktors entbrannte (im Zusammenhang mit der Entwicklung der Dyadic Adjustment Scale, ‚DAS‘ von Spanier). Bei Gangestad & Snyder (1985) und Snyder & Gangestad (1985, 1986) findet sich eine aufschlussreiche Begründung zur Rettung des Konstrukts ‚Self-monitoring‘, das durch Simple Structure-Rotationen verloren gehen würde. Dilemmata dieser Art werden zunehmend durch methodische Kompromisse zu entschärfen versucht, zu denen das Extrahieren von Faktoren zweiter Ordnung (führt zu einem ‚hierarchischen‘ Modell) und die Überprüfung diverser Modellmöglichkeiten durch konfirmatorische und Strukturgleichungsverfahren gehören (Undheim & Gustavsson, 1987: *„Restoring general intelligence“*). Die Ursachen der Probleme werden dadurch verschleiert.
  - 20 Hamilton (1958) bediente sich der seinerzeit noch aktuellen Methode der ‚simple summation‘, die ihm aus einem Angst-Fragebogen einen Generalfaktor der Angststärke lieferte ( $F_1$ ) neben einem schwächeren Faktor  $F_2$  der Angstebene (psychische vs. somatische Symptom-Dominanz). Ein Beruhigungspharmakon, Benactyzin, das den am Versuch teilnehmenden Patienten gegeben wurde, führte, wie erwartet, zu Veränderungen bei  $F_1$ , nicht bei  $F_2$ . Der Verfasser kritisiert Thurstones Simple Structure-Verfahren, bei dessen Verwendung faktorielle Deutungsprobleme auftraten und bei dem der Behandlungseffekt faktoriell nicht mehr als ein eindeutig angstreduzierender in Erscheinung trat. Eine Varimin-Reanalyse der mitgeteilten Korrelationsmatrix konnte die Kritik und Faktorendeutung Hamiltons voll bestätigen (unveröffentlichte Reanalyse des Verfassers).
  - 21 *„It is ... natural to postulate that when a unique simple structure is found for a battery of tests of mental abilities, then the non-vanishing entries in the factorial matrix are positive“* (Thurstone, 1947, p. 341). Thurstone meint an dieser Stelle, dass die einschränkende Bedingung der „positive manifold“ zur Anwendung des Prinzips der „simple structure“ nicht notwendig sei. Doch sei die Rotation zur Einfachstruktur vor allem dann angezeigt, wenn *„the factor loadings shall be positive or zero“*. (23) Berneyer (1957) referiert in diesem Sinne: *„The different methods of [factor] analysis [of mental aptitudes] yield factors which have negative loadings ... Such factors, so Thurstone contends, must be devoid of ‚scientific meaning‘. They do not permit us to ‚interpret the various tests as functions of the mental aptitudes“*

- which those tests elicit“* (p. 23). Auch C. Burt teilt uneingeschränkt Thurstones „positive manifold“-Idee: „... *an ability for x is by definition a dispositional property that facilitates doing x, i.e., it denotes a positive and never a negative tendency. Hence we shall be compelled to seek factors with positive saturations only“* (Burt, 1954, p. 18).
- 22 Die verfälschende Wirkung einer Simple Structure-Rotation bemerkten auch Highmore & Taylor (1954) bei der Analyse von Sporttest-Daten: „... *the ‚basic‘ factor, representing general athletic ability (in which we are primarily interested), necessarily disappears, and the group factors [of simple structure rotation] show little relation to the classification indicated by the [initial] bipolar matrix“* (p. 4).
- 23 Dies wurde geprüft und bestätigt gefunden mit Daten aus dem sportlichen Leistungsbereich (Zehnkampfleistungen, Ertel, b).
- 24 „*If they are able to choose one or the other, ... psychologists tend to prefer uni-directional to bi-polar measurement, probably as a result of the prestige of such measures as the standard metre in physics. ... Thurstone expressed a preference for a positive manifold (without, in the writer’s opinion, giving a fully convincing explanation) ... Bi-polar and unidirectional measurement are both needed in psychology“* (Thompson, 1963, p.22). Die letztere Behauptung findet sich ausführlich in Thompson, 1962, begründet.
- 25 Kritik am Prinzip der Einfachstruktur ist nur in Ausnahmefällen laut geworden: „*Die Entscheidung zwischen den vielen mathematisch möglichen Lösungen wurde durch eine formale Regel, das Minimalprinzip („simple structure“) getroffen, dem jede theoretische Begründung fehlt. Frei von jedem Zwange, grundlegendere Überlegungen anstellen zu müssen, wurde nun Massenarbeit geliefert ... Der Erfolg dieses Vorgehens war, wie vorauszusehen, eine Überproduktion von oberflächlich definierten Faktoren“* (Meili, 1969, p. 278). Auch Revenstorf (1976) hielt Einfachstrukturen, zumal bei „großen Merkmalzusammenstellungen (Fragebögen) ... für im allgemeinen unwahrscheinlich“, weil die „*Merkmale mehr und mehr alle möglichen Kombinationen von Faktorausprägungen aufweisen können. In diesem Falle streuen die Merkmale der Variablenkonfigurationen über den ganzen Faktorenraum und eine Einfachstruktur im Thurstone’schen ... Sinne ist nicht mehr festzustellen“* (p. 321). Unzählige Beispiele aus dem Alltag der Faktoreninterpretation ließen sich anführen: „*Es ist klar, dass die Einfachstruktur zu wünschen übrig lässt“* (Bierhoff, 2000); „... *ist trotz der vergleichsweise hohen Interkorrelationen nicht von einer Einfachstruktur auszugehen“* (Schaper & Baumgart, 2002); „*Insgesamt erscheint die Prägnanz der Einfachstruktur eingeschränkt“* (Beauducel et al., 2003), „... *keine Einfachstruktur nach dem Bargman-Test“* (Herzberg, P. Y., 2002); „*Eine akzeptable Einfachstruktur der Faktorladungsmatrix ließ sich weder durch orthogonale, noch durch schiefwinklige Rotationen der drei Hauptachsen erreichen“* (Schmitt, M. , 2000); „*Die Lösung lässt nicht auf eine Einfachstruktur schließen“* (Lambert et al. 2002).
- 26 Einfachste anthropometrische Merkmale wie Körpergröße und Haarfarbe beruhen auf einem Zusammenwirken mehrerer Gene.
- 27 Methodische Ansätze, die der Kombinatorik latenter Größen bei der Produktion manifester Einheiten Rechnung tragen, findet man außerhalb faktorenanalytischen Denkens, so z. B. in Wiley’s ‚Latent Partition Analysis‘: „... *it is assumed that the manifest categories are derived by dividing and combining the latent categories. We have found that by examining manifest categories it is possible to reconstruct information about the latent partition and about its relation to the manifest partitions“*. (Wiley, 1969, p. 183).
- 28 „*Language learning appears to be based on a relatively plastic mix of neural systems that also serve other functions. I believe that this conclusion renders the mysteries of language evolution ... somewhat more tractable. That is, the continuities that we have observed between language and other cognitive systems make it easier to see how this capacity came about in the first place“*. (Bates, E. Modularity, domain specificity, and the development of language. <http://www.ecs.soton.ac.uk/~harnad/Papers/Py104/bates-1994.html>)
- 29 Unter den zahlreichen zitierbaren Aussagen sei ausgewählt: „*Ein lebendes System ist ein struktur-determiniertes System, und alles in diesem System geschieht aufgrund von Nachbarschaftsrelationen zwischen seinen Bestandteilen ...“* (Maturana, 1998, p. 184).
- 30 Die letzte umfassendere Kritik an der Faktorenanalyse im deutschsprachigen Raum (durch Holz-Ebeling, 1995) nimmt Anstoß an der Vagheit bei der Deutung von Faktoren. Die Autorin fordert mit Recht, dass die Auffassung multipler Bedingtheit (multi-UV-Bedingtheit) der untersuchten Variablen (AV) aufrechterhalten werden sollte und stellt fest, dass die Faktorenanalyse der multi-UV-Bedingtheit der AV nicht gerecht wird. Sie schlägt deshalb methodisch ein der Varianzanalyse verwandtes Verfahren vor, das die Faktorenanalyse „ersetzen oder ergänzen“ soll. Für ein varianzanalytisches Denken ist die Existenz

- von multi-UV, welche auf einzelne AV Einfluss nehmen, selbstverständlich. Der Vorschlag der Autorin, so unpraktisch er sonst auch ist (nach ihrem eigenen Urteil), ist interessant als Versuch einer Korrektur von Fehlern der komplexitätsfeindlichen Simple Structure-orientierten Faktorenanalyse. Die Richtung stimmt ungefähr, doch rührt die Autorin die eigentliche Ursache des Dilemmas nicht an.
- 31 Immer wieder sind Faktorenanalytiker, die sich mit sperrig-komplexen Variablen konfrontiert sahen, an die Grenzen des Prinzips Simple Structure gestoßen. Guilford & Zimmerman (1963) wagen eine Liberalisierung des Sparsamkeitsprinzips, um Komplexität zuzulassen: „*In general, investigators need to relax somewhat their drive to achieve parsimony. Long ago, psychology should have progressed beyond the stage in which investigators continue to look for the ‚philosopher’s stone‘*“ (p. 299). Vertritt man das Komplexstruktur-Prinzip, so sind an der Sparsamkeitsforderung keinerlei Abstriche zu machen. Man hat nur anzuerkennen, dass die Sparsamkeitsstrategien der *Natur* andere sind als die, die Thurstone zur geometrischen Abbildung der von der Natur hervorgebrachten Variablen für geeignet hielt.
- 32 Aus der Perspektive des Komplexstruktur-Prinzips lässt sich auch die von Thurstone schon erfundene und von Cattell vehement propagierte Strategie der schiefwinkligen Rotation besser verstehen. Sie stellt ein missglücktes Spagat dar zwischen der Suche nach einer einfaktoriellen Abbildung von Variablen und der Einsicht, dass man darüber hinaus auch der Komplexität genügen muss, die im Universum der beobachtbaren Variablen natürlicherweise vorhanden ist. „*It is unreasonable to expect that a great variety of influences operating and interacting in the same universe would be completely uncorrelated*“ (Cattell & Dickman, 1962, p. 390). Der schiefwinklig rotierende Faktoranalytiker versucht deshalb, Interaktionen zwischen den Variablen durch Korrelation der Faktoren im nicht-orthogonalen Faktorenraum abzubilden. Die Variablen bleiben dabei Simple Structure-gebunden, und dies in besonderem Maße. Gleichzeitig wird durch die Hintertür, indem man die Korrelierbarkeit der Faktoren untereinander zulässt, einem funktionalen Zusammenhang Einlass gewährt – ein dubioser Kompromiss, den schon Guilford & Zimmermann (1963, p. 289) kritisierten: „*[This amounts to] a hollow and accidental victory for oblique methods of rotation*“, der durch Verzicht auf die theoretisch unverzichtbare Unterscheidbarkeit funktionaler Einheiten erkauft wird. Die Pseudolösung
- des Problems mithilfe schiefwinkliger Rotation wird hier deshalb erwähnt, weil sie die Spannung verdeutlicht, die das frustrierende Streben nach Einfachstruktur zur Folge hat: Man versucht, das selbst geschaffene Übel durch hochproblematische Manöver zu beseitigen.
- 33 Symptomatisch für die irrige Simple Structure-Orientierung ist auch der Versuch von Cattell & Radcliffe (1962), die real vorhandene Komplexität, die durch Rotation zur Simple Structure nicht beseitigt wird, durch Unterdrückung unerwünschter Varianz zu eliminieren: „*If we are right in assuming that behaviour which has a large personality factor variance will generally be factorially complex, then the unwanted common factor variance will, as a rule, be far from negligible*“. Den Autoren liegt daran, „unifactor scales“ zu entwickeln und sie stellen zu diesem Zweck ein Verfahren dar, „... *which will reduce the contribution of unwanted factors by suppression*“ (p. 125).
- 34 Die anschauliche Prägnanz der Einfachstruktur ist zwar weniger durchschlagend wie die des täglichen Sonnen-Umlaufs um die Erde, die bis zur kopernikanischen Wende gravierende kosmologische Verzerrungen zur Folge hatte. Doch scheint die Ursache der Simple Structure-Verirrung auf ähnlichen Mechanismen zu beruhen. Die komplizierten Circumplex-Nachbesserungen bei Simple Structure-rotierten Faktoren erinnern an die Zusatzberechnungen für die planetarischen Anomalien (Epizyklen), mit denen die geozentrische Vorstellung vom Kosmos konserviert wurde. Auch Quine führt als Grund für die generelle Vorliebe für Einfachheit neben dem motivationalen Wunschdenken einen basaleren Wahrnehmungsmechanismus an: „*There is a subjective selectivity that makes us tend to see the simple and miss the complex*“ (Quine, 1963, p. 103).
- 35 „*It may be contended that we should cling at all costs to the conception of simple structure, because we have no satisfactory alternative*“ (Reybourne & Rath, 1949, p. 127).
- 36 Eine Relativierung ist hier nötig: Da die *interindividuelle* Varianz von Verhaltensvariablen in einer kurzzeitig geförderten Forschungsperiode relativ leicht feststellbar ist, bedient man sich in der Forschungspraxis vornehmlich der hierfür zuständigen R-Technik. Ungleich aufwändiger wäre die Ermittlung *intraindividuell* ablaufender Kovarianz, die größere Zeiträume umfasst. Die *P-Technik*, die für die Aufklärung der dabei anfallenden Daten einzusetzen wäre, ist deshalb in der Forschungsszene eine Rarität, eine gleichzeitige Aufklärung *intraindividuell* zusammen mit *interindividuell*er Varianz, die auf einer gemein-

- samen konzeptuellen Basis stattfinden könnte, wird hauptsächlich aus äußeren (ökonomischen) Gründen selten versucht.
- 37 Die Implikationsthese Gigerenzers wurde als „central thesis“ einer Publikation zur multidimensionalen Skalierung von Smith & Jones (1975) vorweggenommen: „*All data analysis and all scaling involve fundamental assumptions about the psychological processes that lead to the data and the scaling solutions under consideration. In particular, current scaling methods, in our view, should be regarded with deepest suspicion, precisely because they are based on doubtful or untested psychological assumptions*“ (p. 44).
- 38 Basilevsky (1994, p. 415) beschreibt die Tendenz in der CFA-Praxis zum Herauspräparieren singulärer Faktoren bei den jeweils untersuchten Variablen u. a. wie folgt: „... *we may wish to impose zero restrictions on the loadings. Values other than zeroes can also be used, but zeroes are most common in practice*“.
- 39 „*Dass keine der latenten Ursachen eindeutig definiert ist, wird im LISREL-Manual taktvoll verschwiegen*“ (Schönemann & Borg, 1996, p. 250). Das „Unbestimmbarkeitsproblem“, ein rein mathematisch-formales Problem, sei im LISREL-Fall „tatsächlich verschlimmert ...“, weil man dort meist wesentlich mehr latente Variablen postuliert als im Modell der multiplen Faktorenanalyse ...“ (p. 250).
- 40 „*Die praktisch schrankenlose Plastizität des LISREL-Modells ... unterhöhlt nicht nur den Anspruch auf statistische Inferenz, sie fordert fast zum Missbrauch auf, denn bei hinreichender Geduld wird sich schon irgendein Kausalmodell finden, das für die gerade vorliegenden Daten nicht abgelehnt werden muss*“ (Schönemann & Borg, 1996, p. 250).
- 41 Rost (2005): „*Der Modellspezialist [Spezialist im Modellieren mit Strukturgleichungen, log-linearen Modellen, Item-Response-Modellen usw.] ... kann mit Daten überhaupt nichts anfangen, wenn man ihm nicht sagt, welche Variable mit welcher anderen interagieren soll ..., welche latenten Variablen es geben soll ... usw.*“. Mein Kommentar hierzu: Der heutige ‚Modellspezialist‘ traut sich a priori zu, was der exploratorische Faktorenanalytiker früherer Jahrzehnte trotz Analyse massenhaft vorliegender Daten nicht einmal im Nachhinein hat gewinnen können, nämlich eine approximative Einsicht in die grundlegenden latenten Parameter menschlichen Denkens, Fühlens und Verhaltens. Diese Parameter, so setzt man heute vielfach voraus, könne man sich ausdenken, man müsse sie nur mit Versuch-und-Irrtum-Strategie je nach Modellfit-Ergebnissen heraus präparieren.
- 42 Das Zitat ist einem Artikel von Barrett (2002) entnommen, der die konventionelle Psychometrie insgesamt kritisiert, weil sie durch Mess-Operationen die Inhaltsebene (Intelligenz, Persönlichkeit) entleert. Allerdings scheint Barrett selbst höhere inhaltliche Ambitionen aufzugeben. Er plädiert, gewissermaßen ersatzweise, für Offenheit gegenüber *Anwendungsfragen*. Barrett propagiert eine sogenannte „applied numerics“, die von schwankenden theoretischen Erwartungen befreien und pragmatisch umso mehr nützen soll (Absage an theoretische Ansprüche).
- 43 Wenn die vernachlässigten Vorgehensweisen *phänomenologische* genannt werden, erscheinen sie möglicherweise wissenschaftshistorisch belastet. Auch mit der Bezeichnung *hermeneutisch* könnte man Assoziationen zu geronnenen Denkstilen hervorrufen und von Meinungsgegnern entsprechend abgestempelt werden. Als Desiderat darf man indessen einbringen, dass in allen Phasen der Problemlösung die inneren Hindernisse überwunden werden sollten, die sich der Verwendung menschlicher Erkenntnis mit ihren weit reichenden Verzweigungen entgegenstellen. Die Bemühungen in den Natur-, Geistes- und Kulturwissenschaften sind in ihrer Gesamtheit Niederschlag eine Entfaltung epistemischer Möglichkeiten. Sie sollten bei Bedarf in jeder einzelnen Disziplin, wenn auch mit unterschiedlichen Akzenten, immer abruffähig bleiben. Dies ist auch das Hauptanliegen des Dissidenten Sigmund Koch, 1999, und anderer ungehörter Rufer, die die ständig ins Objektivistische abgleitenden Wissenschaft kritisieren (Bridgman, 1959).
- 44 S. Jevons (1873): „*Ich neige dazu, an mathematischen Autoren herumzunörgeln, weil sie oftmals bejubeln, was sie alles leisten können, und es unterlassen, darauf hinzuweisen, dass das, was sie machen, nur ein winzig kleiner Teil dessen ist, was gemacht werden könnte. Sie zeigen ... allgemein die Neigung, die Existenz von widerspenstigen Problemen auch nicht einmal zu erwähnen ...*“ (zitiert nach Rescher, 1985, p. 124).
- 45 Die Auffassung der derzeit tonangebenden formalmodellierenden Fachvertreter scheinen einen ‚Paradigmenwechsel‘ kategorisch auszuschließen. So spricht etwa Rost (2005) von „ehernen Gesetzen, die jede Modeströmung überstehen ... oder gar gestählt aus ihr hervorgehen“. Nach Rost können Moden und Zeitgeist der Psychologie den Methodeneinsatz zwar am Rande affizieren, doch dies werde nicht dazu führen, „*dass sich das bisherige Arsenal an Forschungsmethoden als falsch oder unbrauchbar erweist. Vielmehr führen sie [die Moden] nur dazu, dass sich das Methodenarsenal*

um wesentliche Teile erweitert und bereichert“. Dass auch fehlkonstruierte Teile in diesem Methodenarsenal vorhanden sein und u. U. großen Schaden anrichten können und in diesem Fall grundlegend überarbeitet werden müssen, hätte Rost in seiner generalisierenden Rückschau eigentlich zu ergänzen.

- 46 Herrn Kollegen Jürgen Guthke und seiner Mitarbeiterin, Frau Seiwald, sei gedankt dafür, dass sie meinen Beitrag zu einer Problemlösung in ihrer Forschungsarbeit erbat, ebenso für ihre Einladung zur Präsentation des vorliegenden Ansatzes im Rahmen eines Leipziger Kolloquiums. Gestärkt und ermuntert haben mich auf diesem steinigen Dissidenten-Weg auch Zusprüche von Tatjana Schnell, Gerd Lüer, Nikola Kohls und André Beauducel.
- 47 Dies wird mit dem Programm *faktor.exe* ermöglicht, geschrieben in meinem Auftrag von Uwe Engeland. Es wird im Folgeartikel beschrieben. Bisherige Erträge der komplexitätsoffenen Faktorenanalyse wurden niedergelegt in Ertel, a, b, c, d, e.

## Literatur

- Acton, G. S. & Revelle, W. (2004). Evaluation of the psychometric criteria for circumplex structure. *Methods of Psychological Research Online*, 9 (1), 1-29.
- Andresen, B. (1998). *Persönlichkeitsstruktur und Psychosetendenz. Bd. 1: Grundlagen, Methoden, Basisfaktoren*. Westerau: Psychometrie- & Psychodiagnostik-Verlag.
- Arkes, H. R. (1991). Costs and benefits of judgment errors: implications for debiasing. *Psychological Bulletin*, 110 (35), 486-498.
- Barrett, P. (2002). Beyond psychometrics: Measurement, non-quantitative structure, and applied numerics. *Augmented web version eines Artikels aus The Journal of Managerial Psychology*.
- Basilevski, A. (1994). *Statistical factor analysis and related methods. Theory and applications*. New York: Wiley.
- Bates, E., Thal, D., & Janowsky, J. (1992). Early language development and its neural correlates. In I. Rapin & S. Segalowitz (Eds.), *Handbook of neuropsychology*, Vol. 7: Child neuropsychology. Amsterdam: Elsevier.
- Beauducel, A., Strobel, A., & Brocke, B. (2005). Psychometrische Eigenschaften und Normen einer deutschsprachigen Fassung der Sensation Seeking-Skalen, Form V. *Diagnostica*, 49(2), 61-72.
- Beauducel, A. & Wittmann W. W. (2005). Simulation study on fit indices in confirmatory factor analysis based on data with slightly distorted simple structure. *Structural Equation Modeling*, 12, 41-75.
- Bechtel, W., & Richardson, R. C. (1993). *Discovering complexity*. Princeton NJ: Princeton University Press.
- Berneyer, G. (1957). Psychological factors. Their number, nature, and identification. *The British Journal of Statistical Psychology*, 10, 17-27.
- Bierhoff, H. W. (2000). Skala der sozialen Verantwortung nach Berkowitz und Daniels: Entwicklung und Validierung. *Diagnostica*, 46 (1), 18-28.
- Blinkhorn, S. F. (1997). Past imperfect, future conditional: Fifty years of test theory. *British Journal of Mathematical and Statistical Psychology*, 50, 175-185.
- Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, 117 (2), 187-215.
- Bortz, J. (2004). *Lehrbuch der Statistik*. Sechste Auflage. New York: Springer. Erste Auflage 1977.
- Bridgman, P.W. (1959). *The way things are*. Cambridge: Harvard University Press.
- Briggs, S. R. (1989). The optimal level of measurement for personality constructs. In D. M. Buss & N. Cantor (Eds.), *Personality Psychology: Recent Trends and Emerging Directions* (pp. 246-260). New York: Springer.
- Burroughs, G. E. R., & Miller, H. W. L. (1961). The rotation of principal components. *The British Journal of Statistical Psychology*, 14 (Part 1), 35-42.
- Burt, C. (1949). The structure of the mind: A review of the results of factor analysis. *The British Journal of Educational Psychology*, 19, 100-111, 176-199.
- Burt, C. (1954). The sign pattern of factor matrices. *The British Journal of Statistical Psychology*, 7, Part 1, 15-29.
- Buse, L. & Pawlik, K. (1978). Stichprobenfehler und Rotationsambiguität von Faktorladungen. *Zeitschrift für experimentelle und angewandte Psychologie*, 25, 10-34.
- Butler, J. M. (1969). Simple structure reconsidered: Distinguishability and invariance in factor analysis. *Multivariate Behavior Research*, 4, 5.
- Caspi, A. (1998). *Social, Emotional, and Personality Development*. In W. Damon & N. Eisenberg (Eds.), *Handbook of Child Psychology* (Vol. 3, pp. 311-388). New York: John Wiley & Sons.
- Cattell, R. B. (1965). Factor analysis: An introduction to its essentials. II: The role of factor analysis in research. *Biometrics*, 21, 405-435.

- Cattell, R. B., & Dickman, K. (1962). A dynamic model of physical influences demonstrating the necessity of oblique simple structure. *Psychological Bulletin*, 59, 389-400.
- Cattell, R. B., & Radcliffe, J. A. (1962). Reliabilities and validities of simple and extended weighted and buffered unifactor scales. *The British Journal of Statistical Psychology*, 15(2), 115-128.
- Child, D. (1990). *The essentials of factor analysis*. London: Cassell, 1990, 2<sup>nd</sup>. Edition, 1<sup>st</sup> edition: 1971.
- Church, A. T., & Burke, P. J. (1994). Exploratory and confirmatory tests of the big five and Tellegen's three- and four-dimensional models. *Journal of Personality and Social Psychology*, 60, 93-114.
- Cliff, N. (1983). Some cautions concerning the application of causal modeling methods. *Multivariate Behavioral Research*, 18, 115-126.
- Cliff, N. (1992). Abstract measurement theory and the revolution that has never happened. *Psychological Science*, 3, 186-190.
- Comrey, L., & Lee, H. (1992). *A first course in factor analysis*. Hillsdale, NJ: Lawrence Erlbaum, 2<sup>nd</sup> Edition. 1<sup>st</sup> Edition Comrey alone: 1973, New York: Academic Press.
- Davies, W. K. D. (1971). Varimax and the destruction of generality: a methodological note. *Area*, 3, 112-118.
- De Raad, B., Hendriks, A. A. J., & Hofstee, W. K. B. (1994). The Big Five: a tip of the iceberg of individual differences. In C. J. Halverson, Jr., G. A. Kohnstamm, and R. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 91-109). Hillsdale, NJ: Erlbaum.
- Edmonds, B. (2002). *Simplicity is not truth-indicative* (No. CPM Report No. CPM-OL-99).
- Ertel, S. (a). Basismerkmale der Persönlichkeit. Exploratorische Faktorenanalyse mit komplextheoretischer Orientierung. *Nach einem Vortrag auf der Siebten Arbeitstagung der Fachgruppe Differentielle Psychologie, Persönlichkeitspsychologie und Psychologische Diagnostik. Martin-Luther-Universität Halle-Wittenberg, 29. und 30. September 2003.* (unveröffentlicht).
- Ertel, S. (b) Offen für Komplexstruktur. Exploratorische Faktorenanalyse von Zehnkampf- und anderen sportlichen Leistungen (unveröffentlicht).
- Ertel, S. (c) Faktorielle Komplexstruktur durch Varimin-Rotation. Manifestation simulierter Zusatzvarianz in I-S-T Daten. (unveröffentlicht).
- Ertel, S. (d) Intelligenz (g) und Lernkapital. Komplexstrukturelle Faktorenanalyse des I-S-T (unveröffentlicht).
- Ertel, S. (e) Wozu Simple Structure? Fragen an die Experten multivariater Verfahren. (unveröffentlicht).
- Estes, W. K. (1975). Some targets for Mathematical Psychology. *Journal of Mathematical Psychology*, 12, 263-282.
- Eysenck, H. J. (1992). Four ways five factors are not basic. *Personality and Individual Difference*, 13, 667-673.
- Eysenck, H. J. (1997). Personality and experimental psychology: The unification of psychology and the possibility of a paradigm. *Journal of Personality and Social Psychology*, 73, 1224 - 1237.
- Fabrigar, L. R., Wegener, D. Z., McCallum, R. C. & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods* 4, 272-299.
- Ferrando, P. J., & Lorenzo-Seva, U. (2000). Unrestricted versus restricted factor analysis of multidimensional test items: Some aspects of the problem and some suggestions. *Psicologica*, 21, 301-323.
- Finch, J. F., & West, S. G. (1997). The investigation of personality structure: Statistical models. *Journal of Research in Personality*, 31, 439-485.
- Fittkau, B. (1968). Die Mehrdeutigkeit der Varimax-Rotation und einige sich daraus ergebende Folgerungen. *Archiv für die gesamte Psychologie*, 120, 106-114.
- Fleck, L. (1935). *„Entstehung und Entwicklung einer wissenschaftlichen Tatsache. Einführung in die Lehre vom Denkstil und Denkkollektiv.“* Basel: Schwabe.
- Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: A critical review and analysis. *Personnel Psychology*. 39, 291-314.
- Forster, M. (1998). Parsimony and simplicity. *Available from: <http://philosophy.wisc.edu/forster/220/simplicity.html>*.
- Forster, M., & Sober, E. (1994). How to tell when simple, more unified, or less ad hoc theories will provide more accurate predictions. *The British Journal for the Philosophy of Science*, 45, 1-35.
- Funder, D. C. (2001). Personality. *Annual Review of Psychology*, 52, 197-121.
- Gangestad, S. & Snyder, M. (1985). „To carve nature at its joints“: On the existence of discrete classes in personality. *Psychological Review*, 92. 317-349.
- Gawronski, B. (2000). Falsifikationismus und Holismus in der experimentellen Psychologie: Logische Grundlagen und methodologische Konsequenzen. *Zeitschrift für Sozialpsychologie*, 31 (1), 5-17.
- Gibbons, R. D. & Hederer, D. R. (1992). Full-information item bi-factor analysis. *Psychometrika*, 57, 423-436.

- Gigerenzer, G. (1978). Artefakte in der dimensionsanalytischen Erfassung von Urteilsstrukturen. *Zeitschrift für Sozialpsychologie*, 9, 110-116.
- Gigerenzer, G. & Strube, G. (1978) Zur Revision der üblichen Anwendung dimensionsanalytischer Verfahren. *Zeitschrift für Entwicklungspsychologie und Pädagogischen Psychologie* 10 (1), 75-86.
- Glass, G. V. (1991). Eulogy for Henry Kaiser (1927-1992). *Multivariate Behavioral Research Monographs*, 27, 159-171.
- Glass, G. V. & Taylor, P. A. (1966). Factor analytic methodology. *Review of Educational Research*, 36, 566-587.
- Goldberg, L. R. (1981). Developing a taxonomy of trait-descriptive terms. In D. W. Fiske (Ed.), *Problems with language. New directions for methodology of social and behavioral science, No. 6* (pp. 43-65). San Francisco: Jossey-Bass.
- Goldberg, L. R. (1990). An alternative „Description of Personality“: The Big-Five Factor Structure. *Journal of Personality and Social Psychology*, 59, 1216-1229.
- Gorsuch, R. L. (1983). *Factor analysis*. Philadelphia: Saunders. 1<sup>st</sup> Edition: 1974.
- Gould, S. J. (2002). *The Structure of Evolutionary Theory*. Cambridge (MA): Harvard UP-Belknap Press.
- Greif, S. (1972). Gruppenintelligenztests – Untersuchungen am WIT, IST, LPS und AIT. Frankfurt: Peter Lang.
- Groner, R., & Groner, M. T. (1991). Heuristische versus algorithmische Orientierung als Dimension des individuellen kognitiven Stils. In K. Grawe, R. H. Hänni, N. Semmer & F. Tschan (Eds.), *Über die richtige Art, Psychologie zu betreiben*. Göttingen: Hogrefe.
- Guilford, J.P. (1974). Rotation problems in factor analysis. *Psychological Bulletin*, 81, 495-501.
- Guilford, J. P., & Zimmerman, W. S. (1963). Some variable-sampling problems in the rotation of axes in factor analysis. *Psychological Bulletin*, 60, 289-301.
- Guttman, L. (1954). A new approach to factor analysis: The radix. In P. F. Lazarsfeld (Ed.) *Mathematical Thinking in the social sciences* (pp. 258-348). Glencoe, IL: Free Press.
- Guttman, L. (1955). The determinacy of factor score matrices, with implications for five other basic problems of common-factor theory. *British Journal of Statistical Psychology*, 8, 65-81.
- Guttman, L. (1971). Measurement as structural theory. *Psychometrika*, 36, 329-347.
- Guttman, L. (1992). The irrelevance of factor analysis for the study of group differences. *Multivariate Behavioral Research*, 27, 175-204.
- Halberg, F. (1980). Closing remarks: Chronobiology and health. In L. E. Scheving & F. Halberg (Eds.), *Chronobiology: Principles and applications to shifts in schedules* (pp. 541-562). Alphen aan den Rijn: Sijthoff and Noordhoff.
- Hamilton, M. (1958). An experimental approach to the identification of factors. *The British Journal of Statistical Psychology*, 11 (2), 161-169.
- Hamilton, M. (1960). A rating scale for depression. *Journal of Neurological Psychiatry*, 23, 56-62.
- Harman, H. H. (1968). *Modern Factor Analysis* (2<sup>nd</sup> ed.). Chicago: The University of Chicago.
- Highmore, G., & Taylor, W. R. (1954). A factorial analysis of athletic ability. *The British Journal of Statistical Psychology*, 7 (Part I), 1-8.
- Hofstee, W. K. B., de Raad, B., & Goldberg, L. R. (1992). Integration of the Big Five and circumplex approaches to trait structure. *Journal of Personality and Social Psychology*, 63, 147-165.
- Holz-Ebeling, F. (1995). Faktorenanalysen und was dann? *Psychologische Rundschau*, 46, 18-35.
- Horn, J. L. (1967). On subjectivity in factor analysis. *Educational and Psychological Measurement*, 27, 811-820.
- Jäger, A. O., & Hörmann, H. (1981). Demonstrationen von „g“ (der Allgemeinen Intelligenz) und zur Bedeutung des Variablenkontextes bei exploratorischen Faktorenanalysen. *Psychologische Beiträge*, 23, 408-420.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport: Praeger.
- Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, 23, 187-200.
- Kaiser, H. F. (1970). A second generation "Little Jiffy". *Psychometrika*, 35, 401-416.
- Kallina, H. (1967). Das Unbehagen in der Faktorenanalyse. Sonderheft: Münchener Symposium über Faktorenanalyse. *Psychologische Beiträge*, 10, 81-86.
- Kalveram, K. T. (1970). Über Faktorenanalyse. Kritik eines theoretischen Konzepts und seine mathematische Neuformulierung. *Archiv für Psychologie*, 122, 92-118.
- Kanizsa, G. (1975). „Prägnanz“ as an obstacle to problem solving. *Giornale Italiano di Psicologia*, 2(3), 417-425.
- Kaplunovsky, A. S. (2007). Why using factor analysis? Dedicated to the centenary of factor analysis. HAIT Journal of Science and Engineering. Prepublication draft, Internet-zugänglich. Zum gleichen Thema: Kaplunovsky, A. S. (2005). Factor Analysis in Environmental Studies, HAIT Journal of Science and Engineering B 2, 1-2, 54-94.

- Kazdin, A. E. (2006). Arbitrary metrics. Implications for identifying evidence-based treatment *American Psychologist*, *61*, 41-49.
- Koch, S. (1999). *Psychology in human context. Essays in dissidence and reconstruction*. Chicago: The University of Chicago Press.
- Lakatos, I. (1977). *The Methodology of Scientific Research Programmes: Philosophical Papers Volume 1*. Cambridge: Cambridge University Press.
- Lambert, M. J., Hannöver, W., Nisslmüller, K., Richard, M., & Kordy, H. (2002). Fragebogen zum Ergebnis von Psychotherapie: Zur Reliabilität und Validität der deutschen Übersetzung des Outcome Questionnaire 45.2 (OQ-45.2). *Zeitschrift für Klinische Psychologie und Psychotherapie*, *31* (1), 40-46.
- Lawler, D. A., Ebrahim, S., May, M., & Smith, G. D. (2004). (Mis)use of factor analysis in the study on Insulin Resistance Syndrome. *American Journal of Epidemiology*, *159*, 11, 1013-1018.
- Lenk, W. (1985) Faktorenanalyse: ein Mythos? Historische und konzeptionelle Untersuchungen zur Faktorenanalyse und Intelligenzforschung. Weinheim: Beltz.
- Leventhal, D. S., & Stedman, D. J. (1970). A factor analytic study of the Illinois Test of Psycholinguistic Abilities. *Journal of Clinical Psychology*, *26*, 473-476.
- Lienert, G. A. (1969). *Testaufbau und Testanalyse*. Dritte, durch einen Anhang über Faktorenanalyse ergänzte Auflage. Weinheim: Beltz.
- Lukesch, H. & Kleiter, G. D. (1974). Die Anwendung der Faktorenanalyse. Darstellung und Kritik der Praxis einer Methode. *Archiv für Psychologie*, *126*, 265-307.
- Lumsden, J. (1961). The construction of unidimensional tests. *Psychological Bulletin*, *58*, 122-131.
- MacCallum, R. C. (2003). Working with imperfect models. *Multivariate Behavioral Research Monographs*, *38* (1), 115-139.
- MacCallum, R. C., & Austin, J. T. (2000). Application of structural equation modelling in psychological research. *Annual Review of Psychology*, *51*, 201-226.
- Maturana, H. R. (1998). Biologie der Realität. Frankfurt: Suhrkamp.
- Maxwell, A. E. (1959). Statistical methods in factor analysis. *Psychological Bulletin*, *56*, 228-235.
- Maxwell, A. E. (1972). Factor analysis: Thomson's sampling theory recalled. *British Journal of Statistical Psychology*, *25*, 1-21.
- Meili, R. (1969). Die Zwiespältigkeit faktorieller Untersuchungen der Intelligenzstruktur. *Zeitschrift für experimentelle und angewandte Psychologie*, *16*, 278-294.
- Metzger, W. (1941). *Psychologie. Die Entwicklung ihrer Grundannahmen seit der Einführung des Experiments*. Darmstadt: Steinkopf (Letzte Auflage 2002, Wien: Krammer).
- Michell, J. (2000). Normal science, pathological science and psychometrics. *Theory and Psychology*, *10*, 639-667.
- Moosbrugger, H. & Hartig, J. (2002). Factor analysis in personality research: Some artifacts and their consequences for psychological assessment. *Psychologische Beiträge*, *44*, 136-158.
- Mulaik, S. (1972). *The foundations of factor analysis*. New York: McGraw Hill.
- Narens, L. & Luce, R. D. (1993). Further comments on the «nonrevolution» arising from axiomatic measurement theory. *Psychological Science*, *4*, 127-130.
- Neuhaus, J. O., & Wrigley, C. F. (1954). The quartimax method: an analytic approach to orthogonal simple structure. *British Journal of Statistical Psychology*, *7*, 81-91.
- Orlik, P. (1967). Das Dilemma der Faktorenanalyse - Zeichen einer Aufbaukrise in der modernen Psychologie. Sonderheft: Münchener Symposium über Faktorenanalyse. *Psychologische Beiträge*, *10*, 87-98.
- Pawlik, K. (1977). Faktorenanalytische Persönlichkeitsforschung. In G. Strube (Ed.), *Binet und die Folgen. Testverfahren, Differentielle Psychologie, Persönlichkeitsforschung. Bd. 5. Die Psychologie des 20. Jahrhunderts*. Zürich: Kindler.
- Peel, E. A. Factorial analysis as a psychological technique. *Nordic Psychology Monographs*, 7-22.
- Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). The use of factor analysis for instrument development in health care research. Beverly Hills, CA: Sage.
- Quine, W. v. O. (1963). On simple theories of a complex world. *Synthese*, *15*, 103-106.
- Rescher, N. (1985). *Die Grenzen der Wissenschaft*. Stuttgart: Reclam.
- Rescher, N. (1998). *Complexity: A philosophical overview*. New Brunswick: Transaction Publishers.
- Revelle, W. (1985). Factors are fictions, and other comments on individuality theory. *Journal of Personality*, *51*, 707-715.
- Revenstorf, D. (1976). *Lehrbuch der Faktorenanalyse*. Stuttgart: Kohlhammer.
- Revenstorf, D. (1980). *Vom unsinnigen Aufwand*. Stuttgart: Kohlhammer.
- Reyburne, H. A., & Rath, M. J. (1949). Simple structure: A critical examination. *The British Journal of Psychology*, *2*, 125-133.
- Reymert, R., & Jöreskog, K. G. (1993). *Applied Factor Analysis in the Natural Sciences*. Cambridge: Cambridge University Press.

- Rost, J. (2003, Februar). Zeitgeist und Moden empirischer Analysemethoden [45 Absätze]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research [Online Journal]*, 4(2).  
Verfügbar über: <http://www.qualitative-research.net/fqs-texte/2-03/2-03rost-d.htm>.
- Rozeboom, W. W. (1991). Theory and practice of analytic hyperplane optimization. *Multivariate Behavioral Research*, 26(1), 179-197.
- Rummel, R. J. (1970). *Applied Factor Analysis*. Evanston: Northwestern University.
- Ruttkowski, W. V. (1974). *Typologien und Schichtenlehren. Bibliographie des internationalen Schrifttums bis 1970*. Amsterdam: Rodopi.
- Sabourin, S., Lussier, Y., Laplante, B., & Wright, J. (1990). Unidimensional and multidimensional models of dyadic adjustment: A hierarchical reconciliation. *Psychological Assessment. A Journal of Consulting and Clinical Psychology*, 2, 333-337.
- Saunders, D. R. (1962). *Transvarimax: Some properties of the ratiomax and equamax criteria for blind orthogonal rotation*, American Psychological Association Meeting.
- Schaper, N., & Baumgart, C. (2002). TKI-Teamklima-Inventar. *Zeitschrift für Personalpsychologie*, 1(3), 54-158.
- Schmid, J. & Leiman, J. M., (1957). The development of hierarchical factor solutions. *Psychometrika*, 22, 53-61.
- Schmitt, M. (2000). Vorschlag zur Vereinfachung des Beck-Depressions-Inventars (BDI). *Diagnostica*, 46(1), 38-46.
- Schönemann, P. H. (1981). Measurement, scaling, and factor analysis. In I. Borg (Ed.), *Multidimensional data representation: When and why*. Ann Arbor: Mathesis.
- Schönemann, P. H. (1994). Book Review: Louis Guttman on Theory and Methodology: Selected writings. *Applied Psychological Measurement*, 18 (3), 293-297.
- Schönemann, P. H. & Borg, I. (1996). Von der Faktorenanalyse zu den Strukturgleichungsmodellen. In Erdfelder, E., Mausfeld, R., Meiser, T., and Rudinger, G. (Eds): *Quantitative Psychologische Methoden – Ein Handbuch in Schlüsselbegriffen*, 241-252.
- Schönemann, P. H. & Steiger, J. H. (1976). Regression component analysis. *British Journal of Statistical Psychology*, 29, 175-189.
- Schönemann, P. H. & Wang, M. M. (1972). Some new results on factor indeterminacy. *Psychometrika*, 37, 61-91.
- Sechrest, L. B. (2002). Paper without available title presented at the 11<sup>th</sup> European Conference on Personality. Universität Jena, 21.7.-25.7.2002.
- Singer, W. (2003). *Ein neues Menschenbild? Gespräche über Hirnforschung*. Frankfurt a. M.: Suhrkamp
- Smith, P. T. & Jones, K. F. (1975). Some hierarchical scaling methods for confusion matrix analysis. II. Applications to large matrices. *British Journal of Mathematical and Statistical Psychology*, 28, 30-45.
- Snyder, M., & Gangestad, S. (1986). On the nature of self-monitoring: Matters of assessment, matters of validity. *Journal of Personality and Social Psychology*, 51, 125-139.
- Sponsel, R. (1998). Axiome und Konstruktionsprinzipien Differentieller Psychologie der Persönlichkeit in der Allgemeinen und Integrativen Psychodiagnostik, Psychologie, Psychopathologie und Psychotherapie. *Internet Publikation für Allgemeine und Integrative Psychotherapie IP-GIPT. Erlangen: http://www.sgipt.org/gipt/diffpsy/gipt0.htm*.
- Steiger, J. H. (1994). Factor analysis in the 1980's and the 1990's: Some old debates and some new developments. In I. Borg & P. P. Mohler (Eds.) *Trends and perspectives in empirical social research* (pp. 201-223). New York: De Gruyter.
- Steiger, J. H. & Schönemann, P.H. (1975). A history of factor indeterminacy. Unpublished manuscript. Purdue University.
- Stephenson, W. (1956). Methodology of trait analysis. *The British Journal of Psychology*, 47 (5-18).
- Suppes, P. & Zinnes, J. L. (1963). Basic measurement theory. In: Luce, R. D., Bush, R. R. & Galanter, E. (eds.) *Handbook of Mathematical Psychology. Vol. I*. New York: Wiley.
- Thalbourne, M. A. (1998). Transliminality: Further correlates and a short measure. *The Journal of the American Society for Psychical Research*, 92, 402-419.
- Thompson, J. W. (1962). Meaningful and unmeaningful rotation of factors. *Psychological Bulletin*, 59 (3), 211-225.
- Thompson, J. W. (1963). Bi-polar and uni-directional scales. *British Journal of Psychology*, 54, 15-24.
- Thurstone, L. L. (1934). The Vectors of the Mind. *Psychological Review*, 41, 1-32.
- Thurstone, L. L. (1935). *The Vectors of the Mind*. Chicago: The University of Chicago Press.
- Thurstone, L. L. (1947). *Multiple factor analysis. A development and expansion of The Vectors of the Mind*. Chicago: The University of Chicago Press.
- Tinsley, H. E. A. & Tinsley, D. J. (1987). Uses of factor analysis in Counseling Psychology research. *Journal of Consulting Psychology*, 14, 414-424.
- Überla, K. (1971). *Faktorenanalyse. Eine systematische Einführung für Psychologen, Mediziner, Wirtschafts- und Sozialwissenschaftler*. 2. Auflage, Berlin: Springer. Erste Auflage 1968, letzte 1977.

- Undheim, J. O., & Gustavsson, J.-E. (1987). The hierarchical organization of cognitive abilities. Restoring general intelligence through the use of structural relations (LISREL). *Multivariate Behavioral Research Monographs*, 22, 149-171.
- Velicer, W. F. (1977). An empirical comparison of the similarity of principal component, image, and factor patterns. *Journal of Multivariate Behavioral Research*, 12, 5-22.
- Velicer, W. F. & Jackson, D. N. (1990). Component analysis versus common factor analysis: Some issues in selecting an appropriate procedure. *Multivariate Behavioral Research*, 25(1), 1-28.
- Vittadini, G. (1989). Indeterminacy problem in the LISREL model. *Multivariate Behavioral Research*, 24(4), 397-414.
- Vukovich, A. (1967). Faktorielle Typenbestimmung. Sonderheft: Münchener Symposium über Faktorenanalyse. *Psychologische Beiträge*, 10, 81-86.
- Warburton, F. W. (1963). Analytic methods of factor rotation. *The British Journal of Statistical Psychology*, 16, 165-174.
- Weber, E. (1978). *Einführung in die Faktorenanalyse: mit einem Anhang zur Matrizenrechnung*. – 1. Aufl. 1974.
- Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology*, 37, 395-412.
- Wiley, D. E. (1967). Latent partition analysis. *Psychometrika*, 32(2), 183-193.
- Witte, W. (1974). Untersuchungen zur Behinderung des Denkens durch Anschauung. *Psychologische Beiträge*, 16, 277-287.
- Yates, A. (1987). *Multivariate exploratory data analysis: A perspective on exploratory factor analysis*. New York: State University of New York Press.
- Korrespondenz-Adresse:  
Prof. Dr. em. Suitbert Ertel  
Georg-Elias-Müller-Institut für Psychologie  
Georg-August-Universität Göttingen  
Gossler Straße 14  
D-37073 Göttingen  
sertel@uni-goettingen.de

# Hinweise für Autorinnen und Autoren

## Beitragsarten:

Das Journal *Psychologie des Alltagshandelns / Psychology of Everyday Activity* veröffentlicht die folgenden Formate:

### *Originalarbeiten*

Originalarbeiten beinhalten empirische Forschungsergebnisse, Methodenentwicklungen oder theoretische Beiträge (max. 40.000 Zeichen).

### *Sammelreferate (Reviews)*

Sammelreferate fassen den aktuellen Forschungsstand zu einem bestimmten Thema zusammen (max. 50.000 Zeichen).

### *Kurzberichte (Research Notes)*

Kurzberichte stellen Pilotstudien oder Replikationen dar oder informieren über anlaufende Forschungsprojekte oder -programme und deren erste Ergebnisse (max. 20.000 Zeichen).

### *Diskussion*

Diskussionsbeiträge beziehen argumentativ zu einem diskussionswürdigen Thema mit Bezug zur psychologischen Forschung oder Praxis Stellung (max. 20.000 Zeichen).

### *Buchbesprechungen*

Rezensionen zu einem neu erschienen Werk aus der psychologischen Forschung oder Praxis (max. 8.000 Zeichen).

## Einreichung von Manuskripten:

Manuskripte sind generell nach den Richtlinien zur Manuskriptgestaltung der Deutschen Gesellschaft für Psychologie oder der American Psychological Association zu gestalten. Eine Ausnahme bildet jedoch die Kursivsetzung. Im Unterschied zu den DGPs-Richtlinien bzw. APA-Richtlinien bitten wir die Autorinnen und Autoren Textstellen, die in der Druckfassung kursiv erscheinen sollen, bereits im Manuskript kursiv zu setzen. Die Kurzzusammenfassung (Abstract) sollte 1000 Zeichen nicht überschreiten. Beiträge können in deutscher oder englischer Sprache verfasst werden. Bei deutschsprachigen Beiträgen ist neben der deutschsprachigen Kurzzusammenfassung auch ein englischsprachiges Abstract einzureichen. Im Anschluss an das Abstract sind maximal sechs Schlüsselwörter (Keywords) aufzulisten. Bei deutschsprachigen Beiträgen sind die Schlüsselwörter sowohl in deutscher als auch in englischer Sprache anzugeben.

Die Namen der Autorinnen und Autoren sollten nur auf dem Titelblatt erscheinen, um eine anonyme Begutachtung zu gewährleisten.

Tabellen und Abbildungen sind jeweils gesondert und nummeriert dem Manuskript am Manuskriptende beizufügen. Im Manuskripttext ist die Stelle zu kennzeichnen, an der die jeweilige Tabelle oder Abbildung gewünscht wird.

Beiträge sind bitte per E-Mail an die folgende Adresse einzureichen:

*Journal-Psychologie-des-Alltagshandelns@uibk.ac.at.*

Es wird sowohl eine Version im pdf-Format als auch eine Version in einem gängigen Textverarbeitungsprogramm (z.B. Word) benötigt.

### *Belegexemplare*

Erstautorinnen und -autoren erhalten jeweils ein Heft des Journals als Belegexemplar sowie eine pdf-Kopie ihres Beitrages.