

2

Vol. 2

Psychologie  
des Alltagshandelns

---

*Psychology  
of Everyday Activity*

**Editor**  
P. Sachse

JOURNAL

**JOURNAL**  
**Psychologie des Alltagshandelns**  
*Psychology of Everyday Activity*  
Vol. 2 / No. 2, October 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

Recent Developments in Reflexivity Research: A Review .....	2
<i>Pascale S. Widmer, Michaéla C. Schippers, Michael A. West</i>	
Seelische Gesundheit und Entwicklung von Motivation und Leistungsfähigkeit als Funktion der Arbeitsaufgabengestaltung .....	12
<i>Bärbel Bergmann, Ulrike Pietrzyk, Falk Richter</i>	
Vom Schatten herrschender Verhältnisse oder: Was fördert Organisationales Commitment? .....	22
<i>Birgit E. Schmid</i>	
Komplexe Strukturen aufspüren: Faktorenanalyse mit Variminrotation .....	35
<i>Suibert Ertel</i>	

# Recent Developments in Reflexivity Research: A Review

Pascale S. Widmer\*, Michaéla C. Schippers\*\*, Michael A. West\*\*\*

---

\* University of Bern, Switzerland

\*\* RSM Erasmus University, Rotterdam, Netherlands

\*\*\* Aston Business School, Birmingham, United Kingdom

## ABSTRACT

Team reflexivity, or the extent to which teams reflect upon and modify their functioning, has attracted much recent research attention. In the current paper, we identify several predictors as well as consequences of reflexivity by reviewing the last decade of literature on team reflexivity. It is observed that team characteristics such as trust and psychological safety among group members, a shared vision, and diversity as well as leadership style of the team's supervisor influence the level of reflexivity. In addition, team reflexivity is related to a team's output in terms of innovation, effectiveness, and creativity. Explanations for these effects are discussed and a model including all current findings is presented.

## Keywords

reflexivity – review – teams – effectiveness – innovation

## 1 Introduction

Teams are often important building blocks of successful organizations, especially for organizations that operate in dynamic environments. Team-based organizations can respond quickly and effectively in the fast-changing environments, they enable organizations to develop and deliver products and services quickly and cost effectively, enable organizations to learn, and promote positive outcomes such as creativity, innovation, etc. (Cohen & Bailey, 1997; West, 2004). In order to work effectively, however, it is important for teams to coordinate actions of team members. West and his colleagues identified reflexivity as an important determinant of effectiveness of complex decision-making teams (West, 1996; West, Garrod, & Carletta, 1997). West stated that when members collectively „reflect upon their objectives, strategies, processes and wider environments; plan to adapt these aspects of their task functional worlds and make changes accordingly, teams will be more effective“ (West, 2000, p. 151).

An important aim of this paper is therefore to give an overview of recent developments regarding reflexivity. We will start by defining and explaining the theoretical underpinnings of the construct, review the past decade of empirical findings and conclude with a model of what has been found so far.

### *Definition of Reflexivity*

Reflexivity as a group level construct is defined as „the extent to which group members overtly reflect upon, and communicate about the group's objectives, strategies (e.g., decision-making) and processes (e.g., communication), and adapt them to current or anticipated circumstances“ (West, 2000, p. 296). Swift and West (1998) distinguished between task and social reflexivity. Task reflexivity refers to reflexive behaviors with respect to the team's task. Social reflexivity refers to the extent teams reflect on the social processes within the teams. In his later work (e.g., 2000), West no longer distinguished between these two forms of reflexivity. Also, most research to date focuses on task reflexivity. In comparison to reflexive teams, non-reflexive teams show little awareness of the team objectives, strategies and the environment in which they operate. Such teams are inclined to be reactive rather than proactive and react defensively in case of environmental threat. Reflexive teams plan in more detail, pay more attention to long-term consequences and have a larger inventory of environmental cues to which they respond (West, 2000).

Reflexivity is thought of as an iterative process consisting of three components: reflection, planning and action / adaption. The relation between these three components is depicted in Figure 1. Although figure 1 suggests that the different aspects of reflexivity can be clearly discerned, they are probably less sequential and more highly interrelated than depicted.

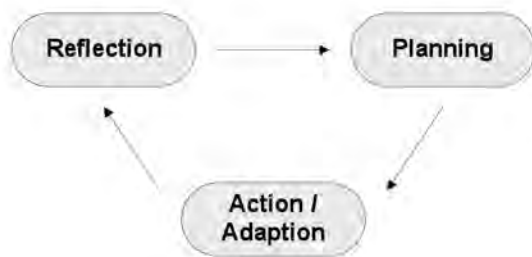


Figure 1: Aspects of reflexivity. Note: From West (2000)

#### Reflection

Reflection refers to the consideration of work-related issues. In the literature on organizational learning, reflection is mentioned as an important factor in learning. Tjosvold (1991), for example, regarded reflection as an important instrument for recognizing how certain present ways of operating can be obsolete, because of environmental changes, thus, reflection is crucial in learning from experience. According to West (2000, p. 4) „reflection includes behaviors such as questioning, planning, exploratory learning, analysis, diverse exploration, making use of knowledge explicitly, planfulness, learning at a meta-level, reviewing past events with self-awareness, and coming to terms over time with a new awareness“. Thus, the variety of behaviors described demonstrates that reflection can vary according to the depth of awareness of inquiry.

#### Planning

Reflection as such does not lead to changes, adaptation needs to occur as well. Planning is seen as the bridge between reflection and action or adaptation (Miller, Galanter, & Pribram, 1960; West, 1996). In the planning phase, goals are presented and ways to achieve these goals are planned. These plans will then be implemented in the action phase. According to Weingart (1992), planning not only takes place before task execution, but often takes place during task execution as well. In that case, only the first actions are usually planned (Faludi, 1973). During task execution subsequent plans are developed and shaped by task feedback (Freidmann, 1966). Research by Weingart (1992) showed that this way of planning occurs relatively more than preplanning. According to West (2000), the planning stage is important in developing implementation intentions and is in going beyond the stage of

reflection and towards action. Research by Sonnentag (1998) showed that detailed planning during task execution is important in coordinating the actions of interdependent team members. Research by Gevers, vanEerde, and Rutte (2001) among student project teams indicates that planning in the orientation phase did not contribute to progress, whereas planning in the execution phase was highly related to the teams' progress. So planning seems to be important for teams, but especially during the execution phase. Thus, planning can be seen as a bridge between reflection and action, indeed.

#### Action

According to West (2000), action refers to goal-directed behaviors relevant to achieving the desired changes in team objectives, strategies, processes, organizations or environments identified by the team during the stage of reflection. Action is seen as an important aspect in most learning cycles and as a way to test assumptions by practical experience. In action theory (Hacker, 1985, 2003), action is even regarded as ‚the core of work psychology‘ per se (cf. Frese & Zapf, 1994). West (2000) asserted that action can be measured on four dimensions: Magnitude, novelty, radicalness, and effectiveness. The first three dimensions describe the innovativeness of the actions, whereas the fourth dimension is related to the performance of the team. The actions carried out by the team members lead to new information, which can lead to further reflection, planning and action as an iterative and ongoing process (West, 2000).

#### Recent Developments

In practice, it is difficult to discern the three components, reflection, planning and action/adaptation, since teams may easily jump between the components, and phases may be not distinguishable, that is they are probably less sequential and more highly interrelated than depicted. When researching the topic of reflexivity, an important development is that researchers tend to focus on one of the aspects. All three components are represented by separate research lines, although the most essential part of reflexivity, but least researched so far, is reflection. Recent research tends to focus more on the reflection aspect of reflexivity (e.g., Schippers, Den Hartog, & Koopman, 2007).

In the following, we present a model of findings so far.

## 2 Antecedents of Reflexivity

It has been noted that teams generally do not engage in reflexive behavior spontaneously (Schippers, 2003). Several factors have been identified to enhance or diminish the possibility of reflective activities in teams. Following, we will present findings concerning the antecedents of reflexivity. In doing so, we will address findings about team characteristics first, followed by findings about leadership style.

### *Team characteristics*

Several characteristics of teams have been found to have an influence on team reflexivity. The most important ones are trust, psychological safety, shared vision, and diversity. Empirical evidence concerning these characteristics will be discussed in the following.

### *Trust and psychological safety*

The concepts of „trust“ and „psychological safety“ similarly refer to the climate within a team in regard to the expectancy of cooperative or non-harming behavior of other team members (Kramer & Tyler, 1996). Such a climate seems an important condition for reflexive behavior. A higher level of reflexivity in a team is associated with greater probability for detecting discrepancies between actual and desired conditions. This might evoke uncertainty or even anxiety in individuals. Furthermore, by openly revealing feelings and thoughts related to team processes, team members make themselves vulnerable (Cunliffe & Easterby-Smith, 2004). It follows that an important prerequisite for the interpersonal risk taking involved in reflexive behavior is psychological safety among team members. Members who trust each other will not be afraid of speaking up freely because they do not worry that other team members will take advantage of them. In a study among 51 work teams, Edmondson (1999) found that, indeed, psychological safety was the most important predictor of team learning. With team learning, she referred to „an ongoing process of reflection and action, characterized by asking questions, seeking feedback, experimenting, reflecting on results, and discussing errors or unexpected outcomes of actions“ (Edmondson, 1999, p. 353). This concept is very close to the concept of reflexivity and demonstrates the relevance of psychological safety to promote an open and constructive atmosphere for reflective group discussions. Similarly, Hoegl and Parboteeah (2006) found that team members' social skills that support the forming of trust, e.g. comprehension or sensitivity, positively predicted reflexivity. Schippers (2003) tested the direct relationship between trust and reflexivity in several studies and found high correlations. Trust also seems to have an important mediating function. In a study among 60 teams, Schippers (2003) demon-

strated that the more inspirational leadership lead to trust (and a shared vision), the more it lead to evaluation and learning as well as better error management. Furthermore, in her study among 59 school management teams (Schippers, 2003), she found a mediating role for trust between attachment style and reflexivity demonstrating that teams are higher on trust when their members score lower on anxious and on avoidant attachment which in turn lead to increased evaluation/learning.

Trust might be particularly important when conflicts arise between team members. Conflicts, on one hand, create moments of potential awareness during which the group can step back from involvement in task related issues and take the opportunity to attend to team, organizational or environmental issues. On the other hand, conflicts – especially if they concern interpersonal issues – might also negatively affect the team climate. West (2000) expects that conflicts are likely to induce reflective behavior. However, Tjosvold, Hui, and Yu (2003) found that not the conflicts themselves, but the way they were managed, i.e. cooperatively, lead to more reflexive behavior. Jehn (1995) found that in groups with norms concerning the acceptance of conflicts in a group, promoting an open and constructive atmosphere for group discussion, conflicts were indeed beneficial. Similarly, Gurtner, Tschan, and Bogenstätter (2009) found no direct effect of social stressors on reflexivity but a moderating effect of task interdependence, indicating that in teams with high task interdependence, conflicts did favourably affect reflexivity. Surprisingly, however, the opposite was true for team stability; members of stable teams reported less reflexive behaviour, especially in the presence of social stressors. Thus, it seems that contrary to primary assumptions (e.g., West, 2000), conflicts do not induce reflexivity per se but can be a trigger for reflexive behavior if the necessary circumstances, i.e. high task interdependence or a climate of cooperative conflict-management, are given. Also, the role of conflicts in relation to reflexivity depends on team characteristics. The most crucial team characteristics are feelings of safety within the team and trust among its members. Task related conflicts within a psychological safe environment will lead to innovation by encouraging debate and consideration of alternative interpretations (West & Richter, 2008). Importantly, conflict acceptance or the way conflicts are dealt with plays a crucial role in either enhancing or inhibiting reflexivity in a team.

### *Shared Vision*

A vision is „an idea of a valued outcome, a higher order goal, which is a motivating force for a work group“ (West, 1990). For organizations to be innovative and to reach their long-term goals with the contribution of many teams and all team members, it is important that

this vision is shared among all employees. With a shared vision, there is a greater possibility for team members to implement an innovation that goes with this vision (Agrell & Gustafson, 1994). Thus, a shared vision creates a commitment to the team, a common identity, and, what is most important concerning reflexivity, fosters risk taking and experimentation (Senge, 1990). We therefore assume that teams with a shared vision are more likely to engage in reflective behavior than teams without a shared vision. This assumption has indeed found some support. Schippers (2004) found high correlations between shared vision and reflexivity in a study among 59 school management teams. Importantly, her results showed that inspirational leaders, i.e. leaders who exert behaviors such as acting as a role model for subordinates or communicating a vision (DenHartog, VanMuijen, & Koopman, 1997), manage to create such a shared vision among team members, highlighting the importance of the team leader's behavior on team processes.<sup>1</sup> A shared vision was also shown to mediate the positive relationship between transformational leadership and reflexivity (Schippers, Den Hartog, Koopman, & van Knippenberg, 2008). A study among 60 teams from different organization supported this perspective and also showed that having a shared understanding of the team goal may counteract possible negative consequences of diversity (Schippers, Den Hartog, Koopman, & Wienk, 2005). Thus, one can conclude that creating a shared vision among team members, which can be encouraged by inspirational or transformational leadership behavior, leads to increased reflective behavior.

### *Diversity*

In today's time of multicultural collaboration and global exchange, managing diverse groups has become a difficult but important challenge for organisations. Previous research has found evidence of both, stimulating and inhibiting effects of diversity in a team (van Knippenberg, De Dreu, & Homan, 2004; van Knippenberg & Schippers, 2007). The social categorization perspective proposes that overall group performance will be higher when groups are homogeneous rather than heterogeneous (Jehn, Northcraft, & Neale, 1999) as there will be fewer relational conflicts (Jehn et al., 1999; Pelled, Eisenhardt, & Xin, 1999) or higher member commitment (Riordan & McFarlaneShore, 1997). Besides, so-called deep-level diversity<sup>2</sup> is said to lead to communication difficulties and therefore poor team functioning (Williams & O'Reilly, 1998). On the other hand, according to information and decision-making theory diversity can have a direct positive impact on

team performance (Hinsz, Tindale, & Vollrath, 1997). This approach suggests that diverse groups possess a broader range of knowledge, skills, and abilities and have to integrate different opinions and perspectives on the task which will lead to more creative and innovative ideas (Ancona & Caldwell, 1992; van Knippenberg et al., 2004). This „value-in-diversity“ approach has also found broad empirical support (Cox, Lobel, & McLeod, 1991; Hambrick, Cho, & Chen, 1996; Jehn et al., 1999). One thought of diversity research is particularly interesting for the concept of reflexivity; members with different views will bring usefully differing perspectives on issues to the group (Williams & O'Reilly, 1998) which stimulates consideration of non-obvious alternatives (González-Romá & West, 2004) and therefore fosters communication and, in turn, reflective behavior. Some studies provide supporting findings for this assumption. West, Utsch, Borrill, and Dawson (2002) found a positive relationship between knowledge diversity and reflexivity in a study among 100 health care teams. They also showed that reflexivity moderates the relationship between knowledge diversity and innovation, demonstrating that in teams with high diversity, in which there is no conformity per se, it is more necessary to communicate and behave reflective to collectively achieve a goal. A similar finding is presented by Fay, Borrill, Amir, Haward, and West (2006); with two independent samples of 66 and 95 teams they found a moderating effect of team processes, such as reflexivity and shared vision, on the relationship between task-related (professional) diversity and innovation quality. Thus, professional diversity was positively related to the quality of team innovation if teams had good team processes, i.e. high reflexivity. Providing further support, Gibson and Vermeulen (2005) found positive relations between team heterogeneity in terms of demographic variables (cf. diversity) and team learning behavior (cf. reflexivity) among 156 teams and Van der Vegt and Bunderson (2005) found that team learning behavior partly mediated the relationship between expertise diversity and team performance. The positive relation between professional diversity in a team and reflexivity was also found by Haward et al. (2003) in a study on 72 breast cancer teams. Schippers et al. (2005) had a closer look on the direct relationship between diversity and reflexivity and their results in a study among 54 work teams from 13 different organizations reveal that the relationship is moderated by outcome interdependence and group longevity; groups that were highly outcome-interdependent and diverse were more reflective than those that were less outcome-interdependent. Having

1 The role of leadership behavior for reflexivity will be discussed in more detail in the section „leadership style“.

2 Deep-level, i.e. psychological diversity as opposed to surface-level, i.e. demographic diversity (Harrison, Price, Gavin & Florey, 2002).

a common goal might therefore be helpful for diverse teams to behave reflective, or, put differently, reflectivity is more required to understand the goal and work towards it together if the team is heterogeneous. Less diverse teams, on the other hand, do not need to reflect on a shared goal because they have a common understanding of it and each other per se, however, if their objectives are not clear, they might engage more in reflective behavior.

Overall, although the results about the direct impact of diversity on reflexivity are somewhat ambiguous, it seems that there is a relation between these two factors. Under certain conditions which require further examination, diversity can affect reflexivity. Reflexivity, in turn, seems to be an important moderator between diversity and team outcomes. This shows that the direction of the relationship between reflexivity and diversity is not clear yet.

#### *Leadership style*

The patterns and styles of leadership are a very important factor to enhance or diminish reflexivity in a team. Although certain incidents might evoke situations in which the teams steps back from its original task and discusses e.g. team processes (e.g. when errors or team conflicts occur) the impact of these events on reflexivity is dependent on leadership style. A team's leader has to react to those incidents in an adequate way to achieve this positive effect. For instance, if the leader does not discuss errors within the team or only solves conflicts at short notice without considering long-term consequences, he might rather restrain the team from acting reflective than facilitate reflecting behavior. What is more, the leader has to create the conditions for exploration, experimentation and risk-taking. The team leader can also influence team reflexivity more directly by encouraging the team to reflect on their objectives, strategies, and processes and stimulating their communication. Several studies have supported the idea of an impact of leadership behavior on reflexivity. In a study of 50 Research and Development (R&D) teams, Hirst, Mann, Bain, Pirola-Merlo, and Richter (2004) showed that facilitative leader behavior – i.e. promoting respect and positive relationship between team members, productive conflict resolution, and open expression of ideas and opinions – was positively associated with team reflexivity, which in turn affected customer ratings of team performance. Similarly, in a study of 136 primary health care teams, Somech (2006) found participative leadership style – i.e. fostering joint decision making or shared influence by superior and employees – was positively related with team reflection, though, only in high functionally heterogeneous teams. On the other hand, directive leadership style – i.e. providing the team with a framework for decision making and action in alignment with the superior's

vision – only increased team reflection when teams were low functionally heterogeneous. Hirst and Mann (2004) analyzed the relations between leadership role performance, team boundary spanning, communication safety, team reflexivity, and task communication developing a five-factor model of team communication which they tested in a one-year longitudinal study with 56 R&D teams. They found that team reflexivity mediated the link between innovative leadership and team performance whilst a cross-lagged analysis demonstrated that innovative leadership behavior led to team reflexivity. However, they did not dwell upon the question of why leadership does affect reflexivity. To fill this gap, Schippers et al. (2008) conducted a field study among 32 work teams looking at the processes through which leadership affected reflexivity in more detail. They showed that transformational leadership – a style of leadership that transforms followers by stimulating them to go beyond self-interest through altering their morale, values and ideals, and motivating them to perform above expectations – does enhance reflexivity, indeed, therefore supporting the findings from Hirst and Mann (2004). Moreover, in their study the impact of transformational leadership was fully mediated by a shared vision. In her survey among 60 teams, Schippers (2003) also showed this mediating effect not only of shared vision but also of trust in the team. The model she tested suggests that the stronger the effects of inspirational leadership on these two factors, the more it leads to better discussion processes, better error management as well as more evaluation and learning. These, in turn, are the three factors of reflection that she discerned in her model. Therefore, results again support the relationship between leadership and reflexivity.

#### *Other Influences*

Several other factors have been revealed to have an impact on reflexivity. For instance, Haward et al. (2003), in a study among 72 breast cancer teams, found a positive relationship between number of leaders and reflexivity, indicating that a shared leadership style combined with democratic decision making worked best for clinical decision making. Another factor that seems to foster reflexivity is *cooperation*. Given that cooperation, or the desire to cooperate, leads to or rather requires constructive controversy (Tjosvold, Wong, Nibler, & Pounder, 2002), it is a reasonable assumption that it, in turn, provokes reflexivity. Indeed, results from studies among 100 Chinese teams suggest that cooperative conflict management (Tjosvold et al., 2003) and cooperative (but not competitive) goals (Tjosvold, Tang, & West, 2004) promote team reflexivity. Furthermore, several recent studies point to the importance of *knowledge and skills* of team members in terms of ability to interact with each other (i.e. so-



cial skills; Hoegl & Parboteeah, 2006) and in terms of ability to structure and control projects (i.e. project management skills; Hoegl & Parboteeah, 2006; Lee, 2008). Notably, team cohesiveness was a significant moderator indicating that teams with higher levels of cohesiveness and high levels of management skills as well as higher levels of existing knowledge tend to achieve higher levels of team reflexivity (Lee, 2008). Positive relationships have also been found between reflexivity and *personality* (e.g. proactive personality; Schippers et al., 2007).

Most recently, the question arose as to when, i.e. in what situations reflexivity is most useful and whether it is always useful. Similarly to the assumption that conflicts might evoke reflexivity, it can be argued that reflexivity is especially beneficial if circumstances are inauspicious, i.e. if improvements are necessary. Some first results provide support for this reasoning: In a study among 98 primary health care teams (Dawson, Schippers, & West, 2009), reflexivity showed the strongest effect on innovation when the quality of work premises was low. Schippers and Homan (2009) examined this moderating role of reflexivity among 75 student teams. They found that team reflexivity was most important for final performance (i.e. grade of bachelor thesis) after initial detrimental performance (i.e. grade of research proposal). Thus, reflexivity might be important for teams that have detrimental initial performance to improve their performance, as well as for teams that face adverse working conditions such as high demands.

### 3 Consequences of Reflexivity

First indications for the proposition that reflection will predict team productivity and effectiveness can be found broadly in both social and organizational psychology (West, 2000; West et al., 1997). More recently, however, there is increasing substantial evidence which support the direct impact of reflexivity on productivity and related constructs. In this paper, we will focus on studies that examine this direct relationship.

#### *Performance, Effectiveness, and Innovation*

Carter and West (1998) were one of the first to study direct relationship between reflexivity and productivity. They conducted a longitudinal study over 18 months among 19 BBC TV production teams and measured reflexivity, team size and team climate. They found that reflexivity was a significant predictor of senior managers' ratings of the effectiveness and creativity of the programs the team produced. Reflexivity also explained more of the variance than did team climate for innovation. Several studies support their results; in a study among 22 student project groups, Gevers et

al. (2001) found that reflexivity was positively related to the progress of the group in the execution phase of the project. It also played a role in making up arrears; high reflexive teams made up arrears, while low reflexive teams seemed not to be able to do this. In studies among 100 teams in China, task reflexivity, promoted by cooperative conflict management and cooperative goals, was shown to result in team performance as indicated by supervisors (Tjosvold et al., 2005) as well as in team innovation as rated by managers (Tjosvold et al., 2004). Lee (2008) reports comparable results from a study among 132 members of R&D groups in Taiwan: reflexivity had a significant influence on both product innovativeness and on new product performance. Research by De Dreu (2002) among 32 organizational teams performing complex, ill-defined tasks showed a moderating effect of reflexivity on the relationship between minority dissent and team innovation and effectiveness. Results showed more innovation and greater effectiveness in teams with low levels of minority dissent, but only when there was a high level of team reflexivity. In a more recent study, De Dreu (2007) found positive direct effects of reflexivity on team effectiveness as rated by supervisors and on learning. Moreover, reflexivity had a moderating function insofar as it was a necessary condition to foster positive effects of team interdependence: Outcome interdependent teams engaged in more information sharing, learned more, and had higher levels of team effectiveness, but only if task reflexivity was high. Hoegl and Parboteeah (2006) conducted a study among 145 software development teams and found further proof of the positive relationship between reflexivity and effectiveness as rated by team members. Surprisingly, however, in their study team reflexivity was not related with higher efficiency as measured by self-rating questionnaires. They explain this unexpected result by the fact that engaging in reflexive actions involves additional time and other costs (such as training of reflexive behavior, altering work strategies etc.). The thought that reflexivity might have a downside in terms of resource consumption is interesting, indeed, and points to the fact that reflexivity might be more helpful in certain circumstances, such as detrimental performance and/or adverse working conditions, as discussed before. The predominance of studies highlighting the positive effect of reflexivity on performance, however, indicate that this downside is usually offset by its gains in terms of effectiveness benefits.

### 4 Enhancing Reflexivity

Taken the strong empirical evidence for the positive effects of reflexivity and considering the fact that teams are often not spontaneously reflexive (Schippers, 2005),

it might be important to know whether and how reflexivity in teams can be initiated and enhanced. Indeed, lately, there have been promising attempts to intentionally induce reflexivity in order to foster productivity. Gurtner, Tschan, Semmer, and Nägele (2007) tested the effects of an intervention to provoke reflexivity within hierarchically structured groups of three that have been given a complex and dynamic task. Reflexivity was implemented by instructions to reflect on the task either individually or communicating with the group. Their results confirm that reflexivity instructions are effective in initiating the postulated effect of reflexivity on performance with groups in the reflexivity condition clearly performing better. Surprisingly, however, individual reflexivity was superior to group reflexivity. Additional analysis suggested that group reflexivity increased discussion of strategies that were rather too general to be helpful. Thus, this study emphasizes the importance of focusing on task-specific strategies when implementing team reflexivity. Similarly, Müller, Herbig, and Petrovic (2009) studied 48 students of mechanical engineering work on a product development task in groups of three. They operationalized team reflexivity as the explication of and reflection on implicit team knowledge and found that those groups that have been instructed to collectively explicate their implicit individual knowledge (‘individual explication’) or to stimulate team communication about the task (‘collective explication’) produced qualitatively better and more innovative products. Also, Vashdi, Bamberger, Erez, and Weiss-Meilik (2007) qualitatively analyzed the impact of the implementation of briefing-debriefing sessions as a team-based reflexive learning and found three paths

by which structured team reflexivity enhances team outcomes: (1) by means of problem identification and solution, (2) by questioning taken-for-granted procedures, and (3) by generating and entrenching a team learning culture. Hence, these results have significant practical implications as, given that reflexivity instructions do not necessarily require high effort or much time, they can be regarded as a very promising tool for fostering effective group performance.

To sum up, the results presented overall support the link between team reflexivity and team effectiveness. It has been shown that reflexivity leads to more innovation (Tjosvold et al., 2004), higher team performance (Schippers, 2005; Tjosvold et al., 2003) and fosters team processes (Gevers et al., 2001). These results and especially the promising approaches to intentionally evoke reflexivity in teams (Gurtner et al., 2007; Müller et al., 2009; Vashdi et al., 2007) suggest that reflexivity not only should but can be fostered in teams in order to help them work more effectively.

## 5 Discussion

The purpose of this paper was to provide a review of the last decade of research focused on the rather new concept of reflexivity. The evidence from our review supports the claim that reflexivity can be important to guarantee and foster team functioning. In addition, the evidence reviewed points to the significance of certain environmental characteristics, such as trust in teams or diversity that provide opportunities for or promote reflexive behavior. Finally, we note that some team characteristics, such as team stability or task interdependence may be necessary

to allow these environmental characteristics to lead to more reflexivity in teams. Figure 2 summarizes the findings reviewed in this paper.

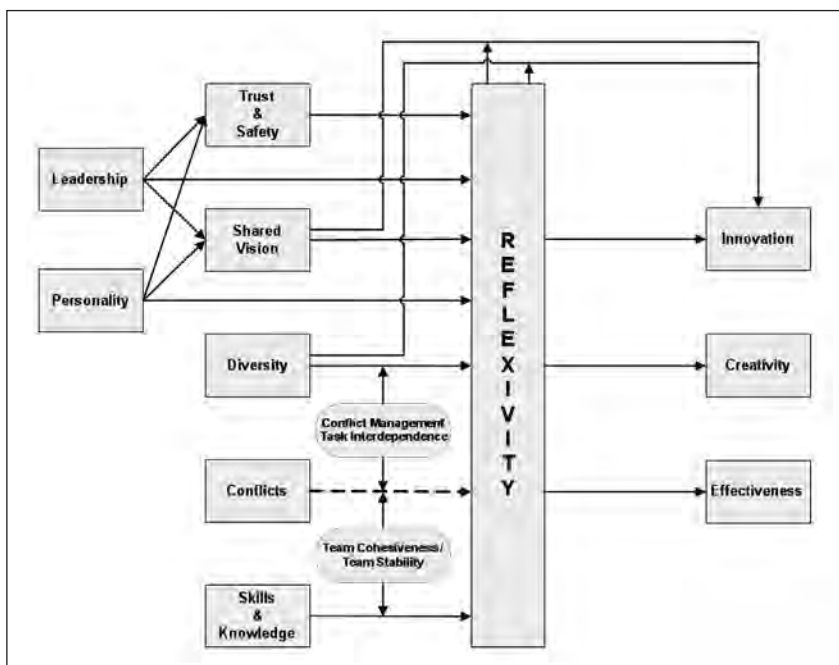


Figure 2:  
Summary of findings

Team working is demanding and complex since it requires team members to coordinate and integrate their efforts towards shared objectives. Teams are formed because their diverse skills are needed in order to accomplish a task that individuals working alone could not or would only do so less effectively or efficiently. However, the experience of team work for many is often characterized by an overwhelming workload leading to a survival mentality. High workload and acceptance of routines and standardized processes are both the enemies of innovation. It is therefore necessary for teams to find the cognitive, social and temporal space to enable them to reflect on the appropriateness of their objectives, the wisdom of their strategies, the efficiency of their processes and the reality of their changing environment. This reflection is likely to reveal the discrepancies in their ways of working which may produce work overload or argue against past routines and, consequently, define the space and direction of appropriate remedial or innovative action. Understanding whether and how reflexivity can enable team effectiveness is of huge importance in an organisational world increasingly dominated by team and inter-team working. The review above offers insights into our understanding in this area in order that our knowledge of how to develop and maintain truly effective team working can continue to be advanced.

## References

- Agrell, A., & Gustafson, R. (1994). The Team Climate Inventory (Tci) and Group Innovation - a Psychometric Test on a Swedish Sample of Work Groups. *Journal of Occupational and Organizational Psychology*, 67, 145-151.
- Ancona, D. G., & Caldwell, D. F. (1992). Demography and Design - Predictors of New Product Team Performance. *Organization Science*, 3(3), 321-341.
- Carter, S. M., & West, M. A. (1998). Reflexivity, effectiveness, and mental health in BBC-TV production teams. *Small Group Research*, 29(5), 585-601.
- Cohen, S. G., & Bailey, D. E. (1997). What makes teams work: Group effectiveness research from the shop floor to the executive suite. *Journal of Management*, 23(3), 239-290.
- Cox, T. H., Lobel, S. A., & McLeod, P. L. (1991). Effects of ethnic group cultural difficulties on cooperative and competitive behavior on a groups task. *Academy of Management Journal*, 34(4), 827-847.
- Cunliffe, A. L., & Easterby-Smith, M. (2004). From reflection to practical reflexivity: Experimental learning as lived experience. In M. Reynolds & R. Vince (Eds.), *Organizing Reflection* (pp. 30-46). Aldershot: Ashgate.
- Dawson, J., Schippers, M. C., & West, M. A. (2009). *When do teams need to innovate? Reflexivity is key when the quality of work premises is low*. Paper presented at the European Congress of Work and Organizational Psychology, Santiago de Compostela, Spain.
- De Dreu, C. K. W. (2002). Team innovation and team effectiveness: The importance of minority dissent and reflexivity. *European Journal of Work and Organizational Psychology*, 11(3), 285-298.
- De Dreu, C. K. W. (2007). Cooperative outcome interdependence, task reflexivity, and team effectiveness: A motivated information processing perspective. *Journal of Applied Psychology*, 92(3), 628-638.
- DenHartog, D. N., VanMuijen, J. J., & Koopman, P. L. (1997). Transactional versus transformational leadership: An analysis of the MLQ. *Journal of Occupational and Organizational Psychology*, 70(1), 19-34.
- Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, 44(2), 350-385.
- Faludi, A. (1975). *Planning theory*. New York: Pergamon Press.
- Fay, D., Borrill, C., Amir, Z., Haward, R., & West, M. A. (2006). Getting the most out of multidisciplinary teams: A multi-sample study of team innovation in health care. *Journal of Occupational and Organizational Psychology*, 79, 555-567.
- Freidmann, J. (1966). The institutional context. In M. Gross (Ed.), *Action under planning* (pp. 31-67). New York: McGraw-Hill.
- Frese, M., & Zapf, D. (1994). Action as the core of work psychology: A German approach. In H. C. Triandis, M. D. Dunette & J. M. Hough (Eds.), *Handbook of Industrial and Organizational Psychology* (Vol. 4, pp. 271-340). Palo Alto, CA: Consulting Psychologists Press.
- Gevers, J. M. P., vanEerde, W., & Rutte, C. G. (2001). Time pressure, potency, and progress in project groups. *European Journal of Work and Organizational Psychology*, 10(2), 205-221.
- Gibson, C., & Vermeulen, F. (2005). A healthy divide: Subgroups as a stimulus for team learning behavior. *Administrative Science Quarterly*, 48(2), 202-239.
- González-Romá, V., & West, M. A. (2004). *Agreeing to Disagree: Climate strength and innovation in work teams*. Unpublished manuscript, Valencia, Spain.
- Gurtner, A., Tschan, F., & Bogenstätter, Y. (2009). *Do social stressors trigger or inhibit task reflexivity? The moderating role of team stability, task interdependence and goal clarity*. Paper presented at the Academy of Management Annual Meeting. Chicago, Illinois.

- Gurtner, A., Tschan, F., Semmer, N. K., & Nägele, C. (2007). Getting groups to develop good strategies: Effects of reflexivity interventions on team process, team performance, and shared mental models. *Organizational Behavior and Human Decision Processes*, *102*(2), 127-142.
- Hacker, W. (1985). Activity: A fruitful concept in industrial psychology. In M. Frese & J. Sabini (Eds.), *Goal directed behavior: The concept of action in psychology* (pp. 262-283). Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Hacker, W. (2003). Action regulation theory: A practical tool for the design of modern work processes? *European Journal of Work and Organizational Psychology*, *12*(2), 105-130.
- Hambrick, D. C., Cho, T. S., & Chen, M.-J. (1996). The influence of top management team heterogeneity on firm's competitive moves. *Administrative Science Quarterly*, *41*(4), 659-684.
- Harrison, D. A., Price, K. H., Gavin, J. H., & Florey, A. T. (2002). Time, teams, and task performance: Changing effects of surface- and deep-level diversity on group functioning. *Academy of Management Journal*, *45*(5), 1029-1045.
- Haward, R., Amir, Z., Borrill, C., Dawson, J., Scully, J., West, M. A., et al. (2003). Breast cancer teams: The impact of constitution, new cancer workload, and methods of operation on their effectiveness. *British Journal of Cancer*, *89*, 15-22.
- Hinsz, V. B., Tindale, R. S., & Vollrath, D. A. (1997). The emerging conceptualization of groups as information processors. *Psychological Bulletin*, *121*(1), 43-64.
- Hirst, G., & Mann, L. (2004). A model of R&D leadership and team communication: The relationship with project performance. *R&D Management*, *34*(2), 147-160.
- Hirst, G., Mann, L., Bain, P., Pirola-Merlo, A., & Richter, A. (2004). Learning to lead: The development and testing of a model of leadership learning. *The Leadership Quarterly*, *15*(3), 311-327.
- Hoegl, M., & Parboteeah, K. P. (2006). Team reflexivity in innovative projects. *R & D Management*, *36*(2), 113-125.
- Jehn, K. A. (1995). A Multimethod Examination of the Benefits and Detriments of Intragroup Conflict. *Administrative Science Quarterly*, *40*(2), 256-282.
- Jehn, K. A., Northcraft, G. B., & Neale, M. A. (1999). Why differences make a difference: A field study of diversity, conflict, and performance in work groups. *Administrative Science Quarterly*, *44*(4), 741-763.
- Kramer, R. M., & Tyler, T. R. (1996). *Trust in organizations*. Thousand Oaks, CA: SAGE.
- Lee, L. T. (2008). The effects of team reflexivity and innovativeness on new product development performance. *Industrial Management and Data Systems*, *108*(4), 548-569.
- Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. London: Holt.
- Müller, A., Herbig, B., & Petrovic, K. (2009). The explication of implicit team knowledge and its supporting effect on team processes and technical innovations: An action regulation perspective on team reflexivity. *Small Group Research*, *40*(1), 28-51.
- Pelled, L. H., Eisenhardt, K. M., & Xin, K. R. (1999). Exploring the black box: An analysis of work group diversity, conflict, and performance. *Administrative Science Quarterly*, *44*(1), 1-28.
- Riordan, C. M., & McFarlaneShore, L. (1997). Demographic diversity and employee attitudes: An empirical examination of relational demography within work units. *Journal of Applied Psychology*, *82*(3), 542-558.
- Schippers, M. C. (2003). *Reflexivity in teams*. Amsterdam: Vrije Universiteit Amsterdam.
- Schippers, M. C. (2004). *Learning to learn at school: Reflexivity, team composition and school performance*. Paper presented at the Academy of Management Meeting, New Orleans, Louisiana.
- Schippers, M. C., Den Hartog, D. N., & Koopman, P. L. (2007). Reflexivity in teams: A measure and correlates. *Applied Psychology-An International Review*, *56*(2), 189-211.
- Schippers, M. C., Den Hartog, D. N., Koopman, P. L., & van Knippenberg, D. (2008). The role of transformational leadership in enhancing team reflexivity. *Human Relations*, *61*(11), 1593-1616.
- Schippers, M. C., Den Hartog, D. N., Koopman, P. L., & Wienk, J. A. (2005). Diversity and team outcomes: the moderating effects of outcome interdependence and group longevity and the mediating effect of reflexivity. *Journal of Organizational Behavior*, *24*(6), 779-802.
- Schippers, M. C., & Homan, A. C. (2009). *Breaking the negative performance spiral: The role of team reflexivity and team learning*. Proceedings of the Sixty-Eighth Annual Meeting of the Academy of Management (online program).
- Senge, P. M. (1990). *The fifth discipline: The art and practice of the learning organization*. New York: McGraw-Hill.
- Somech, A. (2006). The effects of leadership style and team process on performance and innovation in functionally heterogeneous teams. *Journal of Management*, *32*(1), 152-157.
- Sonntag, S. (1998). Expertise in professional software design: A process study. *Journal of Applied Psychology*, *83*(5), 703-715.

- Swift, T. A., & West, M. A. (1998). *Reflexivity and group processes: Research and practice*. Sheffield: The ESRC Center for Organization and Innovation.
- Tjosvold, D. (1991). *Team organization. An enduring competitive advantage*. New York: John Wiley and Sons.
- Tjosvold, D., Hui, C., & Yu, Z. (2003). Conflict management and task reflexivity for team in-role and extra-role performance in China. *International Journal of Conflict Management*, 14(2), 141-163.
- Tjosvold, D., Tang, M. M. L., & West, M. A. (2004). Reflexivity for team innovation in China: The contribution of goal interdependence. *Group & Organizational Management*, 29(5), 540-559.
- Tjosvold, D., Wong, A., Nibler, R., & Pounder, J. S. (2002). Teamwork and controversy in undergraduate management courses in Hong Kong: Can the method reinforce the message? *Swiss Journal of Psychology*, 61(3), 131-138.
- Van der Vegt, G. S., & Bunderson, J. S. (2005). Learning and performance in multidisciplinary teams: The importance of collective team identification. *Academy of Management Journal*, 48(3), 532-547.
- van Knippenberg, D., De Dreu, C. K. W., & Homan, A. C. (2004). Work group diversity and group performance: An integrative model and research agenda. *Journal of Applied Psychology*, 89(6), 1008-1022.
- van Knippenberg, D., & Schippers, M. C. (2007). Work group diversity. *Annual Review of Psychology*, 58, 515-541.
- Vashdi, D. R., Bamberger, P. A., Erez, M., & Weiss-Meilik, A. (2007). Briefing-debriefing: Using a reflexive organizational learning model from the military to enhance the performance of surgical teams. *Human Resource Management*, 46(1), 115-142.
- Weingart, L. R. (1992). Impact of Group Goals, Task Component Complexity, Effort, and Planning on Group-Performance. *Journal of Applied Psychology*, 77(5), 682-695.
- West, M. A. (1990). The social psychology of innovation in groups. In M. A. West & J. L. Farr (Eds.), *Innovation and creativity at work. Psychological and organizational strategies* (pp. 309-334). Chichester: John Wiley & Sons Ltd.
- West, M. A. (1996). Reflexivity and work group effectiveness: A conceptual integration. In M. A. West (Ed.), *Handbook of Work Group Psychology* (pp. 555-579). Chichester: John Wiley & Sons Ltd.
- West, M. A. (2000). Reflexivity, revolution and innovation in work teams. In D. A. Johnson & S. T. Beyerlein (Eds.), *Advances in the interdisciplinary study of work teams: Product development teams* (pp. 1-29). Stamford, Connecticut: JAI Press.
- West, M. A. (2004). *Effective teamwork: Practical lessons from organizational research* (2nd ed.). Oxford: Blackwell Publishing.
- West, M. A., Garrod, S., & Carletta, J. (1997). Group decision-making and effectiveness: Unexplored Boundaries. In C. L. Cooper & S. E. Jackson (Eds.), *Creating tomorrow's organizations* (pp. 293-317). Chichester: Wiley.
- West, M. A., & Richter, A. (2008). Climates and cultures for innovation and creativity at work. In J. Zhou & C. E. Shalley (Eds.), *Handbook of Organizational Creativity* (pp. 211-236). New York: Psychology Press.
- West, M. A., Utsch, A., Borrill, C. S., & Dawson, J. F. (2002). *Diversity, demands and team innovation*. Unpublished manuscript, Birmingham.
- Williams, K. Y., & O'Reilly, C. A. (1998). Demography and diversity in organizations: A review of 40 years of research. In B. Staw & L. Cummings (Eds.), *Research in organizational behavior* (Vol. 20, pp. 77-140). Greenwich, CT: JAI Press.

Correspondence to:  
Pascale Widmer, lic.phil.  
University of Bern  
Department of Work and Organizational Psychology  
Muesmattstrasse 45  
CH-3000 Bern 9  
pascale.widmer@psy.unibe.ch

# Seelische Gesundheit und Entwicklung von Motivation und Leistungsfähigkeit als Funktion der Arbeitsaufgabengestaltung

Bärbel Bergmann, Ulrike Pietrzyk, Falk Richter

Technische Universität Dresden

## ZUSAMMENFASSUNG

Die Sicherung von seelischer Gesundheit, arbeitsorientierter Motivation und Leistungsfähigkeit der Arbeitenden werden als wichtige Ziele von Unternehmen begründet. Mit Bezug auf theoretische Konzepte der Arbeits- und Organisationspsychologie werden Wirkungspfade zwischen gestaltbaren Eigenschaften der Arbeitsaufgaben und den genannten Personenmerkmalen erläutert. Die Klassifikation von Arbeitsaufgaben in die von Peter Warr (1996) unterschiedenen Typen, mit denen längerfristige Zusammenhänge zwischen Eigenschaften von Arbeitsaufgaben und der Leistungsfähigkeit von Arbeitenden beschrieben werden, wird genutzt, um zum einen Fragen nach Niveauunterschieden bei Personenmerkmalen Erwerbstätiger und zum anderen Fragen nach differenzierten Alterstrends abzuleiten. In einer Feldstudie an 385 Erwerbstätigen wird eine Analyse der Arbeitsaufgaben und die Erhebung von Indikatoren arbeitsorientierter Motivation, Leistungsfähigkeit und Gesundheit durchgeführt.

Hoch signifikante Zusammenhänge zwischen Eigenschaften der Arbeitsaufgaben und Personenmerkmalen bestehen. Die Klassifikation der Erwerbstätigen in die Warr-Typologie bestätigt signifikante ungünstige Ausprägungsgrade der Personenmerkmale bei Erwerbstätigen, deren Aufgaben in den Typ 4 mit geringen Anforderungen klassifiziert sind. Die Berechnung von Alterskorrelationen der Personenmerkmale zeigt jedoch, dass diese Kennwerte infolge der hohen Wertestreuung nicht sensibel sind. Mehrheitlich bestehen Nullkorrelationen. Die Studie stützt Hypothesen, die in der Gestaltung lernhaltiger Arbeitsaufgaben eine Stellschraube für das Erhalten und Entwickeln der Beschäftigungsfähigkeit Erwerbstätiger und für eine Gesundheitsprävention sehen.

## Schlüsselwörter

Arbeitsgestaltung – Beanspruchung – Motivation – Psychologie – Wissen

## ABSTRACT

Occupational health, motivation and competence of employees are important for companies. According to theoretical and empirical results of work and organisational psychology effects of job design on employees will be explained. The classification of job activities according to the typology of Peter Warr (1996) is used to deduce questions about differences in levels of personality characteristics. In a field study working tasks of 385 employees were analyzed and a measurement of criteria of motivation, competence and occupational health took place. Highly significant correlations exist between characteristics of working activities and motivation, competence and occupational health. The classification of employees according to the Warr-typology confirms more adversarial levels of motivation, competence and occupational health of employees who belong to type 4. Age correlations of person characteristics are not sensitive as a result of the variance of data. The study confirms the hypothesis that job design is an approach to develop motivation, competence and occupational health of employees.

## Keywords

job design – knowledge – motivation – psychology – stress

## 1 Einführung

Der aus der Globalisierung resultierende steigende Wettbewerbsdruck fordert Unternehmen zu rascheren Veränderungen heraus. Technologien, Produkte, aber auch Kunden und Märkte wechseln. Das effiziente Umgehen mit Veränderungen wird zur entscheidenden Kompetenz von Unternehmen. Für Mitarbeiter resultieren aus beschleunigten Veränderungsprozessen wachsende Lernanforderungen. Für deren Bewältigung sind Weiterbildungsmaßnahmen nur partiell hilfreich, weil diese in der Regel nicht auf den konkreten Lernbedarf zugeschnitten sind. Sehr spezielle unternehmensspezifische Lernanforderungen bestehen. Die Lernziele und Lerninhalte sind oft selbst zu bestimmen. Selbstorganisiertes Lernen ist gefordert. Dies gelingt nur mit hoher Motivation. Aber deutsche Arbeitnehmer haben wenig Spaß an der Arbeit, berichtet das Heft 6 (2006) der Personalführung: „Ein Viertel arbeitet nur des Geldes wegen, gerade einmal 27 Prozent freuen sich jeden Tag auf ihren Job“ (S. 13). Die Entwicklung von Motivation der Mitarbeiter wird deshalb zu einer erstrangigen Aufgabe von Unternehmen. Führungskräfte sind gefordert, mit geeigneten Maßnahmen die Motivationsentwicklung ihrer Mitarbeiter zu unterstützen. Der gestiegene Wettbewerbsdruck führt gleichzeitig zu wachsender Unsicherheit. Neue Risiken und Belastungen entstehen. Auf den Gebieten des Gesundheitsschutzes gibt es einerseits Fortschritte. Aber Arbeitsverdichtungen und Überforderung sind zu beobachten. Prekäre Beschäftigung nimmt zu (Dörre, 2006) und daraus resultiert nicht nur Unsicherheit als Folge der Einschränkung der Zukunftsplanung, weil befristete Arbeit nur einen geringen Planungshorizont ermöglicht und geringfügige Beschäftigung keine dauerhafte Existenzgrundlage sichert. Es kommt hinzu, dass in so genannten flexiblen Beschäftigungsformen Routinearbeit dominiert (Pietrzyk, 2005), so dass die Erwerbstätigen häufig Arbeitsaufgaben unterhalb ihres erworbenen Qualifikationsniveaus ausführen. Damit sind sie aber auch von wichtigen Lern- und Kompetenzentwicklungsprozessen abgekoppelt, was ihre künftige Beschäftigungsfähigkeit im erworbenen Beruf verschlechtert. Es reicht nicht aus, eine Qualifikation zu erwerben. Der Kompetenzerhalt ist an deren Nutzung durch ein Lernen im Prozess der Arbeit bei der Bearbeitung qualifikationsangemessener Aufgaben gebunden. Ungünstigere Zukunftschancen als Folge prekärer Beschäftigung führen zu psychischen Belastungen. Die Zunahme psychischer Erkrankungen bei den Gründen für Arbeitsunfähigkeit ist ein Indikator für veränderte Risiken. Die Zahl der psychisch

bedingten Arbeitsunfähigkeitsfälle stieg zwischen 1994 und 2003 bei den Männern um 81,6 und bei den Frauen um 56,8 Prozent. Jede dritte Frühberentung geht auf das Konto eines psychischen Leidens zurück (Robert-Koch-Institut, 2006). Bei Zeitarbeitsbeschäftigten resultieren gravierende Belastungen aus der Sicht der Betroffenen aus mangelnden Entwicklungsmöglichkeiten, unzureichendem Einkommen und aus der Arbeitsplatzunsicherheit (Technikerkrankenkasse, 2009). Die traditionelle arbeitswissenschaftliche Aufgabe der Erhaltung und Förderung von Gesundheit wandelt sich folglich. Zusätzlich zu den gut dokumentierten körperlichen Erkrankungen und Krankheitsfolgen in Form von Arbeitsunfähigkeit und Frühberentung muss die Beschaffenheit von Indikatoren psychischer Gesundheit sorgfältig analysiert werden. Zusammenhänge ihrer Ausprägungen mit gestaltbaren und damit veränderbaren Merkmalen der Arbeitssituation sind zu beschreiben, um Ansatzpunkte für Maßnahmen einer den neuen Belastungen gerecht werdenden Gesundheitsförderung zu begründen. Die Bundesregierung fordert, die Rahmenbedingungen in Unternehmen und Einrichtungen so zu gestalten, dass gesundheitliches Verhalten gefördert und Krankheiten vorgebeugt wird (BMGS, 2005; [www.bmgs.bund.de](http://www.bmgs.bund.de)). Legt man dieser Forderung die generelle Gesundheitsdefinition zugrunde, die 1948 in der Präambel der Charta der WHO Gesundheit als Zustand vollständigen körperlichen, psychischen und sozialen Wohlbefindens beschrieben hat, so wird klar, dass Gesundheitsförderung einschließt, die Quellen für psychische Fehlbeanspruchungen zu beseitigen sowie die Leistungsfähigkeit und Kompetenz von Personen zu steigern.

Die geschilderte Situation macht deutlich, dass es in Unternehmen darauf ankommt, nach Wegen zu suchen, die sowohl eine Motivationsentwicklung als auch eine Gesundheitsförderung, insbesondere im Hinblick auf psychische Belastungen, für Mitarbeiter leisten.

In diesem Aufsatz werden Feldstudien vorgestellt, in denen gleichzeitig Eigenschaften von Arbeitsaufgaben und Personenmerkmale Erwerbstätiger analysiert werden. Die von Peter Warr (1996) für den Zweck der Vorhersage der Entwicklung der Leistungsfähigkeit Arbeitender über die Spanne des Erwerbslebens begründete Typologie von Arbeitsaufgaben haben wir dabei als Rahmenkonzept für die Beschreibung von Zusammenhängen zwischen Eigenschaften von Arbeitsaufgaben und Personenmerkmalen Erwerbstätiger genutzt.

## 2 Theoretische Konzepte über Zusammenhänge zwischen Arbeit auf der einen und Motivation, Leistungsfähigkeit und psychischer Gesundheit auf der anderen Seite

Ein anerkanntes und empirisch bestätigtes Konzept der Beschreibung von Zusammenhängen zwischen Arbeitsaufgaben und intrinsischer Motivation ist das Job-Characteristics-Modell von Hackman und Oldham (1976). Fünf Eigenschaften von Arbeitsaufgaben, nämlich Aufgabenvielfalt, Aufgabengeschlossenheit, Aufgabenbedeutung, Autonomie und Rückmeldungen im Arbeitsprozess werden als Voraussetzungen für hohe interne Arbeitsmotivation begründet. Autonomie und Rückmeldung haben dabei die Qualität notwendiger Bedingungen. Fehlen diese Eigenschaften, so kann sich Motivation nicht entwickeln.

Das in der Arbeitspsychologie bekannteste Konzept mit dem Anspruch, Zusammenhänge zwischen Eigenschaft von Erwerbsarbeit und psychischen Beanspruchungen der Arbeitenden zu illustrieren, ist das Job-Demand-Control-Modell von Karasek (Karasek, 1979; Karasek & Theorell, 1990). Karasek hat in Längsschnittstudien an einer repräsentativen Stichprobe schwedischer erwerbstätiger Männer nachgewiesen, dass Beeinträchtigungen des Wohlbefindens, der Medikamentenverbrauch und die Arbeitsunfähigkeitstage mit steigenden Arbeitsanforderungen (demands) und geringer werdenden Entscheidungsspielräumen größer werden. Empirisch ist gezeigt, dass eine hohe Arbeitsbeanspruchung mit psychischen Beeinträchtigungen (z.B. Depression oder Erschöpfung) einhergeht. Entscheidungsspielräume in der Arbeit können negative Effekte hoher Arbeitsbeanspruchung abpuffern. Das Job-Demand-Control-Modell beschreibt Arbeit damit nicht nur als Belastungsquelle sondern auch als Ressource. Das Modell begründet, dass es von den Eigenschaften der Arbeit abhängig ist, welche Folge auf den Menschen bei längerer Ausführung zu erwarten sind. Diese Interpretation legen auch Ergebnisse von Längsschnittstudien zu Zusammenhängen zwischen der inhaltlichen Komplexität der Arbeit und den intellektuellen Fähigkeiten älterer Arbeitender dar. Schooler, Mulatu und Oates (1999) begründen an Befunden aus Längsschnittstudien an Arbeitenden eine Abhängigkeit der intellektuellen Leistungsfähigkeit Erwerbstätiger von der Komplexität ihrer Arbeitsaufgabe. Die Komplexität von Arbeitsaufgaben ist eine Funktion des Entscheidungsspielraums.

Auch die Arbeitsgruppe von Illmarinen (Illmarinen, 1999; Illmarinen & Rantanen, 1999; Illmarinen & Tempel, 2002a, 2002b) kommt auf der Basis umfangreicher Längsschnittstudien zu der Schlussfolgerung, dass Arbeitsinhalte, die einen vielseitigen Gebrauch erworbener Qualifikationen zulassen und die Wei-

terentwicklung von Fähigkeiten herausfordern, das Erhalten und die Entwicklung der Arbeitsfähigkeit im Prozess der Arbeit gewährleisten. Die Gründe für eine gute Entwicklung der Arbeitsfähigkeit im Verlauf der Erwerbstätigkeit werden in den Arbeitsinhalten und der beruflichen Qualifizierung gesehen. Damit werden Arbeitsinhalte auch von diesen Autoren als Ressource klassifiziert. Diesen letzten Gedanken hat Warr (1994, 1996, 1998) entscheidend theoretisch untermauert. Mit Bezug auf entwicklungspsychologische Erkenntnisse über unterschiedliche Alterstrends der fluiden und kristallinen Intelligenz sowie in der Allgemeinen und in der Arbeitspsychologie, speziell der Expertiseforschung gewonnene Befunde, macht Warr Vorhersagen über die Entwicklung der Leistungsfähigkeit Erwerbstätiger während der Spanne des Erwerbsalters. Als Wirkmechanismus des Erhaltens oder sogar des Verbesserns der Leistungsfähigkeit begründet er Möglichkeiten einer erfahrungsabhängigen Selbstoptimierung von Arbeitsmethoden bei existierenden Tätigkeitsspielräumen. Damit begründet er nicht nur aktuelle Zusammenhänge zwischen Arbeit und Person sondern längerfristige. Er argumentiert für Wirkungen der Arbeit auf die Person, welche die Spanne des Erwerbslebens betreffen. Diese Wirkungen auf die Person beschreibt er als Wirkungen auf ihre Leistungsfähigkeit. Warr unterscheidet vier Typen von Tätigkeiten mit spezifischen Wirkungen. Er erwartet positive Alterstrends bei wissensintensiver Arbeit, die ohne Zeitdruck erfolgt, aber einen mehrjährigen Erfahrungsaufbau erfordert (Typ 1: Mit dem Alter Leistungsverbesserungen ermöglichende Tätigkeiten). Negative Alterstrends werden erwartet, wenn Arbeitsaufgaben fortlaufende schnelle Informationsverarbeitungen enthalten, die schnelles Lernen erfordern und die Gefahr der Überforderung der aktuellen Verarbeitungskapazität enthalten, sowie bei schwerer körperlicher Arbeit (Typ 2: Mit dem Alter zu Leistungsver schlechterungen führende Tätigkeiten). Bei qualifizierten manuellen und geistigen Facharbeiten, die mit zunehmendem Alter das Risiko der Überschreitung von Basiskapazitäten bergen, aber Gelegenheiten zu einer erfahrungsabhängigen Kompensation von Leistungseinbußen bei Basiskomponenten durch eine Veränderung der Arbeitsmethoden bieten, werden Nullkorrelationen von Leistungsindikatoren mit dem Alter erwartet (Typ 3: Mit dem Alter eine Kompensation von Leistungsbeeinträchtigungen ermöglichende Tätigkeiten). Für Arbeitsaufgaben mit geringen Anforderungen werden ebenfalls fehlende Alterstrends, allerdings auf geringerem Niveau, erwartet (Typ 4: Altersneutrale Tätigkeiten).

Diese vier Typen fordern, Wirkungen von Arbeit auf den Menschen unter Beachtung der Arbeitsanforderungen zu betrachten und in der Gestaltung von



Arbeitsaufgaben eine Stellschraube sowohl für eine Gesundheitsprävention als auch für die Entwicklung von Leistungsfähigkeit und Motivation im Prozess der Arbeit zu sehen. Die vorzustellende Studie möchte die theoretischen Argumentationen von Warr empirisch prüfen. Dabei wird unterstellt, dass komplexe Aufgaben die Anwendung und damit ein kontinuierliches Training von Wissen und Fähigkeiten gestatten und auch die Entwicklung von Motivation ermöglichen. Im Unterschied dazu haben Arbeitsaufgaben mit nur geringen Anforderungen mangelnde Lernerfahrungen zur Folge, die längerfristig zur Ursache verminderter Leistungsfähigkeit werden können (Barkholdt, Freichs & Naegele, 1995; Kohler & Plath, 2000).

### 5 Aufgabenstellung, Fragestellung und Hypothesen

Für altersgemischte Stichproben Erwerbsstätiger wird die Lernhaltigkeit der Arbeitsaufgabe beschrieben und es werden Indikatoren für arbeitsorientierte Motivation, Leistungsfähigkeit und Gesundheit erhoben. Die Enge der Zusammenhänge zwischen Aufgabeneigenschaften und Personenmerkmalen wird ermittelt.

Zusätzlich werden die Erwerbstätigen anhand ihrer Arbeitsaufgaben in die von Warr (1996) beschriebenen Aufgabentypen klassifiziert. Entsprechend der skizzierten Konzepte und des vorliegenden Kenntnisstandes wird erwartet, dass es Niveauunterschiede in der Ausprägung von Indikatoren für arbeitsorientierte Motivation, Leistungsfähigkeit und Gesundheit gibt.

Für Erwerbstätige, deren Arbeitsaufgaben in den Typ 1 klassifiziert werden, haben die Arbeitsanforderungen die Qualität von Ressourcen, so dass eine Weiterentwicklung der Leistungsfähigkeit im Prozess der Arbeit erwartet werden kann. Erwerbstätige, deren Arbeitsaufgaben dem Typ 2 zuzuordnen sind, erleben hohen Zeitdruck und dadurch die Gefahr einer permanenten Überforderung ihrer Informationsverarbeitungskapazität, so dass keine Weiterentwicklung der Leistungsfähigkeit möglich ist sondern negative Alterstrends resultieren. Für Erwerbstätige, deren Arbeitsaufgaben dem Typ 3 zuzuordnen sind, wird eine Entwicklung der Leistungsfähigkeit in einem Umfang vorhergesagt, die eine Kompensation von Leistungseinbußen bei Basisfähigkeiten zulässt. Für Erwerbstätige, deren Arbeitsaufgaben in den Typ 4 klassifiziert werden, d.h. Aufgaben mit geringen Anforderungen (Routinearbeit), ist aus Gründen einer nur teilweisen Inanspruchnahme und damit eines nur partiellen Trainings erworbenen Wissens mit geringerer Ausprägung der Leistungsfähigkeit zu rechnen.

Die folgende Hypothese wird geprüft:

Die Ausprägungen von Indikatoren der Leistungsfähigkeit weisen die Relationen auf:

$$\text{Typ 1} > \text{Typ 2} > \text{Typ 3} > \text{Typ 4}$$

Es wird erwartet, dass die Argumentation für die Entwicklung der Leistungsfähigkeit auch für die Entwicklung arbeitsorientierter Motivation zutrifft. Modernde Arbeitsaufgaben sind kundenbezogen auszuführen und durch kleine Losgrößen gekennzeichnet, so dass durch ständige Lernprozesse eine Anpassung an wechselnde Kontexte zu leisten ist und Veränderungen in den Arbeitsaufgaben zu bewältigen sind. Das Erhalten der Leistungsfähigkeit ist an selbst organisiertes Lernen gebunden und funktioniert nur bei guter Motivation.

Es wurde begründet, dass Tätigkeitsspielräume enthaltende komplexe Arbeitsaufgaben Selbstkontrolle ermöglichen und dadurch die Qualität von Ressourcen erhalten, so dass auch für Indikatoren der Gesundheit ähnliche Relationen bestehen. Somit wird angenommen, dass die von Warr begründete Aussage über Alterstrends der Leistungsfähigkeit aus Gründen der gemeinsamen Abhängigkeit von Aufgabeneigenschaften auch auf Alterstrends für Indikatoren der arbeitsorientierten Motivation und Gesundheit übertragen werden können.

Das Ziel der Studie besteht darin, Eigenschaften von Arbeitsaufgaben nicht nur zur Vorhersage von psychischen Fehlbeanspruchungen zu nutzen, wie es durch das Job-Demand-Control-Modell von Karasek geschieht, oder zur Vorhersage der Leistungsfähigkeit, wie es durch Warr mit der Begründung der Aufgabentypologie geschieht, sondern aus Eigenschaften von Arbeitsaufgaben sowohl Niveauunterschiede als auch Alterstrends für Indikatoren der Motivation, der Leistungsfähigkeit und Gesundheit zu begründen. Zusammengefasst hat die Studie das Ziel, die folgenden Fragen zu beantworten:

1. Wie eng sind Zusammenhänge zwischen Eigenschaften von Arbeitsaufgaben und Indikatoren für Motivation, Leistungsfähigkeit und Gesundheit?
2. Lassen sich die theoretisch begründeten Niveauunterschiede zwischen Erwerbstätigen, die in die Aufgabentypen nach Warr klassifiziert sind, für Ausprägungsgrade von Indikatoren für arbeitsorientierte Motivation, Leistungsfähigkeit und Gesundheit empirisch bestätigen und wie groß sind sie?
3. Existieren zwischen den in unterschiedliche Aufgabentypen nach Warr klassifizierten Erwerbstätigen unterschiedliche Alterstrends für Indikatoren der arbeitsorientierten Motivation, Leistungsfähigkeit und Gesundheit?

## 4 Methodik

### 4.1 Design

Zur Beantwortung der Fragen und zur Prüfung der Hypothesen wurden Feldstudien an verschiedenen Stichproben Erwerbstätiger durchgeführt, in denen die Arbeitsaufgaben beschrieben und Erhebungen zu verschiedenen Facetten des Selbstkonzepts der Kompetenz, arbeitsbezogener Motivation sowie zum Gesundheitserleben durchgeführt wurden. Eine Klassifikation der Erwerbstätigen entsprechend der von ihnen ausgeführten Arbeitsaufgaben in die durch Warr (1996) begründeten Aufgabentypen erfolgte. Das Design der Studie entspricht somit einem Ex-post-facto-Design. Es leistet eine Bestandsaufnahme zu Ausprägungen von Indikatoren beruflicher Kompetenz, arbeitsbezogener Motivation und erlebter Gesundheit in Abhängigkeit von Eigenschaften der Arbeitsaufgabe. Mit diesem Design können sowohl die angegebenen Zusammenhangshypothesen als auch die Unterschiedshypothesen geprüft werden. Allerdings sind bestätigte Zusammenhangs- und Unterschiedshypothesen nur im Sinne einer Stützung der diesen Hypothesen zugrunde liegenden theoretisch begründeten Wirkmechanismen zu interpretieren.

### 4.2 Stichproben

385 Erwerbstätige aus Industrie und Dienstleistung nahmen an der Untersuchung teil. Die Mehrheit der Untersuchungspartner (74%) arbeitete zum Zeitpunkt der Erhebung in Dienstleistungsbereichen wie Forschung und Lehre, Handel, Versicherung, Immobilien, Systemgastronomie, Pharma-Außendienst und Reinigung. Der Altersdurchschnitt der befragten 215 Männer und 170 Frauen lag bei 40 Jahre.

Die Untersuchungsteilnehmer wurden in Gruppen klassifiziert, die den Aufgabentypen von Warr entsprechen. Dem Typ 1 (mit dem Alter Leistungsverbesserungen ermöglichende Tätigkeiten) wurden Erwerbstätige mit Führungsaufgaben in der Finanzdienstleistung, der Industrie, der Pharmabranche und der Bildungsdienstleistung zugeordnet. Führungsaufgaben beinhalten neben Aufgaben, die Fachkompetenz verlangen, auch strategische Aufgaben, die einen Erfahrungsaufbau erfordern und von diesem profitieren. Damit sind Leistungsverbesserungen bei längerer Ausführung möglich. In diese Gruppe wurden 77 männliche und 20 weibliche Erwerbstätige klassifiziert, die ihre Arbeitsaufgaben als überdurchschnittlich lernhaltig eingeschätzt haben. Sie bearbeiten Führungsaufgaben mit längerfristiger Dauer und deshalb ohne aktuellen Zeitdruck. Das Durchschnittsalter in dieser Gruppe beträgt 43,9 Jahren (Min.: 24 Jahre,

Max.: 64 Jahre). 50 Prozent der Personen weisen eine abgeschlossene Ausbildung (Facharbeiterabschluss bzw. Fachschulabschluss) und 50 Prozent einen Hochschulabschluss (Fachhochschul- bzw. Universitätsabschluss) auf. Tätigkeiten, bei denen Leistungsbeeinträchtigungen mit zunehmendem Alter zu erwarten sind (Typ 2), wurden in der vorliegenden Studie nicht identifiziert, so dass dieser Aufgabentyp bei der Gruppenzuordnung nicht berücksichtigt werden konnte. In den Typ 3 (Kompensation von Leistungsbeeinträchtigungen ermöglichende Tätigkeiten) wurden Angestellte und Mitarbeiter ohne Führungsfunktion klassifiziert. Die Erwerbstätigen üben Tätigkeiten entsprechend ihrer Qualifikation aus, welche Kenntnisse erfordern, die über die Grundlagenausbildung im Beruf hinausgehen. Die Arbeitsaufgaben sind charakterisierbar durch Komplexität und Planbarkeit und ermöglichen die Entwicklung individueller Arbeitsstile bei der Ausführung der Tätigkeit, so dass damit eine Kompensation von altersbedingten Leistungseinschränkungen, z. B. einem Nachlassen der Geschwindigkeit aktueller Informationsverarbeitungen durch Erfahrungen über Regelmäßigkeiten beim Auftreten von Anforderungen möglich wird. 228 Erwerbstätige (119 männliche und 109 weiblichen Personen) wurden dieser Untersuchungsgruppe zugeordnet. Das Durchschnittsalter beträgt 38,7 Jahre (Min.: 21 Jahre, Max.: 61 Jahre). Eine abgeschlossene Ausbildung haben ca. drei Viertel der Personen. Ein Viertel besitzt einen Hochschulabschluss. In den Typ 4 (Altersneutrale Tätigkeiten) wurden Erwerbstätige klassifiziert, deren Arbeitsaufgaben beschreibbar sind als formalisierbare Tätigkeiten, die sich routinemäßig wiederholen. Sie beinhalten gleich bleibende Problemstellungen bei geringer Aufgabenkomplexität. Die Arbeitsaufgaben sind durch hohe Strukturiertheit und vorgegebene Lösungswege definiert. In diese Gruppe wurden 58 Personen klassifiziert (17 männliche und 41 weibliche). Der Altersdurchschnitt ist mit den anderen Gruppen vergleichbar. Er beträgt 38,46 Jahre (Min.: 19 Jahre, Max.: 58 Jahre). Alle Untersuchungsteilnehmer dieser Gruppe besitzen eine abgeschlossene Ausbildung.

### 4.3 Datenerhebung

Die Analyse der Arbeitsaufgaben erfolgte mit dem Fragebogen zur Erfassung lernrelevanter Merkmale der Arbeitsaufgaben (FLMA, Richter & Wardanjan, 2000). Dieses standardisierte Selbstanalyseinstrument, das über gute bis befriedigende Gütekriterien verfügt, beschreibt die Lernhaltigkeit von Arbeitsaufgaben über den Tätigkeitsspielraum, die Anforderungsvielfalt sowie die Transparenz der Informationsflüsse. Mit diesem Konstrukt sowie der Feststellung der Existenz von Führungsaufgaben wurden die von Warr theoretisch

unterschiedenen Tätigkeitstypen operationalisiert. Eine Lernhaltigkeit oberhalb des theoretischen Mittels führte zur Klassifikation der Arbeitsaufgaben in die Typen 1 und 3. Eine Lernhaltigkeit unterhalb des theoretischen Mittels führte zur Klassifizierung in den Typ 4. Zur Unterscheidung der Typen 1 und 2 diente die Existenz von Führungsaufgaben. Traf dieses Merkmal zu, so wurde eine Klassifizierung in den Typ 1 vorgenommen.

Zur Beschreibung der arbeitsorientierten Motivation werden drei standardisierte Instrumente mit guten Gütekriterien eingesetzt, die Skala Eigenaktivität beim Lernen aus dem Fragebogen zum Lernen in der Arbeit (Richter, 2000), die Involvement-Skala (Moser & Schuler, 1993) und die Skala zur Messung der Selbstwirksamkeitserwartung von Schwarzer (1994). Weil die Untersuchungspartner an unterschiedlichen Arbeitsplätzen unterschiedlicher Branchen tätig sind, wurden zur Beschreibung der Leistungsfähigkeit branchenübergreifend einsetzbare Instrumente zur Erfassung des Selbstkonzepts der Fachkompetenz und Methodenkompetenz eingesetzt. Für diese Instrumente sind ebenfalls gute bis befriedigende Gütekriterien nachgewiesen (Richter, 2000). Die Messung der erlebten Gesundheit erfolgte mit einer Kurzversion des General Health Questionnaire GHQ 12 (Nitsche & Richter, 2003). Mit diesem Instrument wurden erlebte Beschwerden erfasst. Es existiert ein Bewertungsmaßstab. Schmitz, Kruse und Tress (1999) begründen, dass ermittelte Skalenwerte oberhalb von 11 als kritisch, darunter liegende Werte als unkritisch zu interpretieren sind. Für die Beschreibung der erlebten Arbeitsfähigkeit wurde die erste Skala für die Erfassung des Work Ability Index (WAI) genutzt (vgl. Ilmarinen & Tempel, 2002).

#### 4.4 Datenanalyse

Die Prüfung der Zusammenhangshypothesen zwischen der Lernhaltigkeit der Arbeitsaufgabe auf der einen Seite und den Personenmerkmalen auf der anderen Seite erfolgte mittels Berechnung von Produkt-Moment-Korrelationen. Die Prüfung der Unterschiedshypothesen sowohl für die Indikatoren der arbeitsorientierten Motivation, für die Indikatoren der Leistungsfähigkeit sowie für Indikatoren der Gesundheit, der in die drei Aufgabentypen nach Warr klassifizierten Erwerbstätigen, wurde an Hand multivariater Varianzanalysen vorgenommen. Zusätzlich werden für alle erhobenen Personenmerkmale Alterskorrelationen bestimmt, indem für die Zusammenhänge zwischen den standardisiert erhobenen Merkmalen und dem Alter Produkt-Moment Korrelationen berechnet wurden.

## 5 Ergebnisse

### *Zusammenhänge zwischen Aufgabeneigenschaften und Personenmerkmalen*

Die Zusammenhänge zwischen der per Selbstanalyse erfassten Lernhaltigkeit der Arbeitsaufgabe auf der einen Seite und den Indikatoren für arbeitsorientierte Motivation, dem Selbstkonzept der Fach- und Methodenkompetenz, der erlebten Arbeitsfähigkeit und den seelischen Beschwerden auf der anderen Seite enthält die Tabelle 1.

Es existieren durchgängig hochsignifikante Zusammenhänge.

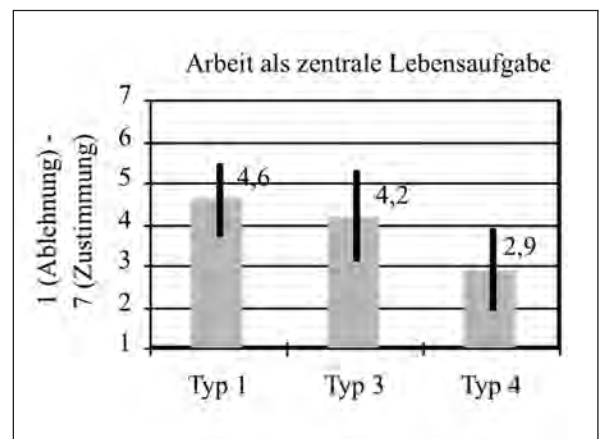
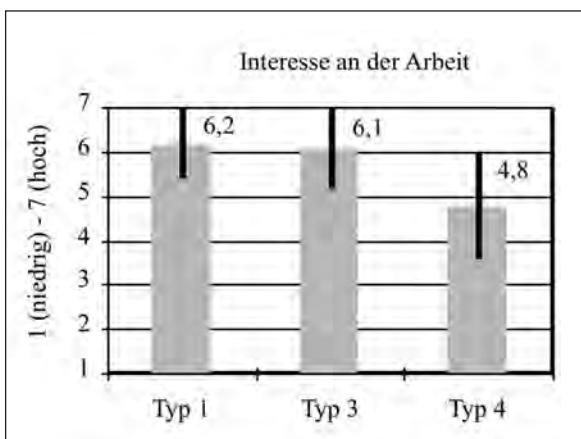
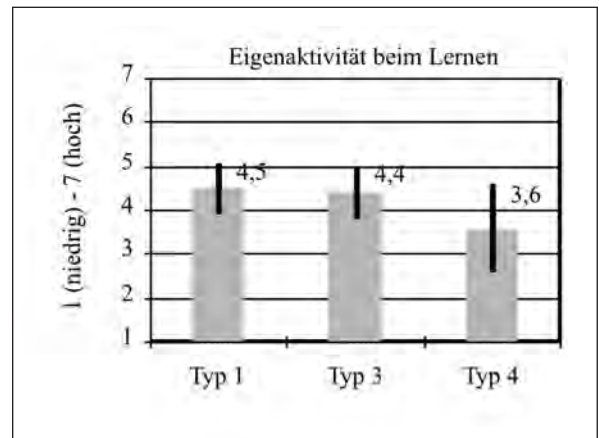
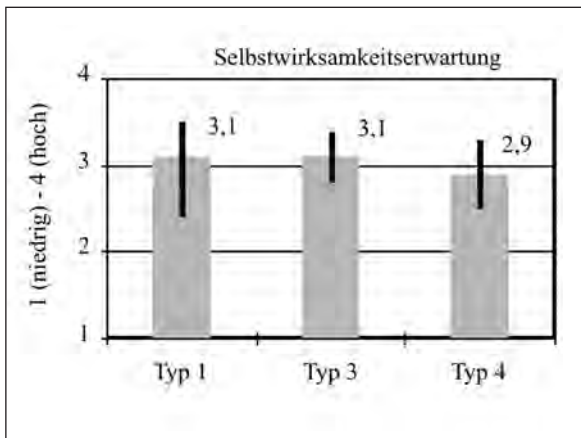
*Tabelle 1: Bivariate Korrelationen zwischen der erlebten Arbeitssituation und den Personenmerkmalen*

Personenmerkmal	Tätigkeitsspielraum r/ p	Lernhaltigkeit der Aufgabe r/ p
Eigenaktivität beim Lernen	0,43/ 0,00	0,52/ 0,00
Job Involvement	0,50/ 0,00	0,63/ 0,00
Selbstwirksamkeitserwartung	0,25/ 0,00	0,30/ 0,00
Selbstkonzept der Fachkompetenz	0,24/ 0,00	0,35/ 0,00
Selbstkonzept der Methodenkompetenz	0,42/ 0,00	0,46/ 0,00
Erlebte Arbeitsfähigkeit	0,18/ 0,00	0,21/ 0,00
Seelische Beschwerden	-0,36/ 0,00	-0,36/ 0,00

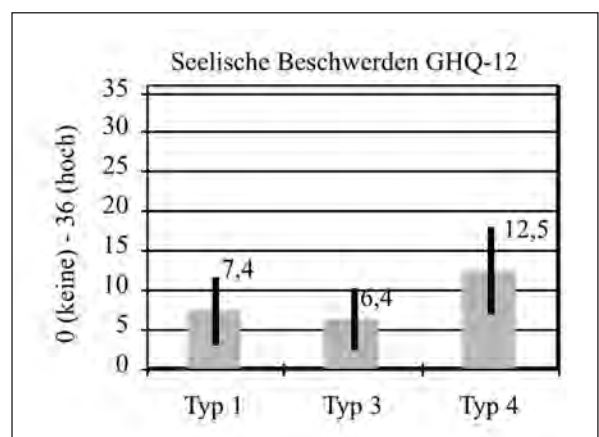
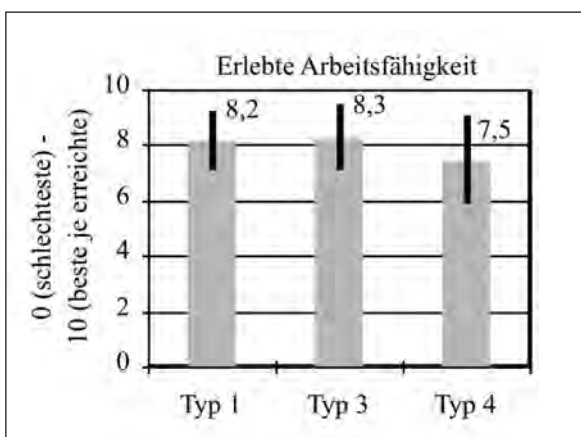
### Unterschiede zwischen den Aufgabentypen nach Warr

Die Abbildung 1 veranschaulicht die Befunde.

a)



b)



c)

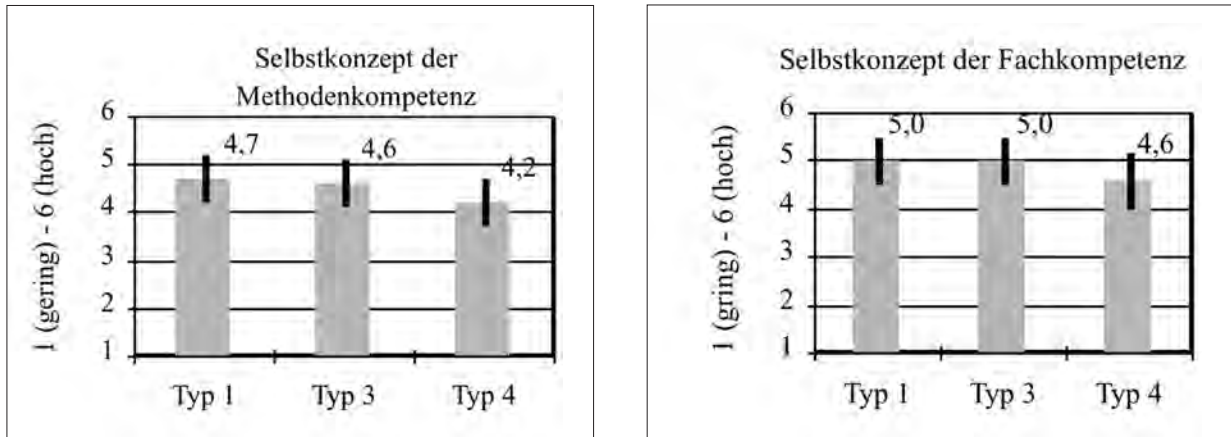


Abbildung 1: Ausprägungsgrade von Indikatoren für arbeitsorientierte Motivation (a), erlebte Arbeitsfähigkeit (WAD) und seelischer Beschwerden (GHQ-12) (b) sowie des Selbstkonzepts der Fach- und Methodenkompetenz (c) bei Erwerbstätigen, die in Aufgabentypen nach Warr (1996) klassifiziert wurden.

Die Ausprägungsgrade aller Indikatoren für arbeitsorientierte Motivation, das Selbstkonzept der Fach- und Methodenkompetenz und der erlebten Arbeitsfähigkeit sind bei den Erwerbstätigen, die in den Typ 1 und den Typ 3 klassifiziert wurden, vergleichbar hoch. Der Umfang an seelischen Beschwerden ist in diesen Gruppen im unkritischen Bereich. Für die Erwerbstätigen, deren Arbeitsaufgaben in den Typ 4 klassifiziert sind, existieren für die Motivationsindikatoren, das Selbstkonzept der Fach- und Methodenkompetenz und die erlebte Arbeitsfähigkeit durchgängig geringere Ausprägungen und für den Umfang an seelischen Beschwerden höhere Ausprägungen.

In der multivariaten Varianzanalyse wurde die Pillai-Spur als Prüfgröße berechnet und als signifikant festgestellt. Sie ist robust gegenüber Verletzungen der Testvoraussetzungen (Rudolf & Müller, 2004). Die Be-

rechnung der Kontraste ergibt für Erwerbstätige des Typs 4 im Vergleich zu Erwerbstätigen des Typs 1 und auch des Typs 3 signifikant geringere Ausprägungen in den Indikatoren arbeitsbezogener Motivation, Leistungsfähigkeit sowie der Einschätzung der Arbeitsfähigkeit. Der Umfang erlebter seelischer Beschwerden ist für die Erwerbstätigen des Typs 4 signifikant größer als bei den Erwerbstätigen des Typs 1 und des Typs 3. Mit einem durchschnittlichen Summenwert von 12,5 wird hier ein als kritisch zu bewertendes Ausmaß erreicht. Die großen Streuungen weisen darauf hin, dass bei diesem Aufgabentyp auch deutliche Überschreitungen des Grenzwertes von 11 vorkommen.

*Alterskorrelationen*

Die Berechnung der Alterskorrelationen ergab überwiegend Nullkorrelationen (s. Tabelle 2). Für Erwerbs-

Tabelle 2: Alterskorrelationen für Indikatoren der arbeitsorientierten Motivation, Leistungsfähigkeit und Gesundheit bei Erwerbstätigen, die nach der Beschaffenheit der Arbeitsaufgabe in die Typen von Warr (1996) klassifiziert wurden.

Indikator	Typ 1 r/ p	Typ 3 r/ p	Typ 4 r/ p
Eigenaktivität beim Lernen	-0,04/ 0,67	0,09/ 0,40	-0,29/ 0,09
Job Involvement	0,05/ 0,78	0,28/ 0,00	0,01/ 0,92
Selbstwirksamkeitserwartung	0,18/ 0,15	0,15/ 0,04	-0,01/ 0,95
Selbstkonzept der Fachkompetenz	-0,07/ 0,96	-0,13/ 0,07	0,03/ 0,86
Selbstkonzept der Methodenkompetenz	0,18/ 0,09	0,16/ 0,02	-0,05/ 0,71
Erlebte Arbeitsfähigkeit	0,09/ 0,57	0,12/ 0,15	-0,02/ 0,91
Seelische Beschwerden	-0,12/ 0,28	0,07/ 0,27	0,14/ 0,29

tätige, die in den Typ 3 klassifiziert wurden, existieren für die Indikatoren Job Involvement, Selbstwirksamkeitserwartung und das Selbstkonzept der Methodenkompetenz leicht positive Alterskorrelationen. Tendenziell gibt es auch bei Erwerbstätigen des Typs 1 leicht positive Alterstrends und für den Umfang der seelischen Beschwerden negative Alterstrends, d.h. eine Abnahme der Beschwerden mit dem Alter. Aber aufgrund der hohen Streuungen wird die Signifikanzschranke verfehlt.

## 6 Diskussion

Die Studie bestätigt die erwarteten Zusammenhänge zwischen der Lernhaltigkeit von Arbeitsaufgaben und Personenmerkmalen Erwerbstätiger. Eine höher eingeschätzte Lernhaltigkeit der Arbeitsaufgabe steht im Zusammenhang mit höheren Ausprägungen von Indikatoren arbeitsorientierter Motivation, des Selbstkonzepts der Fach- und Methodenkompetenz, der erlebten Arbeitsfähigkeit und in geringerem Umfang angegebenen seelischen Beschwerden. Die Zusammenhänge mit Motivationsindikatoren sind tendenziell enger. Eine Klassifizierung der Erwerbstätigen in die Typologie von Warr bestätigte die Hypothese, dass mit dieser Klassifizierung Niveauunterschiede bei Indikatoren arbeitsorientierter Motivation, der Leistungsfähigkeit und Gesundheit einhergehen. Die hypothetisch vorhergesagten ungünstigeren Ausprägungsgrade der Personenmerkmale bei Erwerbstätigen, deren Arbeitsaufgaben dem Typ 4 angehören, die Routinearbeit verrichten, bei der kaum Möglichkeiten zu einem Wissens- bzw. Kompetenzaufbau mit Verwertungsmöglichkeiten bei der Entwicklung von Arbeitsmethoden bestehen, haben sich empirisch klar bestätigt. In dieser Gruppe existieren signifikant schlechtere Ausprägungsgrade bei den gemessenen Personenmerkmalen. Die Unterschiede sind bei Motivationsindikatoren größer als beim Selbstkonzept der Fach- und Methodenkompetenz und der erlebten Arbeitsfähigkeit. Die erlebten Beschwerden erreichen bei den in den Typ 4 klassifizierten Erwerbstätigen ein kritisches Niveau und signalisieren damit Gestaltungsbedarf.

Die Hypothese über angenommene Niveauunterschiede zwischen Erwerbstätigen, deren Arbeitsaufgaben dem Typ 1 und dem Typ 3 angehören, bei Indikatoren der arbeitsorientierten Motivation, Leistungsfähigkeit und Gesundheit hat sich nicht bestätigt.

In dieser Studie wurden die Arbeitssituation und die Personenmerkmale gleichzeitig erhoben. Das Design ist nicht geeignet, Wirkrichtungen zu begründen. Die längere Verweildauer der Erwerbstätigen unter den dargestellten Arbeitsbedingungen, kann aber mit Hilfe der Heuristik, dass die Ursache der Wirkung zeit-

lich vorausgeht, für die gemessenen aktuellen Ausprägungsgrade bei Personenmerkmalen als ein Argument für eine Begründung der Wirkrichtung von der Arbeit auf die Person gelten. Tatsächlich sind Wirkungspfade von der Arbeit auf die Person nicht unidirektional sondern durch komplexe Interaktionen gekennzeichnet. Motivation, Wissen, Handlungsfähigkeit und Gesundheit sind nicht statisch. Sie werden immer wieder hergestellt und für diese Prozesse sind die Eigenschaften von Aufgaben, der Umfang in dem Wissen eingesetzt und damit trainiert oder gar die Entwicklung neuen Wissens herausgefordert wird, entscheidend. Rückkopplungen in Form von Anerkennung sind für die Entwicklung von Motivation, aber auch für seelische Gesundheit eine Voraussetzung. Die Studie zeigt, dass bei längerfristiger Beschäftigung mit Arbeitsaufgaben, die geringere Ausprägungen in der Lernhaltigkeit aufweisen, der Umfang an seelischen Beschwerden ein kritisches Niveau erreicht.

Die Berechnung von Alterskorrelationen für die Personenmerkmale erbrachte im Wesentlichen nicht signifikante Alterskorrelationen, für Erwerbstätige, deren Arbeitsaufgaben in den Typ 3 klassifiziert sind, in drei Fällen leicht positive Alterstrends. Generell existieren aber bei allen Personenmerkmalen große Streuungen, so dass die Alterskorrelationen keine sensiblen Kennwerte darstellen.

Die Studie macht deutlich, Motivation, Leistungsfähigkeit und Gesundheit der Erwerbstätigen sind voneinander und von der Beschaffenheit von Arbeitsaufgaben abhängig. Sie stützt Hypothesen, die in der Gestaltung lernhaltiger Arbeitsaufgaben einen Ansatz für das Erhalten und Entwickeln der Beschäftigungsfähigkeit und für eine Gesundheitsprävention sehen.

## Literatur

- Antonovsky, A. (1979). *Health, stress and coping. New perspectives on mental and physical well-being*. San Francisco: Jossey-Bass.
- Antonovsky, A. (1997). Salutogenese – Zur Entmystifizierung der Gesundheit. *Forum für Verhaltenstherapie und psychosoziale Praxis*. Tübingen: dgvt-Verlag.
- Barkholdt, C., Frerichs, F. & Naegele, G. (1995). Altersübergreifende Qualifizierung – eine Strategie zur betrieblichen Integration älterer Arbeitnehmer. *Mitteilungen aus der Arbeitsmarkt- und Berufsforschung*, 3, 425-436.
- BMGS (2005). URL: [www.bmgs.bund.de](http://www.bmgs.bund.de).
- Dörre, K. (2006). Prekäre Arbeit. Unsichere Beschäftigungsverhältnisse und ihre sozialen Folgen. *Arbeit, Zeitschrift für Arbeitsforschung, Arbeitsgestaltung und Arbeitspolitik*, 1 (15), 181-195.

- Hackman, R. & Oldham, G. (1976). Motivation through the design of work: test of a theory. *Organizational Behavior and Performance*, 21, 289-304.
- Ilmarinen, J. (1999). *Ageing workers in the European Union – Status and promotion of work ability, employability and employment*. Helsinki: Finnish Institute of Occupational Health.
- Ilmarinen, J. & Tempel, J. (2002a). Erhaltung, Förderung und Entwicklung der Arbeitsfähigkeit – Konzepte und Forschungsergebnisse aus Finnland. In Bernhard Badura, Christian Vetter & Henner Schellschmidt (Hrsg.), *Fehlzeiten-Report*. Berlin: Springer.
- Ilmarinen, J. & Tempel, J. (2002b). *Arbeitsfähigkeit 2010. Was können wir tun, damit Sie gesund bleiben?* Hamburg: VSA-Verlag.
- Ilmarinen, J. & Rantanen, J. (1999). Promotion of work ability during aging. *American Journal of Industrial Medicine Supplement*, 1, 21-23.
- Karasek, R. A. (1979). Job demands, job decision latitude and mental strain: Implications for job redesign. *Administrative Science Quarterly*, 24, 285-308.
- Karasek, R. A. & Theorell, T. (1990). *Healthy work, stress, productivity, and the reconstruction of working life*. New York: Basic Books.
- Koller, B. & Plath, H.-E. (2000). Qualifikation und Qualifizierung älterer Arbeitnehmer. *Mitteilungen aus der Arbeits- und Berufsforschung*, 1, 112-125.
- Nitsche, I. & Richter, P. (2003) *Tätigkeiten außerhalb der Erwerbsarbeit – Evaluation des Tauris-Projekts*. Münster: LIT Verlag.
- Personalführung (2006). *Zahlen, Fakten, Meinungen – Aktuelle Studien und Umfragen*, S. 15, Heft 6.
- Pietrzyk, U. (2003). *Flexible Arbeitsform – Zeitarbeit auf dem Prüfstand*. Arbeit, Zeitschrift für Arbeitsforschung, Arbeitsgestaltung und Arbeitspolitik, 2 (12), 112-130.
- Richter, F. (2000). Methodik der Querschnittsuntersuchung. In Bergmann, Bärbel; Fritsch, Andreas; Göpfert, Peggy; Richter, Falk; Wardanjan, Barbara & Wilczek, Susan (Hrsg.), *Kompetenzentwicklung und Berufsarbeit, edition QUEM, Bd. 11*, 55-131, Münster: Waxmann.
- Richter, F. & Wardanjan, B. (2000). Die Lernhaltigkeit der Arbeitsaufgabe – Entwicklung und Erprobung eines Fragebogens zu lernrelevanten Merkmalen der Arbeitsaufgabe (FLMA). *Zeitschrift für Arbeitswissenschaft*, 175-183.
- Robert-Koch-Institut (2006). *Gesundheit in Deutschland. Gesundheitsberichterstattung des Bundes*. Berlin: Statistisches Bundesamt.
- Rudolf, M. & Müller, J. (2004). *Multivariate Verfahren. Eine praxisorientierte Einführung mit Anwendungsbeispielen in SPSS*. Göttingen: Hogrefe.
- Schooler, C., Mulatu, M. S. & Oates, G. (1999). The continuing effects of substantively complex work on the intellectual functioning of older workers. *Psychology and Aging* 17, 31-75.
- Schmitz, N., Kruse, J. & Tress, W. (1999). Psychometric properties of the General Health Questionnaire (GHQ-12) in a German primary care sample. *Acta Psychiatrica Scandinavica*, 100, 462-468.
- Moser, K. & Schuler, H. (1993). Validität einer deutschsprachigen Involvementsskala. *Zeitschrift für Differentielle und Diagnostische Psychologie*, 14, 27-36.
- Schwarzer, R. (1994). Optimistische Kompetenzerwartung: Zur Erfassung einer personalen Bewältigungsressource. *Diagnostica*, 40 (2), 105-123.
- Sonntag, K. & Schäfer-Rauser, U. (1995). Selbsteinschätzung beruflicher Kompetenzen bei der Evaluation von Bildungsmaßnahmen. *Zeitschrift für Arbeits- und Organisationspsychologie*, 37 (4), 165-171.
- Technikerkrankenkasse (2009). *Gesundheitsreport 2009 – Veröffentlichungen zum betrieblichen Gesundheitsmanagement der TK, Bd. 21*, Hamburg.
- Tuomi, K., Ilmarinen, J., Jahkola, A., Katajarinne, L. & Tulkki, A. (1998). *Work ability Index*. Helsinki: Finnish Institute of Occupational Health.
- Warr, P. (1994). Age and employment. In Triandis, Harry C., Dunette, Marvin D. & Hough, Laetta M. (Ed.), *Handbook of industrial and organizational psychology* (S. 485-550). Palo Alto.
- Warr, P. (1996). Younger and older workers. In Peter Warr (Hrsg.), *Psychology at work* (4. Aufl.) (S. 308-332). Harmondsworth: Penguin Books.
- Warr, P. (1998). Age, work and mental health. In Schaie K.W. and Schooler C. (Eds.), *Impact of work on older adults*. Societal Impact on Aging Series. New York: Springer.
- Korrespondenz-Adresse:  
Prof. Dr. Bärbel Bergmann  
Technische Universität Dresden  
Fakultät Mathematik und Naturwissenschaften  
Institut für Psychologie I  
Professur für Methoden der Psychologie  
D-01062 Dresden  
bergmann@psychomail.tu-dresden.de

# Vom Schatten herrschender Verhältnisse oder: Was fördert Organisationales Commitment?

Birgit E. Schmid

Marienhaus GmbH, Waldbreitbach

## ZUSAMMENFASSUNG

Organisationales Commitment gilt in vielerlei Hinsicht als verheißungsvolles Konstrukt. Der wissenschaftlichen Forschung bietet es Untersuchungsmöglichkeiten vorläufig noch ohne absehbares Ende, dafür sorgt allein schon die Mehrdimensionalität des Merkmals. Für die betriebliche Praxis wird es vor allem durch sein Versprechen interessant zu mehr Anstrengungs- und Leistungsbereitschaft von Mitarbeiterinnen und Mitarbeitern beizutragen. Ausgehend von den Ergebnissen einer Untersuchung zum Thema Beteiligungsorientierung in Unternehmen (Schmid, 2009) und den Erfahrungen aus einem zweieinhalbjährigen Organisationsentwicklungsprojekt in einem Krankenhaus der Grund- und Regelversorgung mit knapp 1.000 Mitarbeitenden in Deutschland, wird im vorliegenden Beitrag der Frage nachgegangen, wie sich Beteiligung, prozedurale Gerechtigkeit und eine soziomoralische Atmosphäre auf das Commitment der Beschäftigten auswirken. Dahinter steht die Überzeugung, dass eine grundsätzlich positive Bindung an das Unternehmen eine strukturelle Verankerung der genannten Merkmale im Unternehmen voraussetzt und es nicht ausreicht, diese nur fallweise oder anlassbezogen zu gewähren.

## Schlüsselwörter

Commitment – Beteiligung – Gerechtigkeit/Fairness – Wertschätzung

## ABSTRACT

Organizational Commitment is in many respects a promising construct. In the scientific field, especially the multidimensionality of the construct offers a plethora of investigations. In business practice, the construct remains interesting due to its contribution to employees' effort and motivation. Based on the results of an investigation of participation orientation in organizations (Schmid, 2009) and experience gained from a two-year organizational development project in a hospital with almost 1,000 employees in Germany, we investigate how participation, procedural fairness, and a socio-moral atmosphere effect employees' commitment. We posit that a fundamentally positive commitment to one's organization requires a structural anchoring of the mentioned characteristics. Occasional or ad hoc grants of these is, however, not sufficient.

## Keywords

Commitment – participation – justice/fairness – estimation



## 1 Einführung

Was bewegt Menschen dazu, sich mit einem Unternehmen, seinen Werten und Zielen zu identifizieren? Unter welchen Bedingungen sind Menschen bereit, sich mit ihrer ganzen Arbeitskraft verantwortungsvoll und eigeninitiativ in ihrem Arbeitsbereich tagtäglich zu engagieren? Diese und ähnliche Fragen bewegen nicht nur Wissenschaft und Forschung, sondern gleichermaßen auch betriebliche Praktiker. In Zeiten sich verknappender finanzieller Ressourcen und einem ansteigenden Mangel an Fachkräften und Experten, entwickelt sich dieses Thema zunehmend zu einem erfolgsentscheidenden Faktor. Es geht dabei nicht nur darum gute Mitarbeiterinnen und Mitarbeiter zu gewinnen und im Unternehmen zu halten, sondern vor allem um deren Einsatz- und Leistungsbereitschaft.

Tatsächlich ist jedes Unternehmen auf das Engagement seiner Mitarbeiter angewiesen. Mehr noch: Im Grunde lebt jedes Unternehmen vom sogenannten Extra-Rollenverhalten, einer Leistungsbereitschaft und -erbringung der Mitglieder, die über die eigentliche Vertragserfüllung hinausgeht. Neuere Befragungsergebnisse zum Mitarbeiterengagement zeichnen gerade diesbezüglich jedoch ein düsteres Bild. Der Gallup Engagement Index 2004 (Gallup, 2005) berichtet, dass lediglich 15 Prozent der Arbeitnehmer in Deutschland ihrer Arbeit mit vollem Engagement nachgehen. Die Global Workforce Studie aus den Jahren 2004 bis 2006 (Towers Perrin, 2006) enthält vergleichbare Ergebnisse. Demnach beschreiben sich nur rund 20 Prozent der Befragten als engagierte Mitarbeiter. Gleichzeitig wird in dieser Studie aber auch darauf verwiesen, dass sich Arbeitnehmer durchaus motivieren lassen.

Als Treiber für Mitarbeiterengagement gelten die Kommunikation von Karrieremöglichkeiten, ausreichende Entscheidungsfreiheit, der Ruf des Unternehmens als Arbeitgeber, faire Vergütung im Vergleich zu Arbeitskollegen, angemessene Nebenleistungen, der Vorgesetzte versteht, was die Mitarbeiter motiviert, Work/Life-Balance, Bindung von erfolgskritischen Mitarbeitern, Programme und Anreize zur Gesundheitsvorsorge, Vorgesetzte sind offen und zugänglich.

### *Commitment als Erfolgsfaktor*

In der organisationspsychologischen wie auch in der betriebswirtschaftlichen Literatur gilt das *Commitment* als ein Schlüsselfaktor für engagierte Mitarbeiter. Ganz allgemein bedeutet Commitment „Bindung“ und beschreibt als Organisationales Commitment die „psychologische Bindungsbeziehung zwischen Mitarbeitern und Unternehmen“ (Matiaske & Weller, 2003) bzw. „a global, systemic reaction that people have to the company for which they work“ (Colquitt, Conlon, Wesson, Porter & Ng, 2001, S. 429). Mit dem positiv an

die Organisation gebundenen Mitarbeiter wird hohe Leistungsbereitschaft ebenso wie hohe Teilnahmebereitschaft in Form geringer Fehlzeiten und geringer Fluktuationsneigung verbunden. Die Erforschung des Organisationales Commitment erhält ihre Bedeutung für die Praxis damit durch ihr implizites Versprechen zu einer Verbesserung der Effektivität im Unternehmen beizutragen. Moser (1997, S. 167) formuliert warum „aus praktischer Sicht ein Interesse daran besteht zu verstehen, was Commitment ist und wie es zustande kommt: Organisationsmitglieder müssen partizipieren und produzieren, und zumindest für die Partizipation scheint Commitment relevant zu sein“. Den Erfahrungen aus einem inzwischen zweieinhalb Jahre laufenden umfassenden Organisationsentwicklungsprojekt in einem Klinikum der Grund- und Regelversorgung mit knapp 1.000 Mitarbeitenden folgernd, scheint Commitment durchaus auch für das Produzieren bzw. Erbringen von Dienstleistungen relevant zu sein. Es drängt sich demnach unwillkürlich die Frage auf, welche Faktoren dazu dienlich sind ein hohes Commitment von Mitarbeitern an ihr Unternehmen zu unterstützen und zu fördern.

Im Folgenden werden Ergebnisse aus einer Untersuchung von Unternehmen mit besonderen Formen der Mitarbeiterbeteiligung vorgestellt. Ergänzt wird die Stichprobe durch Unternehmen mit klassisch bürokratischen Organisationsformen (vgl. Schmid, 2009; Weber, Unterrainer & Schmid, 2009).

## 2 Beteiligungsgrad und Organisationales Commitment

Beteiligung im Sinne von Partizipation/Mitbestimmung an Entscheidungen in einem Unternehmen ist in den westlichen Industrieländern in den Grundzügen gesetzlich vorgeschrieben und geregelt. Arbeitgeber- und Arbeitnehmervertretungen stehen in gemeinsamer Verantwortung für den Erfolg eines Unternehmens. Hierüber besteht gesellschaftspolitischer Konsens, vor allem was die repräsentative Einbindung der Mitarbeiter durch den Betriebsrat in Fragen von Arbeitsplatzert halt und Einkommensgestaltung betrifft. Was dieses Modell nicht verhindern kann, das sind das leise Schwelen bzw. das Aufflackern alter Bilder des Klassenkampfes in Zeiten der Krise. Aktuell häufen sich daher die Stimmen auch aus der Politik, welche eine Beteiligung der Arbeitnehmer am Gewinn ihrer Unternehmen fordern. Dabei scheint es, als sähen politisch Verantwortliche in dieser Beteiligung eine haushaltsneutrale Möglichkeit das Absinken der Reallöhne und ein immer weiteres Auseinanderdriften der Einkommen von Managern/Unternehmern und Arbeitnehmern zu verhindern oder wenigstens zu entschärfen.

Die Forderung mag grundsätzlich zu begrüßen sein, sie greift unter der Frage „Was motiviert meine Mitarbeiter zu besonderer Leistung und bindet sie an mein Unternehmen“ jedoch zu kurz (vgl. Herzberg, Mausner & Snyderman, 1959). Nicht allein monetäre Beteiligungsanreize am Jahresende motivieren, sondern vor allem die kontinuierliche Beteiligung während des Jahres an organisationalen und unternehmerischen Entscheidungen. Auf diese Effekte verweist die inzwischen rund 50jährige Tradition rund um die Organisationale-Demokratie-Forschung (u.a. IDE, 1981; Emery & Thorsrud, 1982; Wilpert & Rayley, 1985; Hacker, 1986; Weber, 1999; Harrison & Freeman, 2004; Foley & Polanyi, 2006).

Je mehr Menschen die Möglichkeit und die Kompetenz haben, ihr Unternehmen verantwortlich mitzugestalten, desto mehr identifizieren sie sich mit diesem und sind bereit sich aktiv dafür einzusetzen (Moser, 1997; Matiaske & Weller, 2003). Organisationales Commitment hängt u.a. von Arbeitsmerkmalen und Arbeitserfahrungen in Zusammenhang mit eigenem Kompetenzerleben und dem Gefühl, einen wichtigen Beitrag zum Erfolg der Organisation zu leisten ab (Meyer & Allen, 1997). Nach Rhodes und Steers (1981) weisen Arbeitnehmer, die in wichtige organisationale Entscheidungen eingebunden bzw. am Unternehmen beteiligt sind, ebenfalls ein höheres Commitment aus. Dies alles spricht dafür, dass eine substanzielle Beteiligung im und am Unternehmen ein hohes positives Commitment befördert.

### 3 Aspekte von Verfahrensgerechtigkeit und Organisationales Commitment

Aus einer anderen Forschungslinie stammt die Erkenntnis, dass Entscheidungs- und Verteilungsverfahren bestimmte Prozesskriterien erfüllen müssen, um unter dem Anspruch von Gerechtigkeit in ihren Ergebnissen Zustimmung und Akzeptanz bei den Betroffenen zu finden und damit deren Identifikation zu unterstützen. Beteiligung allein reicht hierfür nicht aus, so die Erkenntnis aus der Gerechtigkeitsforschung. Neben der Beteiligung Betroffener sind Merkmale wie Transparenz, Unvoreingenommenheit, Genauigkeit und Konsistenz einzuhalten (vgl. Leventhal, 1980; Lind & Tyler, 1988; Greenberg, 1990; Cohen-Carash & Spector, 2001; Colquitt et al., 2001).

Konsistenz meint die Widerspruchsfreiheit der verwendeten Regeln über die Zeit und durch verschiedene Autoritäten, Genauigkeit die Berücksichtigung aller Informationen und Standpunkte, Unvoreingenommenheit die Unterdrückung von einseitigen Urteilen und das Außerachtlassen persönlicher, egoistischer Interessen der Entscheider und Transparenz die

Offenlegung und Begründung der Entscheidungskriterien. Die positive Wirkung prozeduraler Gerechtigkeit zeigt sich in Effekten, die von Unternehmen alle als höchst erfreulich betrachtet werden, da sie in der Regel ergebnis- und identifikationssteigernde Wirkung entfalten. Als gerecht wahrgenommene betriebliche Verfahren tragen zu höherer Arbeitszufriedenheit bei, stärken das Vertrauen in das Management (Alexander & Ruderman, 1987; Dörfel & Schmitt, 1997) und unterstützen Organizational Citizenship Behavior (Organ & Moorman, 1995). Ein Großteil der Autoren ist sich einig, dass die prozedurale Gerechtigkeit wesentlich zur Identifikation mit dem Unternehmen beiträgt und in starkem Maß Einfluss auf das Organisationale Commitment der Beschäftigten nimmt (u.a. Folger & Konovsky, 1989; Greenberg, 1990; Moorman, 1991; McFarlin & Sweeney, 1992; Masterson et al., 2000).

### 4 Aspekte einer soziomoralischen Atmosphäre und Organisationales Commitment

John Rawls bezeichnet in seinem Hauptwerk „A Theory of Justice“ (1971) die Gerechtigkeit als „die erste Tugend sozialer Institutionen“. In sozialen Institutionen leben und arbeiten Menschen zusammen und dieses Zusammenleben erfordert Regeln. Es ermöglicht aber auch Entwicklung. Während Martin Buber vom Reifen des Einzelnen am Du spricht (Buber, 1995), betont Lawrence Kohlberg stärker das Reifen an und durch die Gemeinschaft (Kohlberg, 1996). In seiner kognitiven Entwicklungstheorie zum Moralerwerb betont Kohlberg, dass Moral nicht ohne den Bezug auf eine soziale Gemeinschaft gelernt und praktiziert werden kann. Vielmehr ist die „moralische Entwicklung ... in jeder Hinsicht ein sozialer Vorgang“ und hat „als Entwicklungsziel ... die Schaffung moralischer, d.h. gerechter Interaktionsstrukturen“ (Kohlberg, Wassermann & Richardson, 1978, S. 223). Entscheidend in der Moraltheorie Kohlbergs ist die Betonung des Entwicklungscharakters. Es wird den einen abschließenden vollkommenen Moralstandard nicht geben, aber es gibt Bedingungen, die eine Weiterentwicklung der moralischen Urteilsfähigkeit des Menschen behindern und solche, die sie fördern. In einer Längsschnittstudie zum Thema „Arbeitsbiographie und Persönlichkeitsentwicklung“ gelang es einer Gruppe um Ernst-H. Hoff und Wolfgang Lempert am Max-Institut für Bildungsforschung in Berlin entsprechende Anregungspotenziale zu extrahieren (Hoff & Lempert, 1990; Lempert & Corsten, 1997). Als relevante soziale Prozesse, die die Weiterentwicklung des moralischen Bewusstseins begünstigen, nennen sie

- offene Konfrontation mit sozialen Problemen und Konflikten

- zuverlässig gewährte Wertschätzung, Zuwendung und Unterstützung
- zwanglose Kommunikation und partizipative Kooperation
- angemessene Zuweisung und Zurechnung von Verantwortung.

Um ihre entwicklungsförderliche Wirkung entfalten zu können, müssen diese Anregungspotentiale über einen längeren Zeitraum hinweg, Lempert und Corsten (1997) sprechen von mindestens zwei Jahren, konstant ausgeprägt sein. Das bedingt in der Regel auch deren Abbildung in den strukturellen, prozessualen Bedingungen einer Organisation. Neben der Persönlichkeitsförderlichkeit dieser Merkmale und ihrem bewusstseinsbildenden Charakter, haben diese Anregungspotenziale aber auch gemeinschafts- und identitätsstiftendes Potenzial. Matiaske und Weller (2003) berichten, dass Mitarbeiter, denen Wertschätzung und Vertrauen entgegen gebracht werden, dies mit Loyalität und dem Gefühl der Verpflichtung und Verbundenheit dem Unternehmen gegenüber quittieren. Weitere Ergebnisse aus der Commitment-Forschung belegen die positive Wirkung wahrgenommener Wertschätzung in der Arbeit auf die emotionale Bindung der Mitarbeiter an das Unternehmen (u.a. Meyer & Allen, 1997; Moser, 1997; Dörfel & Schmitt, 1997; McFarlin & Sweeney, 1992).

## 5 Organisationales Commitment

Der Commitment-Begriff ist in der betrieblichen Alltagssprache bereits fest verankert, so zumindest bei den Führungskräften. Kein Projekt, wo im Vorfeld nicht nach dem Commitment der Beteiligten gefragt wird. Kein Veränderungsvorhaben, dem ohne das Commitment des Top Managements Chancen eingeräumt werden. Das „Commitment der Betroffenen einholen“ ist fast schon zu einem Mantra geworden. Vielfach wird es auch so gehandhabt: in dem schönen Glauben, dass allein das tausendfache Wiederholen schon dessen Existenz befördert, wird wieder und wieder das Commitment beschworen. Doch so einfach ist es nicht mit diesem Konstrukt, jedenfalls nicht in der realen betrieblichen Praxis.

Auch die psychologische Forschung hat ihre Geschichte mit dem Organisationalen Commitment. Inzwischen bald 40 Jahre intensive Beschäftigung mit dem Konstrukt haben verschiedene Ansätze hervor gebracht (zum Überblick s. Moser, 1997). Aktuell wird weitestgehend einheitlich auf der Basis des integrativen Ansatzes von Meyer und Allen (1991) gearbeitet. Das Modell integriert die verschiedenen Forschungslinien und unterscheidet drei Commitment-Dimensio-

nen: affektives, abwägendes und normatives Commitment.

Im vorliegenden Forschungsdesign kommt vor allem dem affektiven und dem normativen Commitment Bedeutung zu. Dem abwägenden Commitment, welches vorrangig auf individuellen Abwägungsprozessen der Vor- und Nachteile des Verbleibens in oder Verlassens der Organisation beruht (vgl. Meyer & Allen, 1997), wird sowohl theoriegeleitet als auch auf Grundlage früherer Untersuchungen (Schmid, 2004) in diesem Kontext keine signifikante Wirkung zugesprochen, was entsprechende Prüfungen auch bestätigen (vgl. Schmid, 2009).

Die impliziten Versprechen des Commitments sind vorrangig mit dem affektiven Commitment verbunden. Nach Meyer und Allen (1997) ist affektives Commitment von verschiedenen Faktoren beeinflussbar, am stärksten jedoch durch die konkreten Arbeits Erfahrungen des Mitarbeiters in der Organisation. Kompetenzerfahrung und Wohlbefinden sind dabei die zentralen Erfahrungsdimensionen, die affektives Commitment befördern. Meyer und Allen gehen noch einen Schritt dahinter zurück und verweisen auf die Bedeutung organisationaler Strukturmerkmale, die derartige Erfahrungen erst ermöglichen. Neben strukturell verankerten Beteiligungsmöglichkeiten an betrieblichen Entscheidungs- und Verteilungsverfahren sind dies in unserem Untersuchungskontext Strukturmerkmale gerechter Verfahren und einer wertschätzenden Unternehmenskultur.

Das normative Commitment beschreibt eine Art Verpflichtungsgefühl gegenüber dem Unternehmen. Es resultiert „aus der moralischen Verpflichtung heraus, sich in einer Weise zu verhalten, die den Zielen und Interessen der Organisation“ entspricht (Schmid, Hollmann & Sodenkamp, 1998, S. 95). Dieses Verpflichtungsgefühl wird als Ergebnis eines Internalisierungsprozesses, der im Rahmen familiärer, gesellschaftlicher und betrieblicher Sozialisation stattfinden kann betrachtet, kann aber auch als eine spezielle Form der Reziprozität im Kontext sozialer Austauschbeziehungen konzipiert werden (Eisenberger, Armeli, Rexwinkel, Lynch & Rhoades, 2001). Hier kann auch eine Unternehmenskultur inkludiert werden, welche ein ausgeglichenes Geben und Nehmen von Unternehmen und Mitarbeiter sicherstellt. Nicht zuletzt sind es die Investitionen, die ein Unternehmen in den Mitarbeiter tätigt (Aus- und Weiterbildung, Karrierepfade etc.), die normatives Commitment begünstigen. Ein Unternehmen, das sich seinen Mitarbeitern gegenüber fair verhält, kann ebenfalls mit normativ gebundenen Mitarbeitern rechnen (vgl. Matiaske & Weller, 2003).

Den bisherigen Forschungsergebnissen folgend, gibt es gute Gründe in Unternehmen mit einem hohen Beteiligungsgrad der Mitarbeitenden von affektiv und

normativ gebundenen Mitarbeitern rechnen zu können. Gleichzeitig kann erwartet werden, dass sich in beteiligungsorientierten Unternehmen faire Entscheidungs- und Verteilungsverfahren und ein gutes sozio-moralisches Klima ausprägen können.

## 6 Untersuchungsmethode

### Stichprobe

Für die Überprüfung dieser Hypothesen wurden in 30 Unternehmen, welche sich durch unterschiedliche Formen und Grade der Mitarbeiterbeteiligung auszeichnen, die Mitarbeitenden mittels Fragebogen zu den Merkmalen Beteiligung (Organisationale-Demokratie-Struktur, Weber, 2004), prozedurale Gerechtigkeit (Schmitt & Dörfel, 1995), soziomoralische Atmosphäre (Weber, Iwanowa, Schmid & Unterrainer, 2004) und Organisationales Commitment (Felfe, Six, Schmook & Knorz, 2004) befragt. Ergänzt wurden die Befragungen durch Interviews mit betrieblichen Experten, Dokumentenanalysen (Unternehmensvertrag, Verfassung, Leitbilder, innerbetriebliche Vereinbarungen, Geschäftsberichte etc.) und Unternehmensbesichtigungen. Dadurch konnte von den Unternehmen ein umfassendes Bild gewonnen werden.

Nach Datenbereinigung können an dieser Stelle die Daten von insgesamt 420 Probanden aus 30 Unternehmen berichtet werden. Die Unternehmen können anhand ihres durchschnittlich ausgeprägten Beteiligungsgrades in verschiedene Unternehmenstypen klassifiziert werden, beginnend von U1 (bürokratische Unternehmen) mit dem niedrigsten Beteiligungsgrad bis zu U6 mit dem höchsten Beteiligungsgrad in selbst-

verwalteten Betrieben (s. Tabelle 1). In Tabelle 1 ist die Verteilung der Befragten nach Unternehmenstypen für die Stichprobe dargestellt. Mit Ausnahme des Unternehmenstyps 4b, der tendenziell auch dem Typ 6 zugeordnet werden könnte, ist die Probandenzahl relativ gleichmäßig verteilt. 87 Probanden sind den bürokratischen Unternehmen angehörig, 73 den sogenannten Partnerschaftsunternehmen, 49 arbeiten einem konventionell geführten Belegschaftsunternehmen, 26 in Belegschaftsunternehmen, die sich weitestgehend in Belegschaftsbesitz befinden, 126 Befragte sind in demokratischen Reformunternehmen tätig und 59 in Betrieben, die selbstverwaltet sind. Die Unternehmen stammen aus dem Euregio-Raum Süddeutschland – Österreich – Südtirol/Norditalien.

Für die Beteiligung an der Befragung galten die folgenden Kriterien:

- Unternehmenszugehörigkeit seit mindestens 6 Monaten
- Beschäftigungsmaß mindestens 40 Prozent
- keine Genossenschaftsaufsichtsräte oder sonstigen Genossenschaftsmitglieder oder Kommanditisten, die nicht in der Firma im Arbeitsalltag tätig sind
- keine Temporärarbeiter oder Honorarkräfte, die nur für einen Zeitraum weniger Wochen oder Monate beschäftigt sind
- keine allein auf Außenstellen Arbeitende, die ohne häufigen Kontakt zu anderen Firmenangehörigen tätig sind.

Von den 420 in den Datensatz aufgenommenen Probanden sind gut zwei Drittel Männer (N = 272) und

Tabelle 1: Verteilung der Gesamtstichprobe nach Unternehmenstypen

Typ	Beschreibung	Anzahl	Pb
U1	Bürokratische Unternehmen	8	87
U2	Partnerschaftsunternehmen	3	73
U4a	Konventionell geführte Belegschaftsunternehmen/ Produktivgenossenschaften	3	49
U4b	Konventionell geführte Belegschaftsunternehmen/ Produktivgenossenschaften in überwiegender Belegschaftsbesitz	2	26
U5	Demokratische Reformunternehmen	7	126
U6	Selbstverwaltete Belegschaftsunternehmen/ Produktivgenossenschaften	7	59
		30	420

ein knappes Drittel Frauen ( $N = 133$ ). Rund 66.2 Prozent der Befragten sind im Alter zwischen 30 und 50 Jahren, 21.8 Prozent sind jünger als 30 Jahre und 12 Prozent liegen im Altersbereich über 50 Jahren, wovon lediglich 3 Personen älter als 60 Jahre sind. Die befragten Personen üben ihre aktuelle Tätigkeit seit durchschnittlich 6.5 Jahren aus. Die durchschnittliche Dauer der Organisationszugehörigkeit wird mit 9.4 Jahren angegeben.

237 Personen, das sind 56.6 Prozent der Befragten, geben an aktuell Mitglied eines betrieblichen Mitbestimmungsorgans (Gremien, Ausschüsse, Mitarbeitervertretung, Betriebsrat etc.) zu sein. Rund 43 Prozent der Teilnehmer verfügen über Kapitalanteile an ihrem Unternehmen (s. Tabelle 2).

### *Güte der eingesetzten Verfahren*

Bei den eingesetzten Verfahren handelt es sich um zwei empirisch bereits erprobte Instrumente von Schmitt und Dörfel (1995) und Felfe und Mitarbeitern (2004) sowie zwei neu entwickelte Verfahren von Weber und Mitarbeitern (2004). Die Organisationale-Demokratie-Struktur (Weber, 2004) baut auf dem IDE-

Verfahren aus der europaweiten Industrial Democracy Studie aus den 1980er Jahren auf (Wilpert & Rayley, 1985) und erreicht mit .98 eine sehr zufrieden stellende Reliabilität. Mit dem Verfahren zur Erhebung der soziomoralischen Atmosphäre wurde von den Autoren (Weber et al., 2004) ein Screeningverfahren entwickelt, welches zukünftig auch betrieblichen Praktikern zur Verfügung stehen soll. Mit .90 erreicht die eingesetzte Version ebenfalls eine zufrieden stellende Reliabilität (s. Tabelle 3).

### *Auswertungsmethode*

Der Komplexität des Untersuchungsansatzes wird durch die Anwendung so genannter Strukturgleichungsmodelle Rechnung getragen. Gegenüber anderen multivariaten Analyseverfahren weist die Strukturgleichungsmethode Vorteile auf, die sich vor allem aus der Modellierung der Messfehler ergeben. Damit können auch Kausalbeziehungen auf der Ebene der latenten Variablen geschätzt werden in denen der Einfluss der Messfehler eliminiert wurde. Die Strukturgleichungsmethodik hat zudem den Vorteil, dass sie alle Beziehungen zwischen den Variablen des Modells

*Tabelle 2: Soziodemographische Merkmale Gesamtstichprobe*

	Anz	%		
<b>Geschlecht</b>				
Männer	272	64.8		
Frauen	133	31.7		
<b>Alter</b>				
≤ 30 Jahre	92	21.8		
30 bis 49 Jahre	278	66.2		
≥ 50 Jahre	50	12.0		
<b>Mitglied in einem Mitbestimmungsorgan</b>				
Gesamt	237	56.6		
<i>Davon Männer</i>	160	67.5		
<i>Davon Frauen</i>	72	32.5		
<b>Kapitalanteile</b>				
Gesamt	181	43.1		
<i>Davon Männer</i>	123	70.0		
<i>Davon Frauen</i>	53	30.0		
	MW	SD	Min	Max
Durchschnittliche Dauer Unternehmenszugehörigkeit	9.4	7.5	0.5	35.0
Durchschnittliche Dauer Tätigkeitsausübung	6.5	5.7	0.0	30.0

Tabelle 3: Kennwerte der verwendeten Skalen

		Item Anz.	M	SD	$\alpha$
ODS	Struktur Org. Demokratie (Weber, 2004)	45	3.13	1.27	.98
pG	Prozedurale Gerechtigkeit (Schmitt & Dörfel, 1995)	10	4.22	.95	.92
smA	Soziomoralische Atmosphäre (Weber et al., 2004)	13	4.37	.87	.90
OC	Organisationales Commitment (Felfe et al., 2004)	12	3.88	.89	.82

simultan analysiert. Die Datengrundlage dafür bilden die Varianzen und Kovarianzen der beobachtbaren Variablen.

Als Maße für die Güte des geschätzten Modells steht eine breite Palette an Fit-Indizes zur Verfügung. Von statistischer Seite (Byrne, 2001) gelten bei den Fit-Indizes folgende kritischen Werte: Goodness-of-Fit statistics  $>.90$ ,  $\chi^2$  nicht signifikant, Parsimony-fit-indices  $>.50$  (PNFI, PGFI, CMIN/df, AIC). Als weitere Maße, die für einen akzeptablen Fit sprechen, gelten die Incremental Fit Indices (NFI, TLI  $>.90$ ) sowie RMSEA und RMSR mit  $<.08$ .

### 7 Commitment als Ergebnis von Beteiligung, Fairness und Wertschätzung

Die hier vorgestellte Untersuchung will aufzeigen, ob die Möglichkeit zu demokratischer Teilhabe an betrieblichen und unternehmerischen Entscheidungen die Bindung der Mitarbeitenden an das Unternehmen beeinflusst. Demokratische Teilhabe an Entschei-

dungsprozessen ist per se kein notwendiges Element einer Arbeitsorganisation. Mitbestimmung – sofern sie über das gesetzlich vorgeschriebene Mindestmaß hinausgeht – muss von den Organisationsmitgliedern gewollt und abgesichert werden. Dies erfolgt einerseits durch die Verankerung in der betrieblichen Verfassung (ODS), andererseits aber auch in der entsprechenden Gestaltung von Prozeduren (pG) sowie durch bestimmte Aspekte der Unternehmenskultur (smA).

Bei Aufnahme der Faktoren in das Strukturgleichungsmodell, zeigt sich das Modell wie in Abbildung 1 dargestellt. Die Pfadkoeffizienten folgen den Ergebnissen der bisherigen Ausführungen. Das Modell erreicht relativ zufrieden stellende Fit-Maße. NFI und CFI bewegen sich mit  $.975$  und  $.985$  näher bei  $1.0$  als bei  $.90$ . Der RMSEA liegt mit  $.056$  deutlich unter der kritischen Marke von  $.08$  bzw. bei strenger Prüfung von  $.06$ . Der Chi-Quadrat-Test fällt leider signifikant aus. Byrne (2001) stellt hierfür zwar die Forderung nach Nicht-Signifikanz, weist aber gleichzeitig darauf hin, dass bei größeren Stichproben, wie hier bei

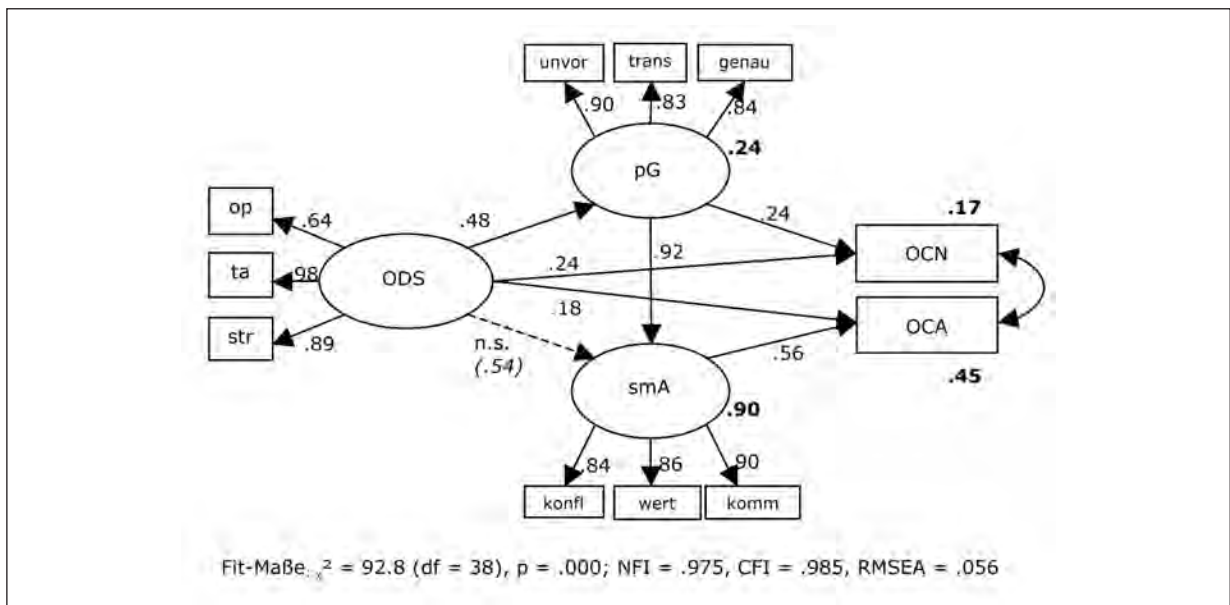


Abbildung 1: Fit-Indizes und Pfadkoeffizienten des Modells

N= 420, die Effekte rasch signifikant werden. Da alle anderen Fit-Indizes die Kriterien ohne weiteres erfüllen, kann die Signifikanz toleriert werden. Insgesamt wird das Modell durch die empirischen Daten gut abgebildet. Die Passung des Modells schlägt sich auch in der aufgeklärten Varianz nieder. Insgesamt werden 62 Prozent der Varianz des Organisationales Commitments aufgeklärt (45 Prozent OCA, 17 Prozent OCN, siehe Abbildung 1).

Bei genauerer Betrachtung lassen sich zwei hauptsächliche Wirkpfade erkennen. Ein wirkkräftiger Weg geht von der Beteiligung (ODS) über die prozedurale Gerechtigkeit (pG) und die sozioralistische Atmosphäre (smA) zum affektiven Commitment (OCA). Zusammen mit einem schwachen, aber signifikanten direkten Effekt von der Organisationales Demokratie Struktur (ODS) auf das affektive Commitment (.18) werden so 45 Prozent der Varianz des affektiven Commitment aufgeklärt (siehe Abbildung 2).

Ein zweiter Pfad zeigt sich ausgehend von ODS über die prozedurale Gerechtigkeit hin zum normativen Commitment (OCN). Ein direkter Effekt lässt sich wieder von ODS auf das normative Commitment hin beobachten. Gemeinsam tragen die Effekte zu einer 17prozentigen Varianzaufklärung des normativen Commitments bei (siehe Abbildung 3).

### 8 Bedeutung der Ergebnisse für die Gestaltung des betrieblichen Alltags

Die Ergebnisse der Untersuchung bieten eine Fülle von Anregungen für vertiefende Analysen und weiterführende Fragestellungen. So ist, um nur ein Beispiel zu nennen, der hohe Zusammenhang zwischen prozeduraler Gerechtigkeit und sozioralistischer Atmosphäre einer Klärung bedürftig (vgl. Schmid, 2009). An dieser Stelle sei der Blick aber stärker auf mögliche

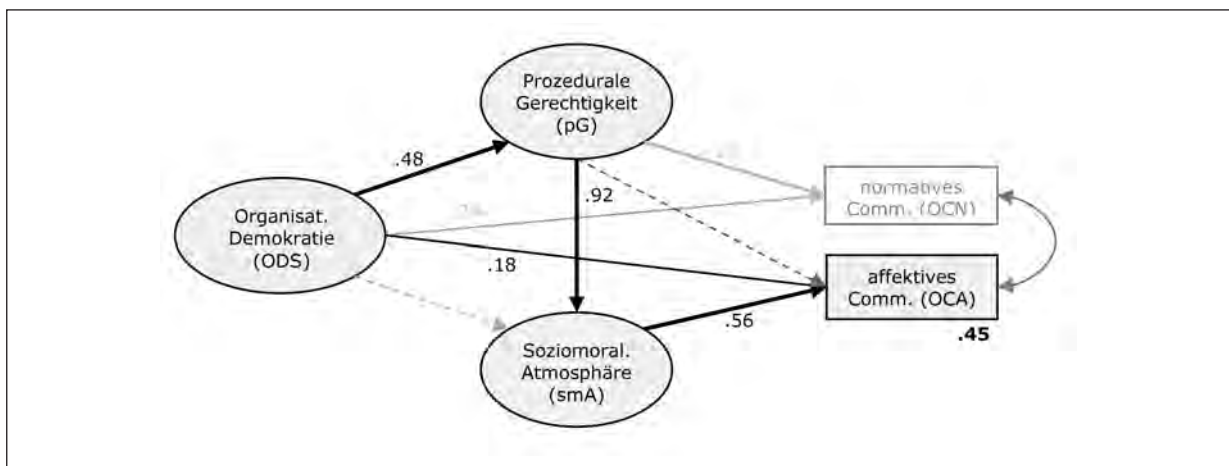


Abbildung 2: Schematische Darstellung Wirkgefüge ODS, pG, smA auf OCA

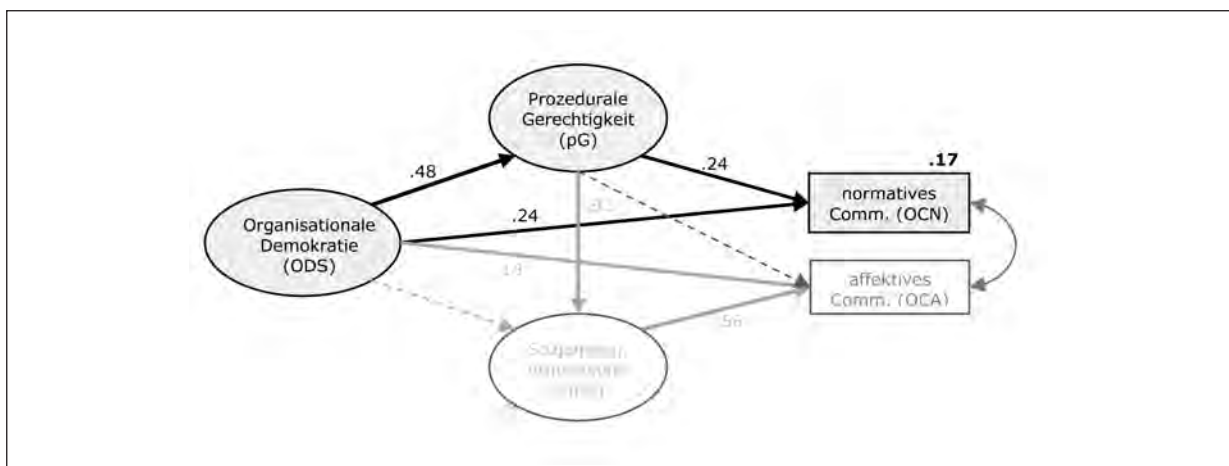


Abbildung 3: Schematische Darstellung Wirkgefüge ODS, pG, smA auf OCN

Ableitungen für den Alltag in Unternehmen gelenkt. Was kann ein Verantwortlicher aus diesen Ergebnissen für sich und seine Aufgaben in der Unternehmensführung und -gestaltung ableiten?

### **Beteiligung**

Zunächst lässt sich festhalten, dass mit zunehmender Beteiligung von Mitarbeitern an betrieblichen Entscheidungen das affektive und das normative Commitment der Unternehmensmitglieder zunehmen. Dies ist dort umso mehr der Fall, wo diese Beteiligungsmöglichkeiten nicht nur fallweise gewährt, sondern in der Verfassung des Unternehmens verankert und in betrieblichen Strukturen abgebildet sind. Für den Unternehmensverantwortlichen kann dies eine Anregung sein, in seinem Unternehmen über neue Beteiligungsformen der Mitarbeiter nachzudenken. Modelle hierfür finden sich sowohl in Deutschland als auch in Österreich und Italien. Das erfolgreichste Projekt diesbezüglich ist mit MONDRAGON aber wohl im spanischen Baskenland zu finden (siehe [www.mcc.es](http://www.mcc.es)).

Deutlich fallen auch die Ergebnisse zum Einfluss beteiligungsorientierter Strukturen und Verfahren auf die Klima- und Verfahrensmerkmale der Organisationen aus. Mit ansteigendem Beteiligungsgrad der Mitarbeiter werden die betrieblichen Verteilungsverfahren zunehmend als fair bewertet. Dies geschieht in den Unternehmen der untersuchten Stichprobe, die sich durch strukturell verankerte Mitbestimmungsmöglichkeiten auszeichnen, konsequenter und für die Mitarbeitenden überzeugender als in bürokratischen Unternehmen. Dies lenkt den Blick auf den bestimmenden Einfluss von Strukturen und Rahmenbedingungen auf die Kultur eines Unternehmens, denn nach Wohland und Wiemeyer (2006) ist Kultur im Wesentlichen „der Schatten der herrschenden Verhältnisse“. So bedeutsam ein wohlwollendes menschliches Miteinander im Unternehmen auch ist: Wertschätzung ersetzt keinesfalls klare Strukturen. Die Rahmenbedingungen, Verhältnisse müssen stimmen, sie stellen eine Notwendigkeit für effizientes Arbeiten dar. Die Kunst moderner Organisationsgestaltung liegt demnach darin, Unternehmensstrukturen zu entwickeln, welche sich durch Klarheit, Transparenz und Verbindlichkeit auszeichnen und damit auch durch Regeln, die dabei aber gleichzeitig den Menschen als wertorientiertes und nach Autonomie strebendes Individuum nicht aus dem Blick lassen, sondern ihm vielmehr die Möglichkeit bieten, seine individuellen Talente zur Lösung der jeweiligen Aufgabenstellung einzubringen. Übertragen auf Organisationsentwicklung bzw. Change Management bedeutet dies für die Führungskraft vor allem an der Veränderung der Verhältnisse zu arbeiten, denn:

*„Kultur ist die unruhige Abbildung der Verhältnisse. Sie provoziert zum Bewahren und zum -> Lernen. Als Abbild (Schatten) entwickelt sie sich nur mit den Verhältnissen, die sie abbildet. Daher kann sich keine Organisation die Kultur aussuchen, die in ihr wirksam ist. Nur wenn sich die Verhältnisse ändern, ändern sich die Erkenntnisse und Erfahrungen, die die Kultur als Gedächtnis aufbewahrt. Es ist daher sinnlos, von einer Organisation eine bestimmte Kultur zu fordern. [...] Weil das Kulturproblem nicht aus der Dummheit oder dem bösem Willen einzelner Personen besteht, kann es durch Belehrung und Ermahnung nicht gelöst werden. (...) Kultur ist der Schatten der herrschenden Verhältnisse. Sie ist -> kausal »hinten« und kann nur indirekt verändert werden. Da sie nicht Ursache, sondern Ergebnis der Verhältnisse ist, braucht sie nicht entwickelt zu werden. Wenn es gelingt, die Verhältnisse zu ändern, verändert sie sich von allein (Wohland & Wiemeyer, 2006, S. 9 ff).“*

### **Fairness**

Mitarbeiter fühlen sich dem Unternehmen gegenüber verpflichtet, das mit ihnen fair umgeht. Fairness wird dort erlebt, wo Mitarbeiter informiert und am Entscheidungsprozess beteiligt werden. Transparenz, Unvoreingenommenheit und Genauigkeit im Prozess der Entscheidungsfindung ermöglichen es Menschen sich zu positionieren und sich mit den Zielen und Ergebnissen des Prozesses zu identifizieren. Führungskräfte können diesen Ergebnissen entnehmen, dass es gut und hilfreich ist, Mitarbeiter konsequent zu informieren, unterschiedliche Interessen und Ziele transparent zu machen und Ergebnisse zu begründen. Dies kann vor allem für so genannte Expertenorganisationen gelten. Dass dieser Weg aufwendig ist, sei unwidersprochen. Er erfordert ein Höchstmaß an Kommunikation und Mensch zugewandtheit. Klaus Doppler spricht in diesem Kontext von Kommunikation als „Mentalmassage“ und unterstreicht damit die Bedeutsamkeit der Einbindung von Betroffenen in den Entscheidungsprozess und die kontinuierliche Bearbeitung von potentiellen Widerständen durch intensive Kommunikation (vgl. Doppler & Lauterburg, 2002).

### **Wertschätzung**

Moderne Organisationen von heute sind gefordert, ihre Mitglieder zu hegen und zu pflegen. Kein Unternehmen kann Identifikation und Commitment erwarten, das seinen Mitarbeitenden nicht Fürsorge und Wertschätzung entgegen bringt. Dies wird durch den deutlichen Effekt der soziomoralischen Atmosphäre in Richtung affektives Commitment unterstrichen. Führungskräfte können daraus entnehmen, dass neben strukturellen und prozessualen Faktoren letztlich immer auch noch das persönliche Verhalten und Handeln



gegenüber den Mitarbeitern deren emotionale Reaktion gegenüber dem Unternehmen mitbestimmen und prägen. Hierfür ist es hilfreich und notwendig zwischen Person und Aufgabe zu unterscheiden. Wertschätzung eines Menschen bedeutet dann u.a. auch kritische Rückmeldungen zu seinem Handeln bzw. seiner Arbeit nicht zu verwehren, sondern im Dialog sowohl Lob als auch Kritik einzubringen. Wertschätzendes Miteinander ist dort möglich, wo die Spannungsfelder der *conditio humana* anerkannt und akzeptiert werden und das bedeutet, keiner ist in allem perfekt und keiner ist in allem unfähig. Menschliches Sein ist per se auf Entwicklung angelegt. Ein Unternehmen mit einer wertschätzenden Grundhaltung wird daher einen besonderen Schwerpunkt auf eine gezielte Personalentwicklung legen und im Sinne der „lernenden Organisation“ an einer positiven Fehlerkultur arbeiten.

### Conclusio

Beteiligung, Fairness und Wertschätzung – drei Signaturen für eine Möglichkeit der Unternehmensgestaltung, die zwar von allen Beteiligten mehr abverlangt als nur ihren Job zu machen, dafür aber überzeugte, weil entschiedene Mitarbeiter hervor bringt, die sich mit den Zielen und Werten des Unternehmens identifizieren und sich dem gemeinsamen Ergebnis verpflichtet fühlen. Dazu ist es allerdings notwendig mehr als nur an den Schatten im Unternehmen zu arbeiten.

### Literatur

- Alexander, S. & Ruderman, M. (1987). *Understanding attitudes and predicting social behavior*. New York: Englewood Cliffs.
- Buber, M. (1995). *Ich und Du*. Stuttgart: Reclam.
- Byrne, B. M. (2001). *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming*. Hillsdale, NJ: Erlbaum.
- Cohen-Carash, Y. & Spector, P. E. (2001). The role of justice in organizations: a meta analysis. *Organizational Behaviour and Human Decision Processes*, 82 (2), 278-321.
- Colquitt, J. A., Conlon, D. E., Wesson, M. J. Porter, C. O. L. H. & Ng, K. Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology*, 86, 425-445.
- Doppler, K. & Lauterburg, C. (2002). *Change Management. Den Unternehmenswandel gestalten*. Campus: Frankfurt a. Main.
- Dörfel, M. & Schmitt, M. (1997). Procedural injustice at the workplace, sensitivity to befallen injustice, and job satisfaction. *Berichte aus der Arbeitsgruppe „Verantwortung, Gerechtigkeit, Moral“*, Nr. 105. Trier: Universität Trier, Fachbereich I – Psychologie.
- Eisenberger, R., Armeli, S., Rexwinkel, B., Lynch, P.D. & Rhoades, L. (2001). Reciprocation of Perceived Organizational Support. *Journal of Applied Psychology*, 86, 42-51.
- Emery, F. & Thorsrud, E. (1982). *Industrielle Demokratie*. Bern: Huber.
- Felfe, J., Six, B., Schmoock, R., & Knorz, C. (2004). Fragebogen zur Erfassung von affektivem, kalkulatorischem und normativem Commitment gegenüber der Organisation, dem Beruf/der Tätigkeit und der Beschäftigungsform (COBB). In A. Glöckner-Rist (Eds.), *ZUMA-Informationssystem. Elektronisches Handbuch sozialwissenschaftlicher Erhebungsinstrumente*. ZIS Version 8.00. Mannheim: Zentrum für Umfragen.
- Foley, J. R. & Polanyi, M. (2006). Workplace Democracy: Why Bother? *Economic and Industrial Democracy*, 27 (1), 173-191.
- Folger, R. & Konovsky, M. (1989). Effects of procedural and distributive justice on reactions to pay raise decisions. *Academy of Management Journal*, 32, 115-130.
- Gallup (2005). *Engagement Index 2004*. Studie zur Messung der emotionalen Bindung von Mitarbeiterinnen. The Gallup Organisation.
- Greenberg, J. (1990). Employee theft as a reaction to underpayment inequity: The hidden costs of pay cuts. *Journal of Applied Psychology*, 75, 561-568.
- Hacker, W. (1986). *Arbeitspsychologie*. Bern: Huber.
- Harrison, J. S. & Freeman, R. E. (2004). Is organizational democracy worth the effort? Special Topic: Democracy in and around Organizations. *Academy of Management Executive*, 18, 49-55.
- Herzberg, F., Mausner, B. & Snyderman, B. (1959). *The Motivation to Work*. New York: Wiley.
- Hoff, E.-H. & Lempert, W. (1990). Kontroll- und Moralbewußtsein im beruflichen und privaten Lebensstrang von Facharbeitern. In E.-H. Hoff (Hrsg.), *Die doppelte Sozialisation Erwachsener. Zum Verhältnis von beruflichem und privatem Lebensstrang* (S. 125-154). Weinheim: Juventa.
- IDE International Research Group (1981). *Industrial Democracy in Europe*. London: Oxford University Press.
- Kohlberg, L. (1996). *Die Psychologie der Moralentwicklung*. Frankfurt/M.: Suhrkamp.
- Kohlberg, L., Wassermann, E. & Richardson, N. (1978). Die Gerechte Schul-Kooperative. Ihre Theorie und das Experiment der Cambridge Cluster School. In G. Portele (Hg.), *Sozialisation und Moral: neuere Ansätze zur moralischen Entwicklung und Erziehung* (S. 215-259). Basel: Beltz.

- Lempert, W. & Corsten, M. (1997). Soziale Bedingungen der Entwicklung moralischer Orientierungen im Beruf. *Zeitschrift für Erziehungssoziologie und Sozialisationsforschung*, 17, 339-355.
- Leventhal, G. S. (1980). What should be done with equity theory? New approaches to the study of fairness in social relationships. In K. Gergen, M. Greenberg & R. Willis (Eds.), *Social Exchange: Advances in Theory and Research* (pp. 167-218). New York: Springer.
- Lind, E. A. & Tyler, T. R. (1988). *The social psychology of procedural justice*. New York: Plenum Press.
- Masterson, S., Lewis, K., Goldman, B. M. & Tyler, M. S. (2000). Integrating justice and social exchange: The differing effects of fair procedures and treatment on work relationships. *Academy of Management Journal*, 43, 738-748.
- Matiasko, W. & Weller, I. (2005). Commitment als Ressource. Beitrag zur Tagung „Nachhaltigkeit von Arbeit und Rationalisierung“, TU Chemnitz, 01/2005.
- McFarlin, D. B. & Sweeney, P. D. (1992). Distributive and procedural justice as predictors of satisfaction with personal and organizational outcomes. *Academy of Management Journal*, 35, 626-637.
- Meyer, J. P. & Allen, N. J. (1997). *Commitment in the Workplace. Theory, Research, and Application*. Thousand Oaks: Sage Publications.
- Meyer, J. P. & Allen, N. J. (1991). A three-component conceptualization of Organizational Commitment. *Human Resource Management Review*, Vol. 1, No. 1, 61-89.
- Moorman, R. H. (1991). Relationship between organizational justice and organizational citizenship behaviors: Do fairness perceptions influence employee citizenship? *Journal of Applied Psychology*, 76, 845-855.
- Moser, K. (1997). Commitment in Organisationen. *Zeitschrift für Arbeits- und Organisationspsychologie*, 41 (4), 160-170.
- Organ, D. W. & Moorman, R. H. (1995). Fairness and Organizational Citizenship Behavior: What are the Connections? *Social Justice Research*, 6, 5-18.
- Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Belknap Press of Harvard University Press.
- Rhodes, S. R. & Steers, R. M. (1981). Conventional vs. worker-owned organizations. *Human Relations*, 12, 1013-1035.
- Schmid, B. (2009). *Beteiligungsorientierung in Unternehmen*. Innsbruck: innsbruck university press.
- Schmid, B. E. (2004). *Betroffene Beteiligte. Organisationale Demokratie, Commitment und prosoziale Handlungsorientierung im Krankenhaus* (unveröff. Diplomarbeit). Innsbruck: Leopold-Franzens-Universität, Naturwissenschaftliche Fakultät.
- Schmidt, K. H., Hollmann, S. & Sodenkamp, D. (1998). Psychometrische Eigenschaften und Validität einer deutschen Fassung des ‚Commitment‘-Fragebogens von Allen und Meyer (1990). *Zeitschrift für Differentielle und Diagnostische Psychologie*, 19 (2), 93-106.
- Schmitt, M. & Dörfel, M. (1995). *Verfahrensgerechtigkeit, dispositionelle Ungerechtigkeitsensibilität und Arbeitszufriedenheit im betrieblichen Kontext* (unveröff. Dipl.-Arbeit). Trier: Fachbereich 1 - Psychologie der Universität Trier.
- Towers Perrin (2006). *Global Workforce Study. Was Mitarbeiter bewegt zum Unternehmenserfolg beizutragen – Mythos und Realität*.
- Weber, W. G., Unterrainer, C. & Schmid, B. E. (2009). Organizational Democracy, Sociomoral Atmosphere, and Prosocial and Community-Related Value Orientations of Employees. *Journal of Organizational Behavior*, 30.
- Weber, W. G., Iwanowa, A. N., Schmid, B. E. & Unterrainer, C. (2006). *ODEM. Abschlussbericht zum Forschungsprojekt Organisationale Demokratie* (unveröff.). Innsbruck: Universität Innsbruck, Institut für Psychologie.
- Weber, W. G., Iwanowa, A. N., Schmid, B. E. & Unterrainer, C. (2004). *ODEM. Erster Zwischenbericht für die bm:bwk Projektträgerschaft* >node< (unveröff.). Innsbruck: Universität Innsbruck, Institut für Psychologie.
- Weber, W. G. (2004). *Entwicklung eines Verfahrens zur Erfassung der soziomoralischen Atmosphäre* (unveröff.). Innsbruck: Universität Innsbruck, Institut für Psychologie.
- Weber, W. G. (1999). *Organisationale Demokratie - Anregungen für innovative Arbeitsformen jenseits bloßer Partizipation?* *Zeitschrift für Arbeitswissenschaft*, 53 (4), 270-281.
- Wilpert, B. & Rayley, J. (1983). *Anspruch und Wirklichkeit der Mitbestimmung*. Frankfurt/M.; New York: Campus Verlag.
- Wohland, G. & Wiemeyer, M. (2006). *Denkwerkzeuge für dynamische Märkte. Ein Wörterbuch*. MV Wissenschaft.
- Korrespondenz-Adresse:  
 Dr. Birgit E. Schmid  
 Marienhaus Klinikum im Kreis Ahrweiler,  
 Bad Neuenahr-Ahrweiler  
 Dahlienweg 5  
 D-53474 Bad Neuenahr-Ahrweiler  
 b.schmid@maw.marienhaus-gmbh.de

# Komplexe Strukturen aufspüren: Faktorenanalyse mit Variminrotation

Suitbert Ertel

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

## ZUSAMMENFASSUNG

Der vorliegende Artikel knüpft an einer Kritik am ‚Dogma Simple Structure‘ an (Ertel, 2009). Ein neues Rotationsverfahren für extrahierte Faktoren wird vorgestellt (Varimin), welches die latenten Kovarianzquellen ermitteln will, die den manifesten Variablen einer Domäne zugrunde liegen. Varimin optimiert das durch die Initialstruktur einer Faktorenanalyse schon entworfene, aber oft noch korrekturbedürftige Modell multifunktionaler Bedingungsverhältnisse. Die konventionell verwendeten Rotationsverfahren, mit denen die latenten Kovarianzquellen als Einfachstruktur modelliert werden sollen, voran das Standardverfahren Varimax, werden bemängelt, weil sie die zugrunde liegende reale Komplexität der latenten Bedingungsverhältnisse verdecken, anstatt sie aufzuklären. Die neue methodologische Ausrichtung wirft Fragen auf, von denen fünf diskutiert werden. Gefragt wird: (1) Wie lassen sich Varimin-transformierte Faktoren deuten? (2) Wird die Komplexität der Kovarianzquellen nicht zumeist mit der Initiallösung schon hinreichend repräsentiert? (3) Lassen sich faktorielle Simple Structure-Lösungen nicht doch befriedigend interpretieren, wie könnte man sie sonst ständig verwenden? (4) Enthält die regelmäßig anzutreffende Bipolarität von Varimin-Faktoren substantielle Information? (5) Sind mithilfe einer Komplexstruktur-orientierten Faktorenanalyse auch methodenbedingte Einflüsse erfassbar? Die Klärung dieser Fragen wird durch Anwendung der Varimin-Transformation auf acht publizierte und zwei neu generierte Datensätze unterstützt. Die Transformation zur Komplexstruktur, mit welcher latente Kovarianzquellen der manifesten Variablen aufgespürt werden (durch Varimin), wird der Transformation zur Einfachstruktur (durch Varimax) gegenüber gestellt. Es zeigt sich, dass Varimax zur Aufklärung latenter Bedingungen keinen nennenswerten Beitrag leistet. Lediglich zum Clustern manifester Variablen sind Simple Structure-Transformationen geeignet. Auf die Grenzen der methodologischen Innovation wird hingewiesen.

## Schlüsselwörter

Faktorenanalyse – Rotation – Varimax – Varimin – Einfachstruktur – Komplexstruktur

## ABSTRACT

Revealing complex structures. Factor analyses with Varimin rotation.

This paper examines a continuation of the critique of „simple structure“ (Ertel, 2009) in which the rotation procedure „Varimax“ which is commonly used to generate simple structure is replaced with „Varimin“. Varimin aims at manifesting latent complex structures. It optimises the model of complexity which, although delineated by initial unrotated structures, is further improved. The new method raises a number of issues of which five are discussed at length. (1) How can varimin factors be interpreted? (2) Do latent sources of covariance not appear sufficiently complex with initial solutions? (3) Are simple structure solutions in common practice not adequately interpretable? (4) How to interpret the commonly encountered bipolarity of Varimin factor loadings? (5) Is factor analysis with complex structure transformation applicable to data affected by method factors? Ten empirical applications of Varimin transformation serve as practical examples. Particular features of transformation to complex structure, revealing latent sources of covariance (by Varimin), are elucidated by comparing pertinent results with those obtained from transformations to simple structure (by Varimax). Varimax will remain useful for clustering objectives. Attention is also drawn to limitations of the new methodical innovation.

## Keywords

Factor analysis – factor rotation – varimax – varimin – simple structure – complex structure

## 1 Einleitung

Der vorliegende Artikel ist die Fortsetzung eines früheren (Ertel, 2009), in dem eine Kritik an der exploratorischen Faktorenanalyse begründet wurde. Es wurde dargelegt, dass das Hauptproblem der Faktorenanalyse, die geometrische Rotation bzw. algebraische Transformation extrahierter Faktoren, durch Thurstones Modell der Einfachstruktur nicht gelöst, sondern verschärft wurde. Die manifesten Variablen einer empirischen Domäne, so wurde argumentiert, resultieren generell aus *multiplen* Varianz- bzw. Kovarianzquellen. Latente Bedingungen, die die einzelnen Variablen hervorrufen, bilden keine Einfachstruktur, sie bilden fast ausnahmslos multifaktorielle Strukturen. Die monofaktorielle Bedingtheit manifeste individueller Variablen, die mit dem Prinzip der Einfachstruktur vorausgesetzt wird, darf nicht generell vorausgesetzt werden. Das Prinzip wurde als fehlerhaft verworfen.

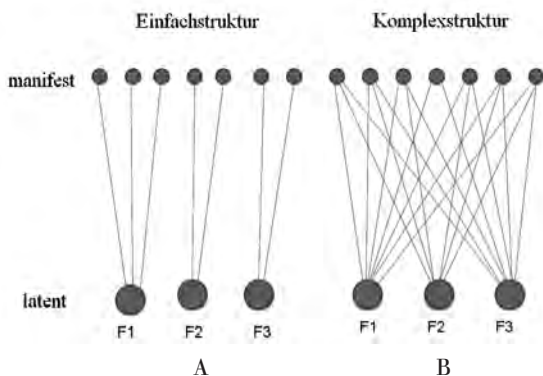


Abbildung 1: Einfachstruktur (A) und Komplexstruktur (B) des Verhältnisses zwischen latenten Kovarianzquellen und manifesten Variablen

Die Abbildungen 1A und 1B veranschaulichen den Unterschied zwischen Einfach- und Komplexstruktur. In beiden Fällen ist auszugehen von der Ebene manifeste Variablen, die interkorreliert werden mit der Absicht, durch Faktorisierung der Korrelationsmatrix die zugrunde liegenden Faktoren aufzuklären. Faktoren sind die Quellen der Kovarianz, die bei den untersuchten manifesten Variablen ermittelt wird. Das einfachstrukturelle Modell Thurstones unterscheidet sich vom hier befürworteten komplexstrukturellen Modell in Folgendem: Eine Rotation der Faktoren zur Einfachstruktur (Abbildung 1A) ordnet einzelnen manifesten Variablen jeweils nur einen extrahierten Faktor zu bzw. so wenig Faktoren wie möglich. Eine Rotation zur Komplexstruktur dagegen (Abbildung 1B) ist darauf aus, individuelle Variablen mit möglichst vielen extrahierten Faktoren hypothetisch zu verknüpfen.

Nach dieser Konzeption wird angenommen, dass die meisten Faktoren in der Regel zur Kovarianz einzelner Variablen einen – je nach Variable verschieden großen – Beitrag leisten.

Unter den Rotationsverfahren, die zur Einfachstruktur führen sollen, werden im Forschungsalltag meist die orthogonalen gewählt, oblique Verfahren nur selten, und von den orthogonalen Verfahren wird fast ausschließlich Kaisers Varimax-Rotation eingesetzt (s. Abbildung 2).

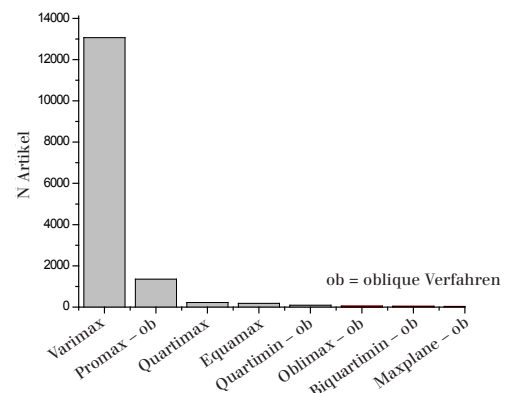


Abbildung 2: Anzahl von Artikeln, in denen auf Faktorenanalyse-Verfahren verwiesen wird. Ergebnis einer Internet-Suche nach Rotationsbezeichnungen in SCIENCE DIRECT (Elsevier, 2008)

Nach den Überlegungen des vorangehenden Artikels sind die Rotationsverfahren, die eine Einfachstruktur erreichen wollen (Abbildung 2), durchweg fehlerhaft. Das Modell der Einfachstruktur verhüllt die in der Initiallösung oft schon erkennbare Komplexität latenter Bedingungen, anstatt diese herauszuarbeiten. Dabei ist anzumerken, dass in der Initiallösung ein befriedigendes Komplexitätsmodell nur beim Sonderfall zweifaktorieller Lösungen zu finden ist. Für Datensätze mit mehr als zwei Faktoren wird mit der Initiallösung die Komplexität der Kovarianzquellen individueller Variablen in der Regel nicht optimal manifestiert. Der Grund ist, dass bei der Extraktion eines Faktors die Varianz, die die Matrix der Korrelationen – bzw. nach einer Faktorenextraktion die jeweilige Restmatrix – enthält, für den betreffenden Faktor überoptimal ausgeschöpft wird. Dies führt zu einem mehr oder weniger großen Überhang an Kommunalität bei den früher extrahierten Faktoren und zu einer entsprechenden Benachteiligung der später extrahierten Faktoren (Gorsuch, 1974, p. 161).<sup>1</sup> Zur optimalen Manifestation der Komplexität der Kovarianzquellen wurde deshalb eine Faktorenrotation auch für notwendig gehalten. Doch soll diese das Gegenteil von dem leisten, was eine Rotation zur Einfachstruktur zum Ergebnis hat.

Aus diesem Grunde wurde *Varimin* konzipiert, ein neues Rotationsverfahren des Verfassers. Varimin ist die Umkehrform von Varimax. Mit Varimax wird durch paarweises Drehen der Faktor-Koordinaten die Varianz der quadrierten Faktorladungen pro Faktor *gesteigert*. Der Vorgang wird durch Iterationen so lange wiederholt, bis sich die Summe der Ladungsvarianzen für die Gesamtheit der Faktoren nicht mehr steigern lässt. Das Kriterium V (s. Gleichung 1), das Varimax-Kriterium, wird maximiert.

$$(1) \quad V = n \sum_{p=1}^m \sum_{j=1}^n (b_{jp} / h_j)^4 - \sum_{p=1}^m \left( \sum_{j=1}^n b_{jp}^2 / h_j^2 \right)^2$$

h = Kommunalität eines Faktors

b = Ladungshöhe eines Faktors

p = Laufindex für Faktoren von 1 bis m

j = Laufindex für Variablen von 1 bis n

Für das *Varimin*-Verfahren werden die Koordinatendrehungen mit dem Ziel einer *Varianz-Verminderung* vorgenommen, iterativ so lange, bis die Summe der Ladungsquadrate, das Kriterium V, sich nicht mehr vermindern lässt.

Angenommen eine initiale Struktur der Abbildung 3A liege vor. Die Anwendung von Varimax transformiert die Struktur 3A in die Struktur 3C. In 3C laufen die Koordinaten durch die Punktwolken, für die einzelnen Faktoren wird die Summe ihrer Ladungsquadrate maximiert. Die Anwendung von Varimin transformiert die Struktur 3A in die Struktur 3B. Der Abstand der Variablenwolken von den Koordinaten wird so weit wie möglich vergrößert. Die Summe der Ladungsquadrate wird damit für die Faktoren minimiert.

Mit der Einführung von Varimin zur Faktorentransformation stellen sich neue Verfahrensfragen. Im Folgen-

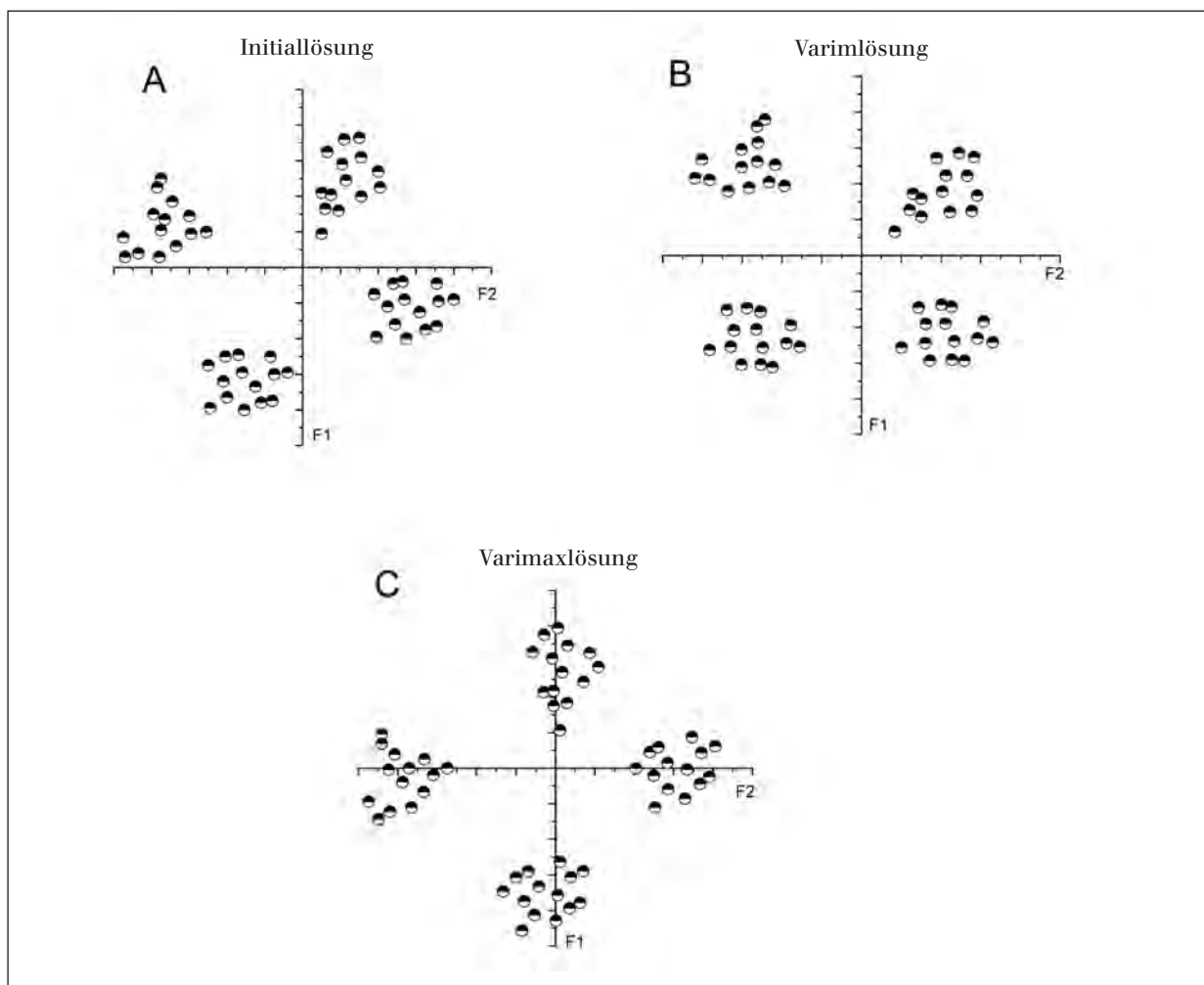


Abbildung 3: Zweifaktorielle Initiallösung mit fingierten Variablen (A), die Varimin-rotiert (B) und Varimax-rotiert wurden (C)

den werden die fünf dringendsten aufgegriffen, wobei empirische Lösungsbeispiele herangezogen werden, die zur Klärung beitragen sollen.

## 2 Fragen und empirische Belege

### Frage I

#### Wie lassen sich Varimin-transformierte Faktoren deuten?

Die Rotation einer Faktorenlösung zur Einfachstruktur wird herkömmlicherweise deshalb für notwendig gehalten, weil die initiale Faktorenlösung komplex ist und Faktoren angeblich nicht gedeutet werden können, wenn die manifesten Variablen, auf die man sich bei der Faktorendeutung stützen muss, mehrfaktoriell geladen sind (Guilford, 1952, p. 27, Burroughs & Miller, 1961, p. 37, Überla 1971, p. 175, Gorsuch, 1974, p. 162, Comrey, 1978, p. 653 f., Reise et al., 2000, p. 292).

Die Behauptung, dass faktorielle Komplexität der Variablen eine Deutung der Faktoren verhindere oder erschwere, lässt sich jedoch nicht aufrecht erhalten, wenn man das Konzept der ‚distinctive features‘ methodisch nutzt. Das Konzept wurde in der Phonologie von Jacobson & Halle (1956) entworfen und fand in dieser Disziplin als ein gewichtiger methodischer Fortschritt weithin Anerkennung. In generalisierter Form ist seine Anwendung auf andere als phonologische und linguistische Inhalte übertragbar.

Die Details: Phoneme lassen sich mithilfe von ‚features‘ (Merkmale) unterscheiden, sofern diese eine distinktive Funktion für die untersuchte Sprache besitzen. Für jedes Phonem hat man auf dem Wege der Selbstbeobachtung und der objektiven Messung drei Klassen von Merkmalen als funktional relevant erkannt, welche die Artikulation der Phoneme spezifizieren: die Artikulationsdauer (kurz vs. lang)<sup>2</sup>, die Sonorität oder Stimmeteiligung (stimmhaft, stimmlos) und den Artikulationsort (bilabial, labio-dental usw. bis uvular). Hier lässt sich die Begrifflichkeit der Faktorenanalyse anwenden, denn man kann sagen: Phoneme sind manifeste Einheiten, distinktive Merkmale sind die zugrunde liegenden (‚latenten‘) Varianzquellen. Jedes Phonem ist durch drei Varianzquellen zu beschreiben, zum Beispiel ist [b] ein kurz artikuliertes (plosives), ein stimmhaftes und bilabiales (mit den Lippen gebildetes) Phonem.

Zu fragen bleibt, wie z. B. beim Phonem [b] die drei Merkmale methodisch aufgespürt werden können. Dies geschieht durch den *Minimalpaar-Vergleich* (*minimal pair comparison*). Zur Paarbildung für [b] ist eine ähnliche andere Einheit der gleichen Domäne mit heranzuziehen. Unter den Phonemen der deutschen Sprache findet sich auch das [p], das mit dem [b] zwei der oben genannten Merkmale teilt (Artikulationsort

und -dauer). Nur hinsichtlich der Stimmhaftigkeit gibt es einen markanten Unterschied. Wenn die Phoneme der deutschen Sprache als Variablen z. B. für eine Beurteilung ihrer Ähnlichkeiten/Verschiedenheiten verwendet werden, dann sollte eine Faktorenanalyse der Phonem-Interkorrelationen, welche Ähnlichkeitsmaße darstellen, mit anschließender Variminrotation Faktoren erkennen lassen, die den bekannten drei distinktiven Merkmalkategorien entsprechen. Die Phoneme [b] und [p] könnten z. B. bei den Faktoren  $F_1$  und  $F_2$  gleiche Ladungen aufweisen, bei  $F_3$  aber müssten die Ladungen markant auseinander treten. Die Aufgabe des Faktorenanalytikers bestünde lediglich darin, den ‚minimalen‘ Unterschied zwischen [b] und [p] für  $F_3$  zu deuten:  $F_3$  wäre dann als Effekt der Stimmhaftigkeit/Stimmlosigkeit anzusehen. Die Deutung würde sich sichern lassen, wenn weitere minimale Paare aus dem Datensatz gebildet werden, z. B. die Paare [d] vs. [t] sowie [g] vs. [k], bei denen der in gleicher Weise zu deutende minimale Unterschied für  $F_3$  auftritt wie bei [b] und [p], während für  $F_1$  und  $F_2$  bei den Paaren [d] vs. [t] sowie [g] vs. [k] wieder unterschiedslose Ausprägungen vorliegen würden, obgleich für diese übrigen Phonempaare beliebige andere  $F_1$  und  $F_2$ -Ausprägungen vorliegen dürfen (nicht müssen) als für [b] und [p].<sup>5</sup>

### Empirischer Beitrag 1:

#### Beurteilung von Phonem-Ähnlichkeiten

Diese Überlegungen wurden für ein Experiment verwendet, an dem zwei deutschsprachige Studierende teilnahmen, eine Psychologiestudentin mit offensichtlicher Sprachbegabung und eine Linguistikstudentin höheren Semesters, zu deren Ausbildung Phonologie gehörte. Die Probandinnen hatten 10 Phoneme der deutschen Sprache ([b], [d], [f], [g], [k], [m], [n], [p], [t], [v] auf bipolaren Likert-Skalen mit sieben Stufen (Mittelstufe = 0) nach Ähnlichkeit zu beurteilen ([v] wird ausgesprochen wie „w“ in z. B. ‚wenn‘). Als Skalenpole wurden die zu beurteilenden 10 Phoneme paarweise in allen Kombinationen (45 Kombinationen) verwendet, die zufällig aufeinander folgten, z. B. waren die Phoneme auf den Skalen

[d] 3--2--1--0--1--2--3 [m]

[b] 3--2--1--0--1--2--3 [d]

und weiteren 43 Skalen nach Ähnlichkeit einzustufen. Wenn das zu beurteilende Phonem gleichzeitig einen Skalenpol definiert, läuft das auf ein Identitätsurteil hinaus, z. B. ist [d] auf der Skala

[d] 3—2—1—0—1—2—3 [m] dem linken Skalenpol maximal ‚ähnlich‘ (mit ihm identisch). In solchen Fällen war laut Instruktion immer die Stufe 3 (maximale Ähnlichkeit) anzukreuzen.

Die zehn Urteilsprofile mit je 45 Urteilen wurden sodann für jede Probandin individuell interkorreliert

Tabelle 1: Korrelationen der Faktorladungen mit objektiven Rangwerten sowie die Tucker-Kongruenzen der Faktoren von zwei Probandinnen

Faktoren	Korrelationen der Faktorladungen mit objektiven Rangwerten			Kongruenz der Faktoren der Vpn	
	Vp Psych	Vp Lingui	beide Vpn		
F1	kurz-lang	.972	.990	.996	.945
F2	stimmhaft-stimmlos	.917	.957	.966	.936
F3	Artikulationsort	.671	.976	.922	.574

und einer Principal Component-Analyse (PCA) unterzogen. Die ersten drei extrahierten Faktoren (= Komponenten) wurden für die Varimin-Rotation berücksichtigt, da drei interpretierbare Faktoren erwartet wurden (Eigenwerte der ersten fünf Faktoren: 2.25, 1.58, 1.40, 1.12, 1.05 für Vp 1 und 4.86, 1.68, 1.09, 1.05, 0.59 für Vp2).

Die Güte der Varimin-Rotation wird danach bemessen, wie gut die Faktorladungen mit den zu erwartenden Klassifizierungen übereinstimmen. Nach phonologischer Klassifizierung haben die folgenden Phoneme eine kurze Artikulationsdauer: [b], [d], [g], [k], [p], [t] (man nennt sie auch ‚plosiv‘), die Phoneme [f], [m], [n], [v] werden mit anhaltendem Luftstrom, also länger artikuliert). Die Phoneme [b], [d], [g], [m], [n], [v] sind stimmhaft, stimmlos sind [f], [k], [p], [t]. Während sich nach Artikulationsdauer und Stimmbeteiligung eine bipolare Unterteilung ohne zusätzliche Differenzierung ergibt, liegen beim Artikulationsort vier Stufen vor, von der Artikulation ‚weit vorn‘ bis ‚weit hinten‘ sind aufzuführen (1) die bilabialen Phoneme [b], [m], [p], (2) die labiodentalen Phoneme [f], [v], (3) die alveolaren [t], [d], [n], und (4) die velaren [k], [g]. Die Ladungen der aus den Urteilsdaten der Probandinnen ermittelten Varimin-rotierten Faktoren werden sodann punktbiserial korreliert mit den linguistischen Variablen Artikulationsdauer (kurz = 1 und lang = 2) und Stimmhaftigkeit (stimmlos = 1 und stimmhaft = 2). Mit den vier ordinalen Stufen des Artikulationsortes (ganz vorn, 1, bis ganz hinten, 4) werden die Faktorladungen Produktmoment-korreliert.

Die Korrelationen zeigt Tabelle 1. Man sieht, dass mit Ausnahme der Korrelation für den Artikulationsort bei der Psychologie-Probandin ( $r = .671$ ) die Korrelationen sehr hoch sind, sie liegen über .90.  $F_1$  repräsentiert bei beiden Probandinnen die Artikulationsdauer,  $F_2$  repräsentiert die Stimmhaftigkeit. Offenbar hat die Linguistik-Studentin aufgrund ihrer Ausbildung ein

feineres Empfinden erworben, was die Wahrnehmung des Artikulationsortes im Mund betrifft (die Tucker-Kongruenz für  $F_3$  bei den beiden Versuchsteilnehmerinnen beträgt nur .574. Da aber die Ladungen bei  $F_3$  für die Linguistik-Studentin wesentlich höher ausfallen als bei der Psychologie-Studentin, ergibt sich nach einer Mittelung der beiden Datensätze mithilfe von Fishers Z-Transformation auch für  $F_3$  eine Korrelation mit dem Artikulationsort von  $r = .922$ , so dass weiterhin die Ergebnisse der vereinten beiden Datensätze für alle drei Faktoren dargestellt werden dürfen.

Abbildung 4 zeigt die Varimin-Ergebnisse. Positive Faktorladungen werden durch dunkle Kreise, negative durch hellere Kreise dargestellt, die variablen Kreisflächen entsprechen den variablen Ladungshöhen. Eine Null-Ladung würde als ausdehnungsloser Punkt wiedergegeben. Wäre die Deutung der drei Faktoren nicht schon durch die Korrelationen mit Expertenurteilen erfolgt, hätte man für  $F_1$  die minimalen Paare [t] vs. [f] und [d] vs. [n] bilden können, wodurch die Artikulationsdauer als distinktives Merkmal in Erscheinung tritt. Bei  $F_2$  wären die minimalen Paare ‚Stimmbeteiligung ja‘ vs. ‚Stimmbeteiligung nein‘ aufgefallen: [b] vs. [p], [d] vs. [t], [g] vs. [k] und [v] vs. [f]. Bei  $F_3$  hätten die minimalen Paare [p] vs. [k] und [b] vs. [g] den vorn-hinten Kontrast des Artikulationsortes manifestiert.

Das Ergebnis der Varimax-Rotation der Faktoren wird ohne Abbildung mitgeteilt, das der Linguistik-Probandin auch nur, weil eine Deutung weniger schwierig ist als das der Psychologie-Probandin. Durch Varimax- $F_1$  werden bei der Linguistik-Probandin die *langen stimmhaften* Phoneme [m], [n] und [v], die hier ein positives Vorzeichen haben, geclustert, mit negativem Vorzeichen gruppieren sich bei  $F_1$  die *kurzen stimmlosen* Phoneme [t] und [k].  $F_2$  gruppiert mit positivem Vorzeichen die *kurzen stimmhaften* Phoneme [g] und [d].

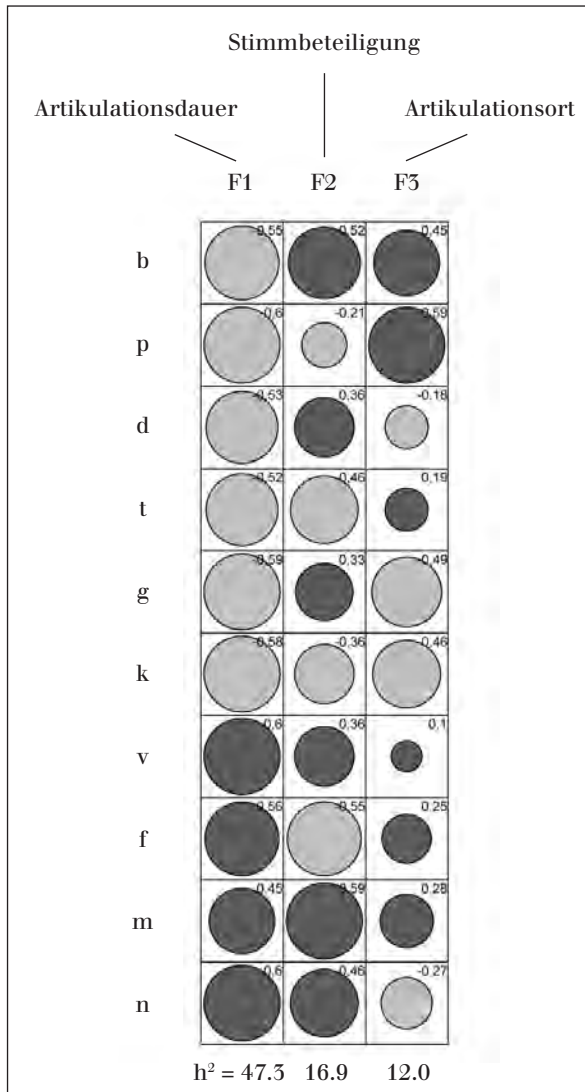


Abbildung 4: Varimin-transformierte Faktorladungen von zehn deutschen Phonemen (aufgrund von Ähnlichkeitsbeurteilungen)

Eine negative  $F_2$ -Ladung hat das *lange stimmlose* [f]. Der unipolare Faktor  $F_3$  gruppiert die bilabialen plosiven Phoneme [p] und [b]. Es ist offenkundig, dass die Varimax-Rotation zwar einander ähnliche Phoneme clustert (die mit Plus-Ladung) und auch die jeweils gegensätzlichen Cluster zusammen bringt (die mit Minus-Ladung). Doch an solchen Clustern, zu denen immer mehrere phonetische Merkmale beitragen, kann man kaum interessiert sein. Denn gesucht werden die latenten Quellen der Kovarianz. Bei der Herstellung von Einfachstrukturen, also Clustern, werden die Varianzquellen *nicht* aufgefunden, sie werden verschleiert.

Eine Deutung von Varimin-transformierten Faktoren ist demnach unproblematisch, sofern minimale Paare vorliegen. Die Deutung von Faktoren der Einfachstruktur-orientierten Faktorenanalysen ist in der Regel ungleich schwieriger, weil bei der Deutung dieser Faktoren *globale Ähnlichkeiten* unter den faktoriell geclusterten Variablen beschrieben werden müssen, wobei immer mehrere zugrunde liegende Merkmale zur Ähnlichkeit bzw. Verschiedenheit beitragen. Bei der Deutung von Komplexstruktur-orientierten Faktoren sind nur *Unterschiede* zwischen jeweils zwei gepaarten Variablen bezüglich *eines einzigen* Merkmals zu ermitteln. ‚Einfachheit‘ im methodischen Vorgehen kommt also – scheinbar paradoxerweise – eher der Complex Structure-Modellierung zu als der Simple Structure-Modellierung.

## Frage II

**Wird die Komplexität der Kovarianzquellen nicht schon durch die Initiallösung hinreichend repräsentiert?**

Man kann bezweifeln, dass es nötig sein soll, die in der Initiallösung schon vorliegende Komplexität noch zu steigern. Man könnte einwenden, das dafür vorgebrachte Argument reiche nicht aus (das Argument war: realitätsverzerrende Bevorteilung der zuerst extrahierten Faktoren gegenüber später extrahierten). Die bislang vorgebrachte Kritik könne zwar das Konzept der Einfachstruktur schwächen, die Forscher würden auf den bislang vernachlässigten Informationswert der Initialstruktur aufmerksam werden, nicht jedoch würde eine Rotation der Initialfaktoren erforderlich sein mit dem Ziel, die Komplexität der initialen Lösung noch zu erhöhen.

Diesen Einwand würde man nur mit empirischen Prüfergebnissen entkräften können. Die Ergebnisse Varimin-rotierter Lösungen, die sich von den Ergebnissen initialer Lösungen unterscheiden, müssten sich als durchweg leichter interpretierbar erweisen, d.h. als befriedigender integrierbar in den jeweils vorliegenden Wissenskonzext, im Vergleich mit den initialen Lösungen, also nicht nur im Vergleich mit Varimax-rotierten Lösungen. Die relative ‚Modelliergüte‘ der Lösungen wäre für die Entscheidung über die anstehende Frage letztendlich maßgebend.

Unter zahlreichen empirischen Belegen, bei denen die Modelliergüte einer Varimin-Rotation die der Initiallösung übertrifft, werden die Ergebnisse einer Analyse der Merkmale beurteilter Münzen ausgewählt.



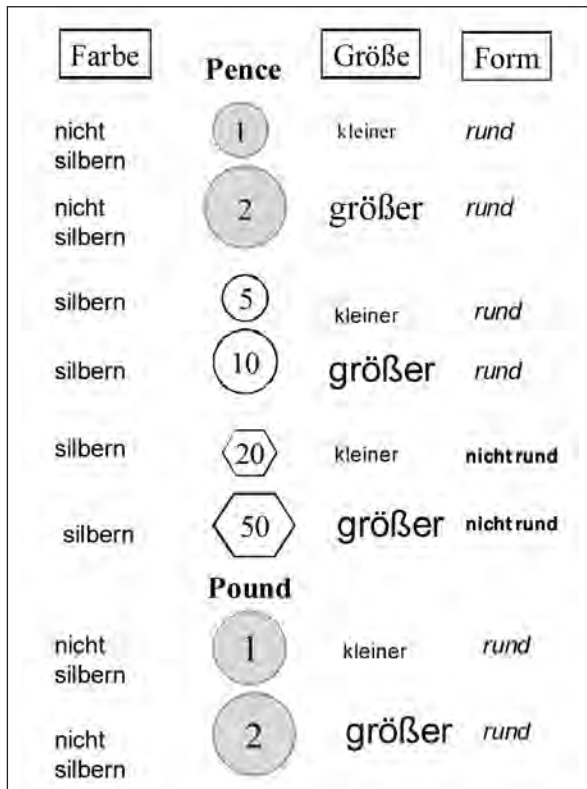


Abbildung 5: Das System der britischen Münzen mit den Eigenschaften Farbe, Größe und Form

**Empirischer Beitrag 2:**

**Ähnlichkeitsbeurteilung britischer Münzen**

Britische Münzen als Urteilsobjekte eignen sich als manifeste Variablen einer faktoriellen Analyse insbesondere deshalb, weil sie eindeutige Merkmale (Varianzquellen ihres phänomenalen Eindrucks) besitzen, nämlich Größe, Form und Farbe. Diese lassen sich mit physikalischen Reizeigenschaften in Verbindung bringen, ihre Feststellung ist nicht auf subjektive Urteile angewiesen (s. Abbildung 5).

Die nach Münzenwert benachbarten Münzen sind größenverschieden (z.B. 1 Pence klein, 2 Pence groß, 5 Pence klein, 10 Pence groß usw.). Dadurch ergeben

sich natürliche Paare, die sich wiederum von anderen Paaren entweder durch die Metallfarbe (5 und 10 Pence z.B. sind silbern, 1 und 2 Pence sind nicht silbern) oder durch ihre Form voneinander unterscheiden (z.B. sind 5 und 10 Pence rund, 20 und 50 Pence sind siebeneckig). Diese Merkmale liegen der Kovarianz der Ähnlichkeitsurteile über die untersuchten Objekte zugrunde. Durchmesser und das Gewicht der Münzen sind aus Tabelle 2 ersichtlich.

Experiment: Eine in Cambridge studierende deutsche Studentin wurde gebeten, die acht Münzen der gegenwärtigen britischen Währung nach Ähnlichkeit einzustufen (1, 2, 5, 10, 20, 50 Pence und 1 und 2 Pound). Die Münzen wurden auf Karton aufgeklebt und in Paaren dargeboten: Jede Münze kam mit jeder anderen Münze gepaart vor, angefangen mit den Paaren 1 Penny<sup>A</sup> – 2 Pence, 1 Pence – 5 Pence usw. bis zum Paar 1 Pound – 2 Pound, insgesamt 28 Paare.

Die Aufgabe der Studentin bestand darin, jede der acht Münzen in die Hand zu nehmen und mit den aufgeklebten Münzen der Karton-Vorlage zu vergleichen. Auf einer bipolaren Siebenpunkt-Skala hatte sie anzugeben, ob die Münze in ihrer Hand eher der links oder der rechts aufgeklebten Münze ähnlich war. Zum Beispiel könnte ihr die 50 Pence-Münze eher der 5-Pence Münze als der 1 Pound-Münze ähnlich erscheinen, worauf sie dann auf der Skala 5 Pence – 1 Pound einen Skalenpunkt in der Nähe des 5 Pence-Pols ankreuzen würde. Ein ganzheitliches Urteil wurde erbeten, alle Ähnlichkeits- und Verschiedenheitsgründe sollten gleichzeitig beachtet werden. Vom finanziellen Münzenwert sollte abgesehen werden.

Auf diese Weise kam für jede Münze ein Urteilsprofil mit 28 Einzelurteilen zustande. Die Urteilsprofile der acht Münzen wurden interkorreliert, die Korrelationsmatrix wurde einer PCA unterzogen, die extrahierten Faktoren anschließend der Varimin- und der Varimax-Rotation zugeführt. Die Erwartung war, dass sich das Merkmalsystem der britischen Münzen nach einer Varimin-Rotation, nicht aber nach einer Varimax-Rotation, faktoriell durchsetzen würde, und dass das Varimin-rotierte Ergebnis Münzen-Merk-

Tabelle 2: Durchmesser und Gewicht der Münzen

	1 Penny	2 Pence	5 Pence	10 Pence	20 Pence	50 Pence	1 Pound	2 Pound
Gewicht (g)	3.28	7.10	3.52	6.50	5.00	8.20	9.48	11.92
Durchmesser (mm)	20	25	18	24	20	28	22	28

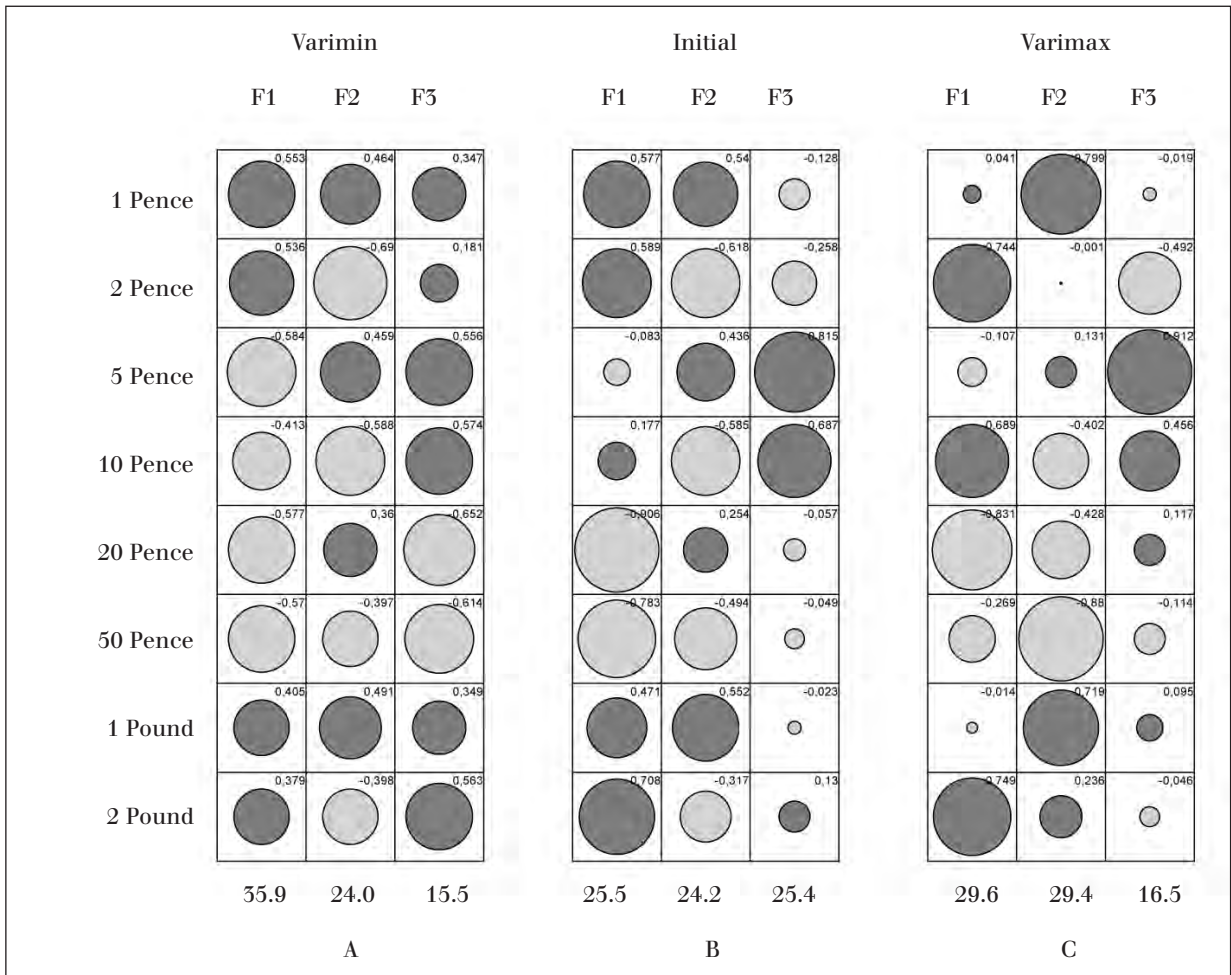


Abbildung 6: Varimin- (A), initiale (B) und Varimax-transformierte Lösung (C) mit Ähnlichkeitsdaten britischer Münzen

male besser hervortreten lassen würde als die initiale Lösung. Das Ergebnis zeigt Abbildung 6.

Das Varimin-Ergebnis fällt erwartungstreu aus: Die Münzen von silberner Farbe heben sich auf dem ersten bipolaren Faktor von den Münzen mit nicht-silberner Farbe eindeutig ab. Dies drücken die Ladungsvorzeichen aus, die Ladungshöhen unterscheiden sich nur wenig. Dasselbe gilt für den zweiten Faktor, der offensichtlich die Größen-Unterschiede der Münzen wiedergibt. Im dritten Faktor schlägt sich offenbar der Varianzanteil der Münzenform nieder.<sup>5</sup> Eine Deutung mithilfe minutiöser Minimalpaar-Vergleiche kann man sich bei der hier vorliegenden Transparenz der ‚latenten‘ Bedingungs-lage ersparen.

Zur Initiallösung: Unter den Faktoren der Initiallösung lässt sich  $F_2$  als Manifestation der Münzengröße deuten. In  $F_1$  könnte man den Farbfaktor vermuten, doch lassen sich die Ladungen von 5 und 10 Pence nicht mit der Münzenfarbe in Zusammenhang

bringen. Eine Deutung von  $F_3$  als Formmerkmal ist ausgeschlossen,  $F_3$  bleibt rätselhaft. Das hier deutlich gewordene defizitäre Ergebnis der Initiallösung führt mit vielen anderen nicht berichteten Ergebnissen aus anderen Untersuchungen zu der praktischen Schlussfolgerung, dass PCA-Faktoren grundsätzlich nach Varimin rotiert werden sollten, auch dann, wenn die Varimin-rotierte Lösung von der initialen nicht sehr abweichen sollte – was bei Faktor-Lösungen mit mehr als zwei Faktoren ohnehin selten vorkommt.

Zur Varimax-Lösung: Das Ideal einer einfaktoriellen Ladung wird nur etwa bei der Hälfte der Variablen erreicht. Die 10 Pence und 20 Pence-Münzen weichen vom solitären Ladungsideal erheblich ab. Multiple Faktorladungen sind in der Simple Structure-Praxis seit jeher bekannt, sie wurden als ein permanent unleidiges Problem hingenommen. Schwerer wiegt die Tatsache, dass Varimax sein Ziel, deskriptive Einfachheit auf der inhaltlichen Ebene zu erreichen, nicht erreicht. Weder Farbe, noch Form, noch Größe

werden durch die Varimax-Faktoren repräsentiert, nicht einmal in Andeutungen.

Um die Generalisierbarkeit des Ergebnisses zu prüfen, wurde das Experiment wiederholt mit acht Studierenden, die mit britischen Münzen keinen Umgang hatten. Die über die Versuchsteilnehmer gemittelten Urteilstwerte wurden genau so analysiert wie die individuellen Werte der Cambridge-Studentin (T. S.). Es zeigte sich, dass bei einem Eigenwertverlauf 2.07, 1.90, 1.30, 1.20 vier Faktoren extrahierbar waren, bei T. S. waren es nur drei Faktoren (Eigenwerte: 2.87, 1.92, 1.24, 0.97). Die faktoriellen Kongruenzen der zu vergleichenden Lösungen, von T. S. und Studenten, waren für  $F_1$  (Farbfaktor) und  $F_2$  (Größenfaktor) beachtlich, sie betragen .981 bzw. .973. Doch fand sich für  $F_3$  von T. S. weder bei  $F_3$  noch bei  $F_4$  der Studenten eine Spur von Kongruenz (.512 bzw. .224). Es stellte sich heraus, dass bei den Studenten die acht  $F_3$ -Gewichte mit den acht Rangplätzen *des Geldwertes* der Münzen hoch korrelierten (.88), während für T. S. die höchste Korrelation mit dem Geldwert nur .14 betrug (bei  $F_1$ ). Offensichtlich haben sich die Studenten bei ihren Ähnlichkeitsurteilen außer von Farbe und Größe noch vom Geldwert der Münzen beeinflussen lassen, während die Münzenform, die sich bei T. S. als  $F_3$  manifestiert hatte, von ihnen unbeachtet blieb.<sup>6</sup>

### Frage III

**Lassen sich faktorielle Simple Structure-Lösungen nicht doch zumeist befriedigend interpretieren, wie könnte man sie sonst ständig verwenden?**

Die Variablen der bisher berichteten Untersuchungen waren Objekte, deren wahrgenommene Merkmale den Ähnlichkeitsurteilen der Probanden zugrunde lagen. Für multivariate Untersuchungen dieser Art werden Faktorenanalysen selten verwendet. Man weiß, dass Faktorenanalysen mit Daten beurteilter Objekte unbefriedigende Ergebnisse liefern. Man weicht auf alternative Verfahren wie MDS (mehrdimensionale Skalierung) aus, mit denen beurteilte Objekte befriedigender dimensional geordnet werden können.

Ein Haupt-Einsatzgebiet für Faktorenanalysen in der Psychologie sind Urteilsdaten über Menschen (Selbst- und Fremdbeurteilungen), Urteile über Persönlichkeitseigenschaften, Verhaltensdispositionen, Einstellungen usw.. Die Varianz der dabei verwendeten Variablen, d. h. der für solche Erhebungen verwendeten Items, ergibt sich durch Messwiederholungen über urteilende Personen. Faktorenanalysen mit solchen Daten führen, wenn sie Simple Structure-orientiert sind, zur faktoriellen Gruppierung verbalsemantischer Urteileinheiten. Bei diesen kann man sich zur Einschätzung der Modelliergüte nicht wie bei Objektbeurteilungen auf anschauliche Kriterien stützen. Semantische Urteileinheiten aus der Erleb-

nis- und Verhaltensdomäne sind weniger streng voneinander geschieden, sie lassen sich leicht kombinieren. Irgendein Sinn wird in den faktoriell gruppierten Variablen fast immer gefunden, das Ungenügen der Simple Structure-Modellierung solcher Variablen wird nicht erkannt. Zwar wird mitunter von signifikanten Korrelationen zwischen Faktorscores aus Selbstbeurteilungen und objektiven Verhaltensdaten berichtet, doch wird dem Faktorenanalytiker die konzeptuelle Arbeit, die er beim Ermitteln latenter Bedingungen der nach Simple Structure gruppierten Urteilsvariablen leisten muss, auch bei anschließender Nutzung von Außenkriterien nicht erspart.

In der folgenden Untersuchung werden Korrelationen von Variablen faktorisiert, bei denen Messwiederholungen über urteilende Personen eine Rolle spielen. Doch handelt es sich hier um inhaltsarme adverbale Variablen (*Formwörter*). Sie stammen aus einer Untersuchung von Carl (1968) über das Antwortstilverhalten. Mit den an diesen Daten gewonnenen Ergebnissen lässt sich aufzeigen, wie Simple Structure-Faktoren und ihre Deutungen entstehen. Zu diesem Zweck werden die mit denselben Daten ermittelten Variminfaktoren herangezogen.

### Empirischer Beitrag 5:

**Differenzierung von Antwortstilen bei der Beantwortung von Fragebögen** (Daten von Carl)

Das Ziel Carls (1968) war die Ermittlung von Antwort-Tendenzen (response sets) bei der Verwendung von Stufenskalen. Er hatte von 100 Personen Fünf-Punkt-Likert-Einstufungen für 580 Items aus dem *Minnesota Multiphasic Personality Inventory* (MMPI) erhoben und für jede Person über alle Items, ohne Rücksicht auf deren Inhalt zu nehmen, die Anzahl der fünf Einstufungen summiert, die von ‚stark zustimmend‘ bis ‚stark ablehnend‘ reichten (die Zwischenstufen der Skalen waren nicht verbalisiert worden). Carl hatte ausgeschlossen, dass aufgrund der gegenseitigen Abhängigkeit der Stufen-Urteile Artefakt-Korrelationen vorkamen: Er ermittelte die Verwendungshäufigkeiten für die fünf Antwortstufen mit jeweils aus dem Gesamtpool gewonnenen Teilstichproben von Items, die inhaltlich parallelisiert worden waren. Aus der Interkorrelationsmatrix Carls, die auch Variablen aus Skalen mit mehr Stufen enthielt (er hatte z.B. auch Sieben-Punkt-Skalen verwendet), wurde für den vorliegenden Zweck die Matrix für das Format mit fünf Urteilsstufen separiert. Für diese wurde mithilfe einer PCA die initiale Faktorenstruktur ermittelt, die dann mit Varimin und zum Vergleich auch mit Varimax rotiert wurde.

Bei dieser Untersuchung sind mit fachlichem Vorwissen bestimmte Ergebnisse erwartbar. Denn bekannt ist die *Ja-Sage-Tendenz* (Akquieszenz), die ei-

nen Faktor hervorbringen müsste, bei dem die beiden zustimmenden Antwortstufen positiv und die beiden ablehnenden Antwortstufen negativ geladen sind. Bekannt ist auch die *Extremantwort-Tendenz*, die einen Faktor hervorbringen müsste, bei dem die beiden Extrem-Stufen der Zustimmung und Ablehnung positiv und die dazwischen liegenden Stufen negativ geladen sein müssten (evtl. nach Vorzeichenumkehr, die Polungsrichtung ist irrelevant). Mindestens diese beiden Faktoren also müsste eine Nachbehandlung der initialen Faktoren durch Varimin zutage fördern.

Das *Varimin-Ergebnis* zeigt Abbildung 7B. Man erkennt in  $F_3$  den erwarteten Akquieszenz-Faktor und in  $F_1$  den Extremantwort-Faktor. Allerdings ist ein weiterer substanzialer Faktor  $F_2$  vorhanden, der nicht erwartet wurde und der zur Deutung herausfordert. Man kann hier bei Herrmann (1965) anknüpfen, der die von ihm so genannte *Urteilsnuanciertheit* als Variante eines Antwortstils bei Fragebogen-Untersuchungen identifizierte. Demnach unterscheiden sich Personen beim Ausfüllen von Fragebögen auch darin, inwieweit sie über eine zustimmende vs. ablehnende Entscheidung hinausgehend zwischen ‚Zustimmung‘ und ‚starker Zustimmung‘ sowie zwischen ‚Ablehnung‘ und ‚starker Ablehnung‘ und auch zwischen ‚Zustimmung/Ablehnung‘ einerseits und ‚Unentschieden‘ bzw. ‚Weiß nicht‘ andererseits differenzieren. Diese Antwortstil-Variante wird in der Literatur seltener diskutiert (gelegentlich unter der Bezeichnung ‚Urteilsvariabilität‘, ‚response variance‘), vermutlich weil sie die Ergebnis-

se weniger verzerrt als die anderen Antwort-Tendenzen (Hinz et al., 2003). Der zur Klärung anstehende Faktor  $F_2$  der Varimin-Analyse lässt sich m. E. risikolos als Ausdruck der ‚Urteilsnuanciertheit‘ deuten.

Sind die drei Response Sets schon in der *initialen* Faktorenstruktur zu finden (s. Grafik 7A)? Die initiale Faktorstruktur ist der Varimin-Faktorstruktur sehr ähnlich. Faktor  $F_1$  ist in den beiden Lösungen nahezu identisch, also ist auch der *initiale*  $F_1$ -Faktor als Ausdruck der Extremantwort-Tendenz zu deuten. Doch bei  $F_3$ , Akquieszenz, ist die initiale Lösung weniger gut, da hier die mittlere Kategorie der Urteils-Enthaltung eine beachtliche negative Ladung aufweist, während doch einer Urteils-Enthaltung andere Bedingungen zugrunde liegen sollten als einem Urteil, das eine Ablehnung beinhaltet. Auch für  $F_2$  ist die Varimin-Lösung prägnanter. Bei der initialen Lösung wird bei  $F_2$  zwischen ‚Ablehnung‘ und ‚starke Ablehnung‘ nicht differenziert. Auch ist die numerische Differenz zwischen ‚Zustimmung‘ und ‚starke Zustimmung‘ in der initialen Lösung deutlich schwächer als bei der Varimin-Lösung. Von Urteilsnuanciertheit kann man beim initialen  $F_2$ -Faktor kaum sprechen.

Unser Hauptinteresse aber gilt hier dem *Varimax-Ergebnis* (Grafik 7C). Wie lässt es sich interpretieren? Darüber gibt ein Vergleich der Varimax- mit der Varimin-Lösung (vgl. Abbildung 7C mit 7B) Aufschluss. Die Varimax-Rotation hat zur bipolaren Clusterung der Variablen ‚starke Zustimmung‘ (positive Ladung

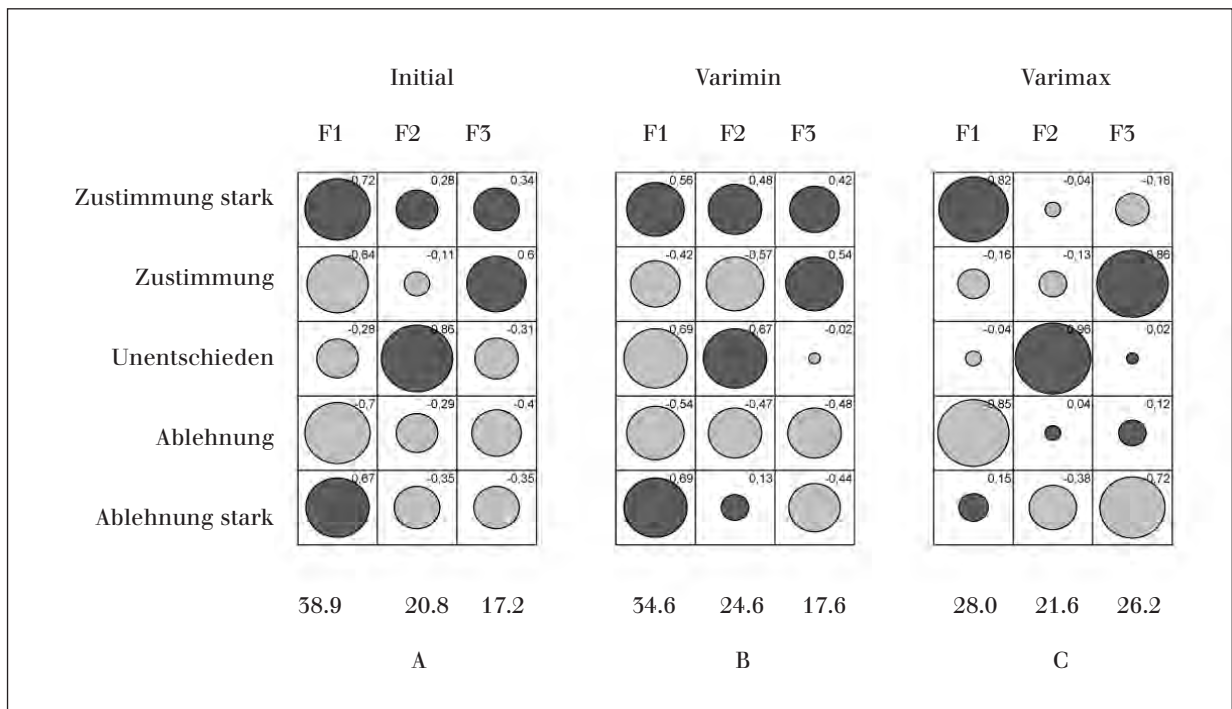


Abbildung 7: Initiale (A), Varimin- (B) und Varimaxstruktur (C) der Response Set-Daten von Carl (1968)

bei  $F_1$ ) und ‚Ablehnung‘ geführt (negative Ladung bei  $F_1$ ). Warum? Es fällt auf, dass die Vorzeichen der Ladungen von ‚starke Zustimmung‘ und ‚Ablehnung‘ bei den drei Merkmalen, die durch Varimin differenziert wurden, kontrastieren. Auch beim Varimax-Faktor  $F_2$ , der ebenfalls bipolar ist, finden sich bei den Varimin-Merkmalprofilen Kontraste, diesmal bei den Variablen ‚Zustimmung‘ und ‚starke Ablehnung‘. Für die Urteilstkategorie ‚unentschieden‘ gibt es bei den anderen Urteilstufen im Varimin-Ergebnis keinen polaren Gegensatz, so bleibt ‚unentschieden‘ auf Varimax- $F_2$  ziemlich isoliert.

Diese Ergebnisse entsprechen dem, was der Varimax-Varimin Faktorenvergleich bei den Münzen erbracht hat. Kurz: Varimax-Transformationen clustern Variablen mit gleichartigen Merkmalprofilen.<sup>7</sup> Im Falle bipolarer Faktoren clustern sie zusätzlich, mit dem jeweils anderen Vorzeichen, die Variablen mit gegensätzlich ausgeprägten Profilen. Die Merkmalsgrundlage selbst wird nicht analysiert, eine Differenzierung auf der Merkmalebene wird durch eine Varimax-Clustering systematisch verunmöglicht.<sup>8</sup>

Was würde ein konventioneller Faktorenanalytiker nach Durchführung einer Varimax-Analyse der Response-Set-Daten möglicherweise publizieren? Er könnte behaupten, Akquieszenz sei nicht, wie man bislang dachte, ein einfaktorielles Konstrukt, man habe vielmehr zwischen einer Akquieszenz I mit Extremheitstendenz ( $F_1$ ) und einer Akquieszenz II ohne Extremheitstendenz ( $F_3$ ) zu unterscheiden. Dann

könnte er noch  $F_2$  als ‚Unentschiedenheit‘ interpretieren und sich darüber hinweg setzen, dass  $F_2$  auch bei ‚starke Ablehnung‘ eine beachtliche negative Ladung aufweist. Er könnte der statistischen Datenverarbeitung voll vertrauen und glauben, drei neue psychologische Konstrukte entdeckt zu haben (Akquieszenz I und II und Unentschiedenheit). Da diese nicht sinnlos sind, würde man nicht bemerken, dass sie aber ziemlich nutzlose Variablengruppierungen darstellen, wie sie durch Anwendung des Modells der Einfachstruktur fortwährend hervorgebracht werden.

Mit der folgenden Reanalyse *verbaler* Daten aus einer MDS-Untersuchung wird die unverbindliche Deutungspraxis, die bei Verwendung semantischer Einheiten nach ihrer Faktorisierung durch Varimax möglich ist, noch deutlicher.

**Empirischer Beitrag 4:  
Semantische Merkmale bei Verwandtschaftsbezeichnungen** (Daten von Marx & Hejj)

Für 16 Verwandtschaftsbezeichnungen, die den Probanden auf Wortkarten dargeboten und von ihnen hierarchisch sortiert worden waren, ermittelten Marx & Hejj (1989) eine Matrix der Ähnlichkeiten (s. Tabelle 3, bei den Autoren Tabelle 2.5, S. 112). Die Autoren selbst verwendeten ein NMDS-Verfahren, um aus den Häufigkeiten der zusammen gelegten Wortkarten die semantischen Merkmale der Verwandtschaft zu gewinnen, was ihnen auch einigermaßen gelang.<sup>9</sup> Für den vorliegenden Zweck wurde die Originalmatrix der

Tabelle 3: Daten des hierarchischen Sortierens von Marx & Hejj (1989)

	1	2	3	4	5	6	7	8	9	10	11	12	15	14	15	16
1 Bruder																
2 Cousin	759															
3 Cousine	858	451														
4 Enkel	777	770	866													
5 Enkelin	842	865	792	454												
6 Großmutter	860	882	850	702	570											
7 Großvater	785	822	877	564	695	444										
8 Mutter	747	942	895	856	805	679	854									
9 Nefte	787	558	674	802	882	925	865	970								
10 Nichte	900	685	542	875	785	869	926	927	465							
11 Onkel	819	594	711	804	875	856	772	915	592	719						
12 Schwester	432	857	764	841	781	774	855	675	884	797	892					
15 Sohn	632	892	955	721	869	864	764	710	900	969	912	745				
14 Tante	905	717	602	896	821	766	845	840	725	576	446	810	960			
15 Tochter	728	872	890	865	735	775	869	561	971	902	972	625	476	906		
16 Vater	614	896	947	787	894	825	695	477	915	981	851	744	571	924	718	

ANH 2.5  
Gruppenmatrix der Distanz-Koeffizienten

Sortierhäufigkeiten, welche als Indikatoren der Distanz oder Unähnlichkeit (UÄ) unter den 16 Begriffen zu betrachten sind, diagonal gespiegelt und zu einer Vierecksmatrix aufgefüllt. Sodann wurden die UÄ-Maße durch  $\ddot{A} = 1 - U\ddot{A}/1000$  in Ähnlichkeitsmaße transformiert.<sup>10</sup> In die Diagonale wurde der aus der jeweils zugehörigen Spalte/Zeile zu entnehmende höchste Ähnlichkeitswert eingesetzt. Die Spalten der Ähnlichkeitsmatrix wurden sodann untereinander korreliert und die Interkorrelationsmatrix einer PCA unterworfen. Fünf substantielle Faktoren wurden extrahiert<sup>11</sup> und nach Varimin und Varimax transformiert.

Abbildung 8A zeigt das Ergebnis der Varimin-Lösung. Der Generalfaktor  $F_1$  hat keine inhaltlich-differenzierende Funktion (minimale Ladungsvarianz), er ist im wesentlichen verfahrensbedingt und kann ignoriert werden.<sup>12</sup> Die Faktoren  $F_2$  bis  $F_5$  sind bipolar, sie lassen mit kontrastierenden Ladungsvorzeichen die erwar-

teten semantischen Merkmale erkennen: Linearität ( $F_2$ ), Kernfamilie ( $F_3$ ), Geschlecht ( $F_4$ ) und Alter oder Generation ( $F_5$ ).<sup>15</sup> Das minimale Wortpaar *Bruder* und *Schwester* z. B. zeigt einen Unterschied der Ladungsvorzeichen lediglich bei Faktor  $F_4$ , der als Geschlechtfaktor (männlich vs. weiblich) zu deuten ist. Andere minimale Wortpaare lassen sich leicht finden, so z. B. *Vater* und *Sohn* oder *Mutter* und *Tochter*, die beide auf  $F_5$ , dem Faktor ‚Generation‘, mit ihrer Ladungsrichtung kontrastieren. In der Faktorenstruktur kommen zwar subtilere Unterschiede nicht zum Ausdruck, so etwa nicht die zwischen der jüngsten, mittleren und der ältesten Generation (etwa Sohn, Vater, Großvater). Auch werden Unterschiede, die durch die Ego-Perspektive bedingt sind, faktoriell ignoriert, z. B. die zwischen (*mein*) *Bruder* und (*meines Vaters*) *Sohn* – die gleiche Person ist gemeint. Doch sind die Hauptmerkmale der Verwandtschaft im Resultat aufzufinden.

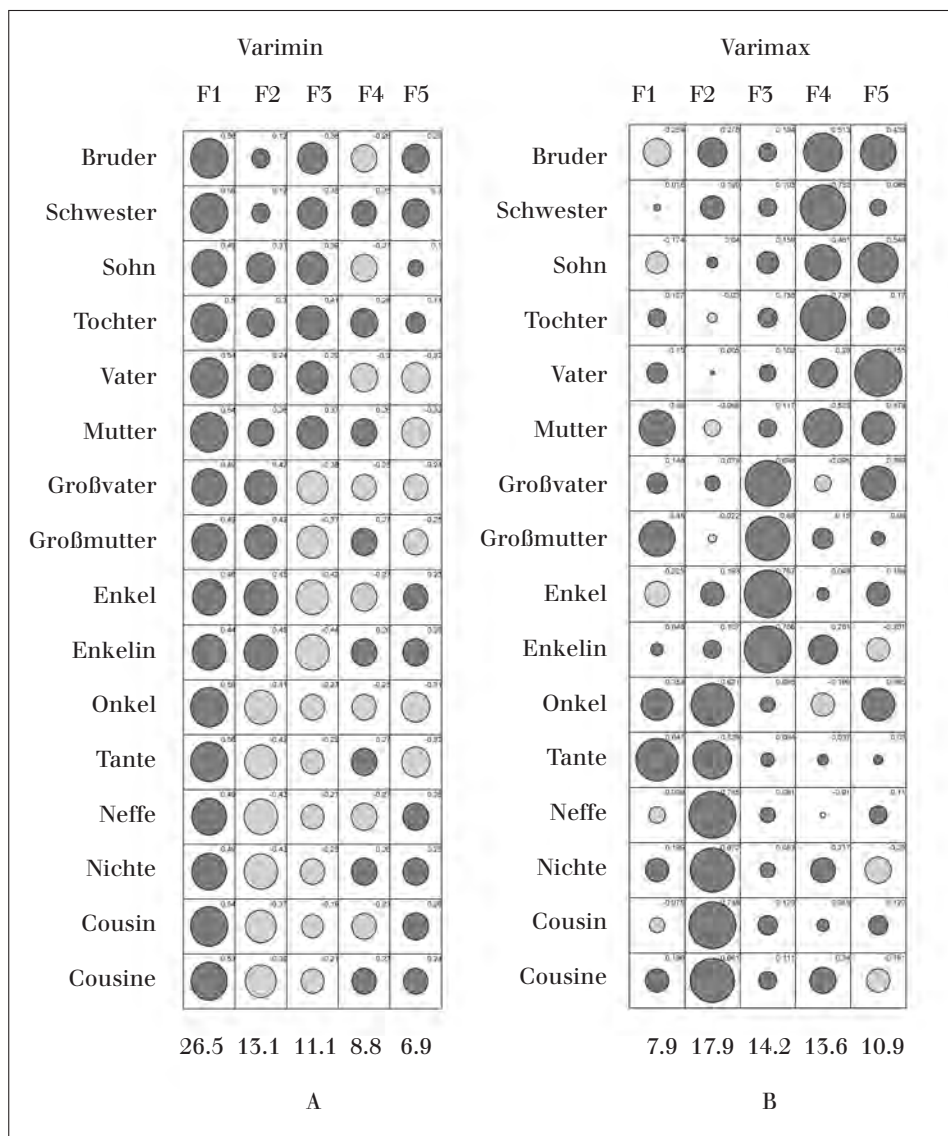


Abbildung 8: Varimin- (A) und Varimax-rotierte Faktoren (B) der Analyse von Verwandtschaftsähnlichkeiten (Daten von Marx & Heij, 1989)

Abbildung 8B zeigt das Ergebnis einer *Varimax*-Rotation der Faktoren, die für Abbildung 8A Variminrotiert wurden. Von den fünf rotierten Faktoren lassen sich  $F_1$  und  $F_5$  gar nicht und  $F_3$  nur mit einiger Willkür interpretieren.  $F_3$  vereint Großeltern und Enkel, worin man einen Sinn hineinlegen kann, denn damit werden die generationsmäßig extremen Vertreter innerhalb derselben Linie gruppiert. Mit ‚sehr jung oder sehr alt in der dominanten Abstammungslinie‘ würde man diese Einheit bezeichnen können.

Bei den Faktoren  $F_2$  und  $F_4$  hat das Varimax-Ergebnis eine gewisse Ähnlichkeit mit dem Varimin-Ergebnis. Mit Varimax- $F_2$  sind die Verwandtschaftswörter der *nicht-linealen* Abstammung geladen (‚entferntere Verwandte‘). Im Unterschied zur Varimin-Lösung fehlt aber den Wörtern der linealen Abstammung (‚nähere Verwandte‘) als faktorielle Kennzeichnung der Vorzeichen-Gegensatz ([+] ‚Merkmal vorhanden‘ vs. [-] ‚Merkmal nicht vorhanden‘), alle Varimax-Faktoren sind unipolar.  $F_4$  gruppiert die Vertreter der Kernfamilie, doch die Nicht-Mitglieder bleiben bei  $F_4$  wieder ohne Kennzeichen (kein negatives Vorzeichen). Außerdem haben *Enkelin* und *Nichte* und *Cousine* noch beträchtliche  $F_4$ -Ladung, womit man wenig anfangen kann.

Es ist also nicht so, dass sich Varimax-Faktoren einer Interpretation völlig entziehen. Doch wegen semantischer Überlappungen lässt ihre Interpretation meist zu wünschen übrig, was mithilfe der transparenten Semantik der Verwandtschaftswörter leicht zu erkennen ist. Vor allem bleiben im Varimax-Ergebnis auch bei den deutbaren Faktoren die *Quellen* der semantischen Ähnlichkeit und Verschiedenheit (Generation, Geschlecht usw.), die eine Faktorenanalyse aufdecken sollte, verborgen. Großeltern und Enkel werden aufgrund der Merkmale Linearität und Generation zusammen geclustert, doch werden diese Merkmale faktoriell nicht manifest. Eklatante Beispiele für die mit Varimax oft einhergehende schwere Interpretierbarkeit von Item- und Skalen-Clustern, die ein weitgehend willkürliches Sinngeben und Benennen ermöglichen, bietet die Big-Five-Persönlichkeitsforschung: Bei einem der fünf Faktoren schwankte man anfangs sogar zwischen drei semantisch höchst heterogenen Bezeichnungen: ‚intellect‘, ‚culture‘ und ‚openness‘ (mehr dazu in Ertel a). Auch ist ‚agreeableness‘ eine ziemlich ungeeignete Bezeichnung, da sie doch eher die Wirkung einer Persönlichkeit A auf eine Persönlichkeit B beschreibt als eine Erlebens- oder Verhaltenstendenz der Persönlichkeit A selbst.

#### Frage IV

##### Enthält die regelmäßig anzutreffende Bipolarität von Varimin-Faktoren substantielle Information?

Die Bipolarität in Varimin-Faktorlösungen, die schon im Vorläufer-Artikel diskutiert wurde (Ertel, 2009, p. 42f) und in den beiden letzten empirischen Beiträgen festzustellen war, verdient besondere Beachtung. Bipolarität tritt in Varimin-Lösungen regelmäßig auf, weit häufiger als in Varimax- und anderen Simple Structure-Lösungen. Das Vorkommen von negativen Ladungen – hier sind nur die mit der Komplexstruktur gewonnenen gemeint und die initialen – kann bedeuten, dass die negativ geladenen Variablen gegenüber den positiv geladenen für die betreffende Quelle der Kovarianz (für den betreffenden Faktor) eine funktional abträgliche Wirkung haben. Bei den Response Set-Daten von Claus z. B. musste man bei den Variablen mit Minus-Faktorgewichten eine hemmende Wirkung annehmen: Wenn jemand viele extreme Ja-Nein-Antworten abgibt, dann gibt er plausiblerweise weniger gemäßigte Antworten ab, und umgekehrt. Das lässt sich seiner Motivation zuschreiben, die gemäßigten Urteile entweder zu meiden oder zu bevorzugen.<sup>14</sup> Eine motivational-funktionale Deutung der Bipolarität bei diesen Response-Set-Faktoren ist erlaubt, weil Carl Vorkehrungen zur Ausschaltung von Artefakt-Korrelationen getroffen hatte. Für die bipolaren Akquieszenz- und Nuanciertheitsfaktoren gilt die motivational-funktionale Deutung des Vorzeichen-Kontrasts bei Faktoraladungen entsprechend.

Anders ist die Bipolarität bei den Verwandtschaftsdaten zu interpretieren. Beim Faktor Geschlecht z. B. kann Bipolarität nicht auch funktional gedeutet werden. Das Vorliegen des Merkmals *männlich* setzt keine Hemmung durch *weiblich* voraus, sondern eine organismische Gegebenheit, die das Vorkommen des Merkmals *weiblich* ausschließt (in der Regel), wenn *männlich* vorliegt. Das Merkmal Linearität setzt für sein Vorkommen oder Nichtvorkommen nur die richtige Antwort auf die Frage nach der entsprechenden Platzierung des Verwandten im Abstammungsbaum voraus, die ein Ja oder Nein ergibt. In solchen Fällen muss man sich darauf beschränken, mit den Vorzeichen das Vorkommen bzw. Nichtvorkommen eines Merkmals ausgedrückt zu finden. Dies wird übrigens auf anderen Gebieten, z. B. in linguistischen Diskursen, viel häufiger so gehandhabt als in der Psychologie.

Der Fall, dass eine negative Ladung eines Faktors das Nichtvorkommen eines Einflusses oder einer Funktion ausdrückt, kann darauf schließen lassen, dass der Faktor für die negativ geladenen Variablen irrelevant ist. Die Entstehung des Faktors ist in solchen Fällen denjenigen Variablen des untersuchten Samples zu verdanken, bei denen er positive Ladungen aufweist.<sup>15</sup>

Man sollte sich bei Faktorenanalysen mit Komplexorientierung darauf einstellen, negative Vorzeichen – wenn sie bei einem Faktor mit positiven zusammen vorkommen – zunächst immer als *Merkmal nicht vorhanden* zu deuten und erst in einem zweiten Schritt, aufgrund eines von anderswoher bezogenen Kontextwissens, eine funktionale Deutung (Hemmung, polar gerichtete Ausprägung usw.) zusätzlich einbringen, wenn man eine solche mithilfe seines Kontextwissen begründen kann.<sup>16, 17</sup>

Zum besseren Verständnis der Bipolarität von Faktoren ist es nützlich, ihr Entstehen anhand eines besonders transparenten Beispielfalls zu verfolgen:

**Empirischer Beitrag 5:  
Intelligenzentwicklung in der Kindheit**

(Daten von Humphreys & Davey)  
In einer Längsschnittstudie testeten Humphreys & Davey (1988) Kinder vom dritten Lebensmonat bis zum 9. Lebensjahr mit insgesamt vier jeweils altersangepassten vergleichbaren Intelligenztests. Sie wollten die Beständigkeit der intellektuellen Entwicklung in diesem Zeitraum ermitteln. Die von ihnen mitgeteilte Interkorrelationsmatrix für 14 Testwiederholungen im Längsschnitt wurde von uns PCA-faktoriert und Varimin- und Varimax-rotiert. Abbildung 9 gibt die Ergebnisse wieder (die nicht-wiedergegebene Initiallösung ist mit der Variminklösung nahezu identisch). Nur zwei Faktoren sind substantiell (Eigenwerte: 5.705, 1.475, 0.827, 0.788 ...).

Der erste Variminfaktor repräsentiert mit einer Kommunalität von 40.7% den stabil bleibenden Anteil der allgemeinen Intelligenz, die Faktorladungen bleiben innerhalb der Untersuchungszeit auf ungefähr gleichem Level. Der zweite Faktor ist bipolar, er repräsentiert mit 10.6 % Kommunalität die zeitabhängige Varianz im Laufe der Intelligenzentwicklung. Diese ist plausiblerweise zurückzuführen auf interindividuell variierende, günstige oder ungünstige Lebensumstände: auf Erziehungs- und Ausbildungseinflüsse, psychische oder körperliche Erkrankungen usw..

Die Bipolarität von  $F_2$  und die monotone Abfolge der Faktorgewichte werden wie folgt interpretiert: Der mit  $F_1$  nicht ausgeschöpfte Kovarianzanteil verteilt sich ziemlich gleichmäßig über die Altersspanne, d. h. die Veränderungen von einem Messzeitpunkt zum nächsten sind geringfügiger als von einem Messzeitpunkt zum übernächsten. Bis zum *über-übernächsten* sind die Veränderungen noch größer, und vom ersten bis zum letzten Messzeitpunkt sind sie am größten. Dementsprechend korrelieren die Testergebnisse der einander benachbarten Messzeitpunkte höher als die der einander entfernteren Messzeitpunkte. Die im mitt-

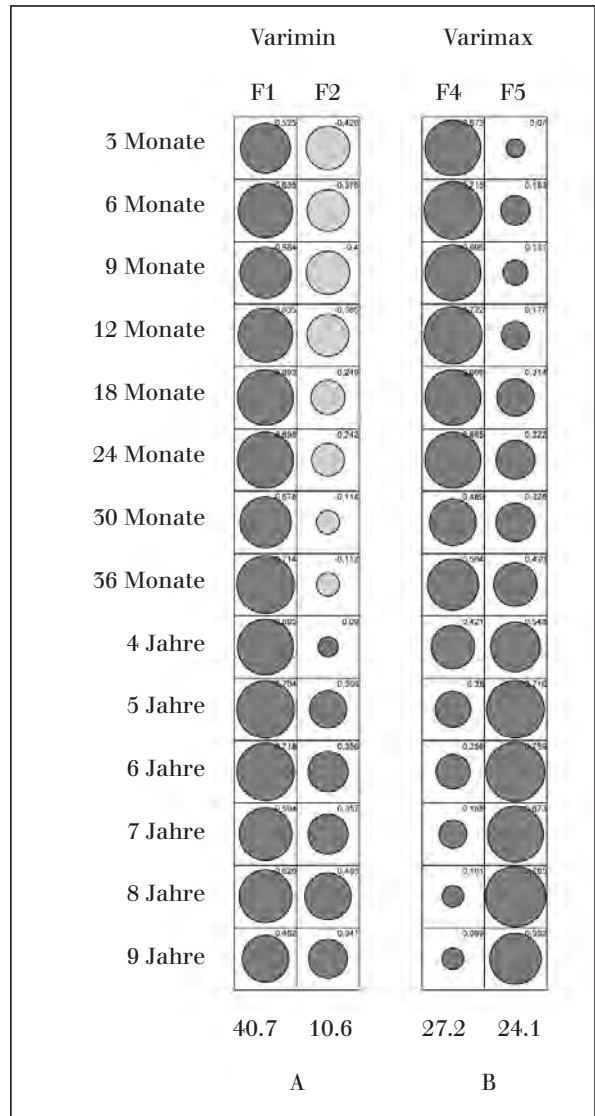


Abbildung 9: Varimin- (A) und Varimaxlösung (B) der Längsschnittdaten zur Intelligenzentwicklung, gewonnen von Humphreys & Davy (1988)

leren Bereich des Längsschnitts liegenden Testwerte haben zu den Testanfangs- und -endwerten ungefähr gleiche Differenzen, die entsprechenden Korrelationen des Testergebnisses von der Mitte des Längsschnitts mit dem ersten und letzten Testergebnis sind ungefähr gleich hoch. Den Testzeitpunkt mit mittlerer Differenz gegenüber dem Anfangs- und Endwert macht das Varimin-Modell zum Nullpunkt des  $F_2$ -Ladungsvektors. Faktor  $F_2$  repräsentiert somit im vorliegenden Fall das Ausmaß an Veränderung der Intelligenz gegenüber dem Mittelwert der Veränderungen.

Das Beispiel ist insofern lehrreich, als es zeigt, dass die Faktorgewichte von  $F_2$  auf die von  $F_1$  bezogen sind, wobei die Art des Zusammenhangs durch einschlägiges Kontextwissen spezifiziert werden muss. Negative



$F_2$ -Ladungen sind in diesem Fall weder Anzeichen von fehlendem oder gegensätzlichem Einfluss noch von logischem Ausschluss, sondern Anzeichen von Differenzen der Intelligenz-Veränderungswerte zum mittleren Veränderungswert. Da die  $F_2$ -Faktorgewichte positiv und negativ gepolt sind, entsteht eine Skala, die anzeigt, dass die Veränderung der Testergebnisse im Laufe der Längsschnitt-Untersuchung monoton zunimmt. Denn wenn z. B. ein genügend hoher konstanter Wert zu den Faktorladungen addiert wird, können die Bipolarität und der Nullpunkt verloren gehen, der Anstieg der Faktorladungen mit zunehmendem Alter bliebe erhalten, dieser würde die seit Beginn der Testperiode allmählich größer werdende Veränderung der Intelligenzleistungen repräsentieren.

Die Varimaxlösung beseitigt den Generalfaktor und ignoriert somit die Tatsache, dass die Intelligenz im getesteten Längsschnitt relativ konstant bleibt. Stattdessen wartet sie mit zwei Faktoren auf, die auf den ersten Blick wie zwei unabhängige Intelligenzen erscheinen, die eine, die sich in der frühesten Kindheit auswirkt, während eine andere Intelligenz die primäre allmählich ablöst – eine absurde Vorstellung, zu der auch Einfachstruktur-Faktoristen kaum neigen werden. Sie würden ihr Verfahren im Falle von Längsschnitt-Wiederholungsdaten vermutlich nur für ungeeignet ansehen.

Ein andersartiges Beispiel, bei dem die negative Ausprägung eines Faktors nicht lediglich Abwesenheit eines Merkmals, sondern ein skalometrisch eigenständiges Merkmal erkennen lässt, bietet die folgende Untersuchung.

#### Empirischer Beitrag 6:

##### Körpervolumen und Körperform von Rindern

(Daten von Rasch)

Von 107 weiblichen Rindern (Färsen) erhob Rasch (1962) 12 Maße ihrer Körperausdehnung. Es handelte sich um Höhen-, Breiten- und Längenmaße. Unsere übliche faktorielle Verarbeitung der von E. Weber mitgeteilten Interkorrelationsmatrix führte zu der in Abbildung 10 wiedergegebenen zweifaktoriellen Lösung (Eigenwerte: 7.69, 1.20, 0.74...). Die nicht wiedergegebene Initialstruktur ist mit der Variminstruktur nahezu identisch.

Der erste Faktor (Kommunalität 64.0%) repräsentiert das Körpervolumen (die Körperausdehnung) ohne Spezifizierung einer Raumdimension. Er besagt, dass ein Rind, das über- oder unterdurchschnittlich hoch gewachsen ist, meist auch über- bzw. unterdurchschnittlich in die Länge und Breite gewachsen ist.  $F_1$  ist dem Generalfaktor der Intelligenz analog zu deuten.

Der zweite Faktor (Kommunalität 10.1%) repräsentiert mit seiner Bipolarität die Varianz, die die Körperform mit sich bringt. Es gibt unter sehr voluminösen wie unter wenig voluminösen Rindern *relativ* schmal gewachsene (schlanke) und *relativ* breit gewachsene (dicke). Das Auftreten von Polarität zeigt an, dass mit mäßigem Spielraum das Wachstum in die Länge „auf Kosten der Breite“ und das in die Breite „auf Kosten der Länge“ vonstatten geht.

In der Varimaxlösung geht der Generalfaktor des Körpervolumens verloren. Die Variablen der Schlankheit und Dicke sind zwar auch erkennbar, allerdings als nicht-polare orthogonale Dimensionen. Die empirische Gegenläufigkeit der Längen- und Breitenausdehnung, die sich durch Varimin aufgrund der Dominanz des Volumenfaktors zu einer bipolaren Einflussgröße der ‚Körperform‘ herausbildet, kommt im Modell der Einfachstruktur nicht zum Tragen.<sup>18</sup>

Die Deutung negativer Ladungen eines Faktors als Ausdruck eines bipolar aufzufassenden Charakteristikums kann auch bei psychologischen Daten nahe liegen, demonstrierbar mit folgendem Beitrag.

#### Empirischer Beitrag 7:

##### Intelligenz- und Leistungstests

(Daten von Holzinger & Swineford)

Die Holzinger & Swineford-Daten werden gelegentlich in Lehrtexten zur Demonstration von Modellrechnungen verwendet, so auch von Jöreskog & Sörbom (2005), von deren Website die Interkorrelationstabelle entnommen wurde (zu ergoogeln mit ‚LISREL 8.52 Jöreskog‘). Es handelt sich um je drei Tests für visuelle, für verbale und für Tempo-Leistungen (speed tests). Die Ergebnisse der vorliegenden Standard-Auswertung für Varimin und Varimax zeigt Abbildung 11.

Varimin- $F_1$  (Abbildung 11 A) lässt sich als Generalfaktor ‚g‘ der allgemeinen Intelligenz ansehen, der für Intelligenztestbatterien generell zu erwarten ist. Der Faktor  $F_2$  kontrastiert durch gegensätzliche Vorzeichen die drei Speed-Test-Leistungen mit den übrigen Testleistungen, bei denen es offensichtlich weniger auf Tempo als auf Konzentration oder ‚power‘ ankommt. Hier liegt ein Fall vor für Bipolarität im Intelligenzbereich, bei dem – mit psychologischen Kontextkenntnissen – negative Ladungen mit funktionaler Bedeutung in Zusammenhang gebracht werden dürfen. Eine erhöhte Fähigkeit und Neigung zu Tempoleistungen wirkt sich vor allem bei geeigneten Speed-Tests förderlich aus. Bei den Tests, die Konzentration und Vertiefung in Probleme erfordern, werden sich Speed-Neigungen und -Fähigkeiten vermutlich eher nachteilig auswirken. Eine entsprechende Gegenwirkung lässt sich auch für die Fähigkeit und Neigung zur Konzentration denken, die sich bei Tests, die Tempofähigkeit und -neigung

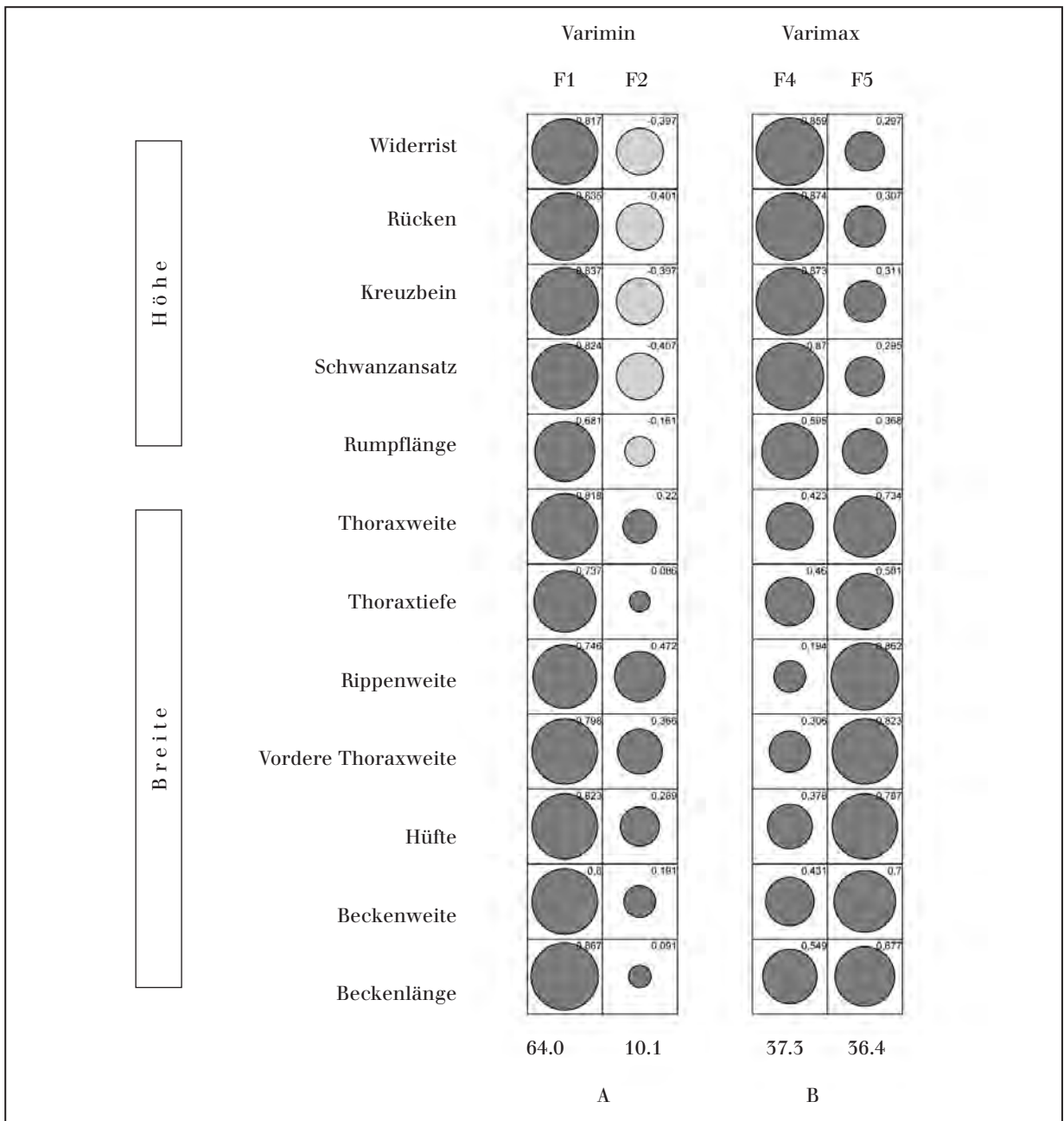


Abbildung 10: Varimin- (A) und Varimaxfaktor (B) von 12 Körpermaßen weiblicher Rinder (Färsen)

erfordern, vermutlich nachteilig auswirken. Diese Gegenläufigkeit kann sich mit der Bipolarität von  $F_2$  manifestieren.

Die Bipolarität von Varimin- $F_3$  hebt eine andere Kovarianzquelle ans Licht: Die Besonderheit verbaler Testleistungen. Bemerkenswert ist, dass sich die Bipolarität von  $F_3$  nicht auf verbale vs. visuelle Tests beschränkt, die beide zur Power-Kategorie gehören, wegen der unterschiedlichen Aufgabeninhalte (Wörter vs. Bildmaterial) bilden verbale und visuelle Tests gute minimale Kontraste. Doch auch die rechnerischen Leistungen aus der Speed-Testserie kontrastieren,

wenn auch etwas schwächer, mit den verbalen Testleistungen. Offenbar liegen den verbalen Leistungen übergreifende Besonderheiten zugrunde. Auch hier lässt sich eine antagonistische Beziehung denken: Die verbalen Fähigkeiten und Neigungen, wenn sie im individuellen Intelligenzprofil dominieren, könnten die rechnerischen als auch die visuellen Aktivitäten etwas benachteiligen und/oder umgekehrt. Doch kann die funktionale Relevanz bei negativen Varianzquell-Ladungen hier wie oft in anderen ähnlichen Fällen nur Hypothese sein.<sup>19</sup>

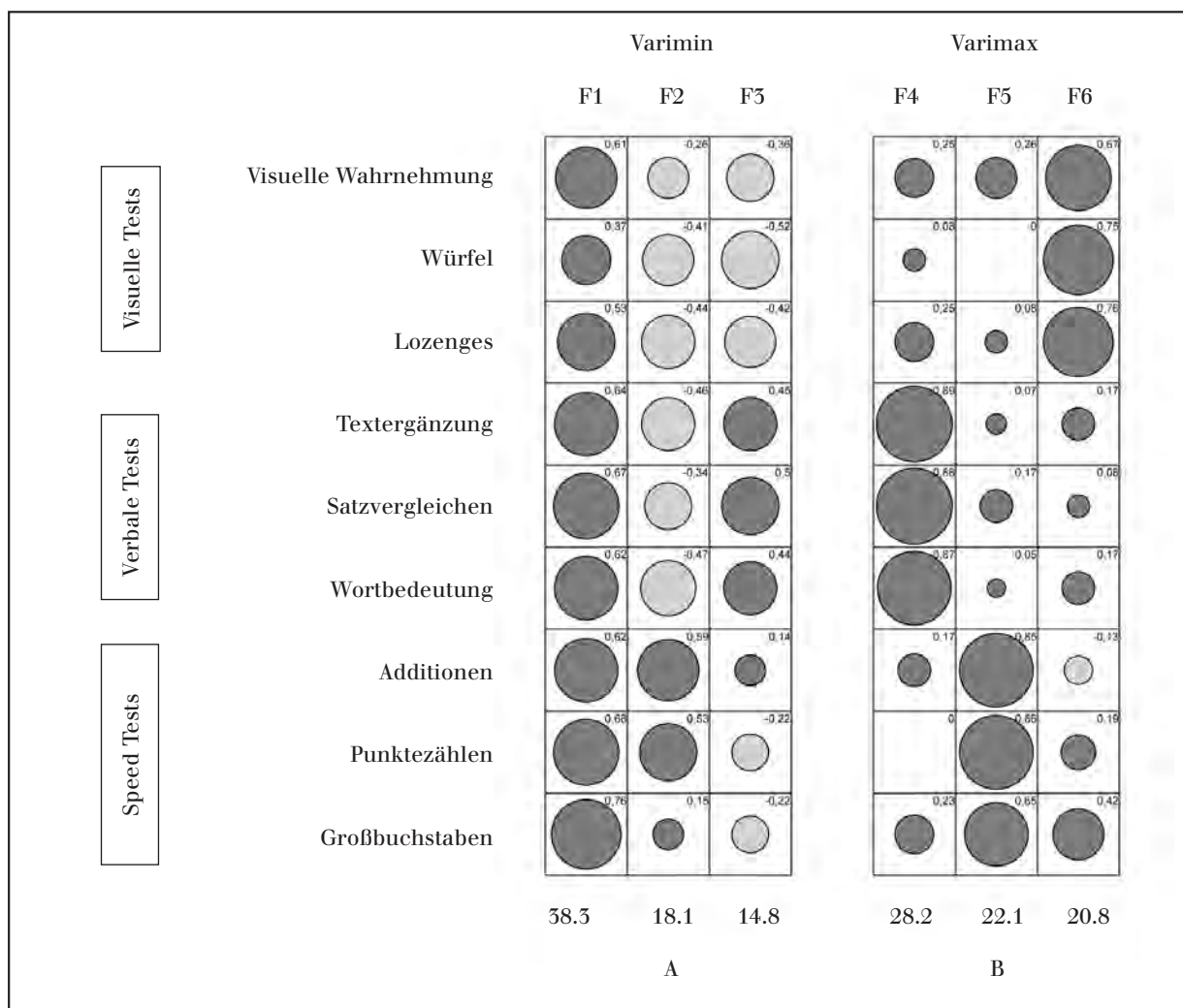


Abbildung 11: Varimin- (A) und Varimaxlösung (B) der Holzinger & Swineford-Daten (1939)

Die Varimax-Lösung (Abbildung 11 B) bringt ‚g‘ zum Verschwinden und separiert die drei Testsorten voneinander, so dass sich mit dieser Lösung kein Antagonismus zwischen verschiedenen Fähigkeiten und Neigungen manifestieren kann. Die LISREL-Lösung von Jöreskog & Sörbom kommt mit größerem Aufwand und mit mehr Eleganz zum gleichen Ergebnis wie die Varimax-Lösung von 11 B. Die Autoren hatten das Simple Structure-Modell mit den drei latenten Varianzquellen, deren Unabhängigkeit vorausgesetzt wurde, konfirmatorisch überprüfen wollen und waren dabei auch erfolgreich – allerdings auf Kosten der naheliegenderen und theoretisch plausibleren Lösung, zu der Varimin geführt hat.

Dass in bipolaren Lösungen nicht immer Variablen nur entweder mit positiven oder negativen Ladungen vorkommen, sondern dass sie auch Nahe-Null-Ladungen haben können, und dass dies informativ sein kann, zeigt das folgende Beispiel.

**Empirischer Beitrag 8:  
Psychophysiologische Aktivationsindikatoren**

(Daten von Köhler & Troester)

Den Autoren Köhler & Troester (1991) ging es um die Validierung des Maßes der Handflächen-Schweißproduktion (PSI, Palmar Sweat Index) als Indikator für psychophysiologische Aktivierung. Sie testeten 50 Personen unter drei Ruhebedingungen und einer Anstrengungsbedingung (von der Zahl 2007 hatten die Probanden sukzessive die Zahl 7 zu subtrahieren) und erhoben psychophysiologische Werte pro Person 16 mal im Laufe der vier Testphasen, nämlich PSI-Werte am Mittel- und Zeigefinger (PSI-M, PSI-F), spontane Fluktuationen der Schweißproduktion (SF), den Hautwiderstand (SCL, skin conductance level) und die Herzschlagfrequenz (HR). Die fünf erhobenen Werte wurden pro Person über die 16 Messwiederholungen intraindividuell korreliert, die Korrelationen wurden über die 50 Probanden gemittelt. Die Ergebnisse un-

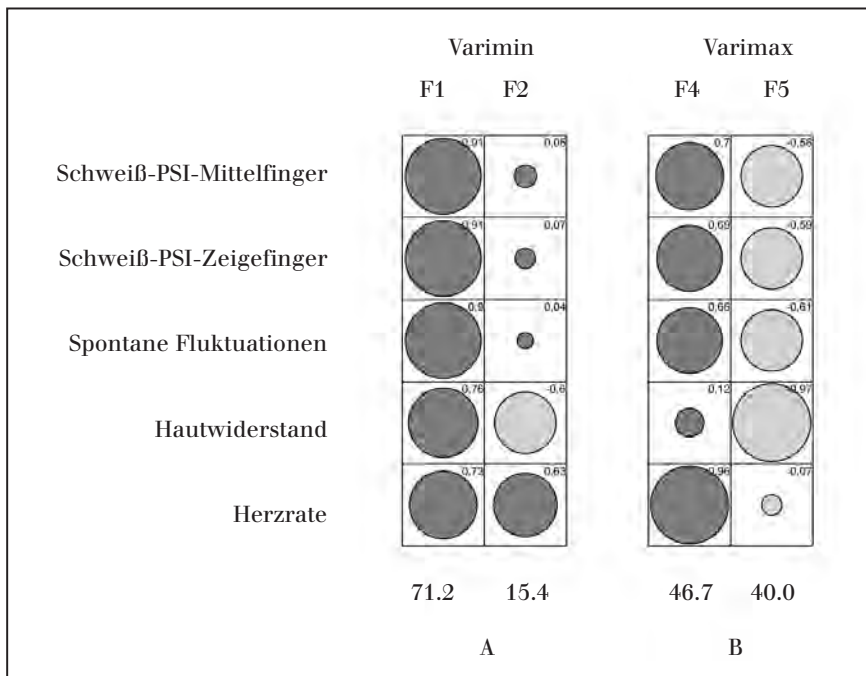


Abbildung 12:  
Varimin- (A) und Varimax-  
lösung (B) der Köhler &  
Troester-Daten (1991)

serer Standardauswertung der faktorisierten Interkorrelationen liefert Abbildung 12. Zwei Faktoren wurden nach dem Eigenwertverlauf 3.50, 0.76, 0.35, 0.20 ... und nach dem Ladungsmuster für interpretierbar gehalten, auf die Wiedergabe der Initiallösung, die der Variminlösung sehr ähnlich ist, wird verzichtet.

Ergebnisse: Varimin- $F_1$  ist ein Generalfaktor, der besagt, dass den individuellen Differenzen der fünf Maße offensichtlich der gleiche Indikatorwert – Aktivierung – zukommt. Varimin- $F_2$  repräsentiert Zusatzvarianz, die sich allerdings auf den Hautwiderstand und die Herzrate beschränkt, die Schweißvariablen haben mit dieser Varianzquelle offenbar nichts zu tun, dies zeigen ihre Um-Null-Ladungen an. Die  $F_2$ -Ladungen mit gegensätzlichen Vorzeichen bei SCL und HR besagen, dass die beiden Variablen offensichtlich ein leicht antagonistisches funktionales Verhältnis zueinander haben: Es gibt offenbar Probanden, bei denen der Hautwiderstand auf Aktivierung stärker reagiert als die Herzfrequenz und andere, bei denen umgekehrt die Herzfrequenz stärker reagiert als der Hautwiderstand. Obgleich beide Variablen mit  $F_1$  im wesentlichen die gleiche Aktivationsfunktion indizieren, bleibt ein geringer Rest an Varianz übrig, der durch eine leichte Präferenz entweder des Hautwiderstands oder der Herzrate aufgeklärt wird. Hier scheint ein Anteil ‚gegabelter Wirkung‘ vorzuliegen: Je mehr Wirkung in Richtung X vorliegt, umso weniger Wirkung ergibt sich für Richtung Y, und umgekehrt. Die Schweißproduktion ist von diesen koordinierten Präferenzwirkungen nicht betroffen.

Die Varimaxlösung stellt den Auswerter vor zwei unlösbare Rätsel. Die Aufspaltung des Aktivierungseffekts auf zwei unabhängige Äste ist das eine Rätsel –

aus physiologischen Gründen sind zwei unabhängige Aktivationsquellen kaum vorstellbar. Das zweite Rätsel ist die Zugehörigkeit von SRL zum einen Aktivationsast und von HR zum anderen Aktivationsast. Varimax vermischt hier die Varianzquellen bis zur Unkenntlichkeit.

Die bisherigen Beispiele zeigten: Die faktorenanalytische Auswertung von Variablen, die der hypothetischen Komplexität ihrer latenten Varianzquellen gerecht werden will, führt zu Ergebnissen, die leichter interpretiert werden können als Ergebnisse, denen man eine Einfachstruktur aufzuzwingen sucht. Die Ebene der latenten Bedingungen wird im Folgenden differenzierter in den Blick genommen.

#### Frage V

#### Sind mit Hilfe der Komplexstruktur-orientierten Faktorenanalyse auch methodenbedingte Einflüsse erfassbar?

Campbell & Fiske legten 1959 einen neuen methodologischen Ansatz vor, der in den darauf folgenden Jahrzehnten die Entwicklung zahlreicher methodischer Verfahren unter dem Label MTMM (multi trait multi method) nach sich zog. Die Autoren nahmen systematisch in Angriff, worauf man schon vorher, aber weniger nachdrücklich, aufmerksam gemacht hatte: Persönlichkeitsforscher sehen sich in ihrer Forschung nicht nur einer Vielheit von Varianz-erzeugenden latenten Traits gegenüber. Mit zusätzlicher Varianz haben sie zu rechnen, wenn sie zu deren Erfassung unterschiedliche Verfahren verwenden. Bald darauf bezog man auch die Varianzanteile bei Fragebögen mit ein, die durch unterschiedliche Informanten (Selbst- vs. Fremdbeurteilung) entstehen. Andere mögliche Vari-

anzeinflüsse, die an den Kovarianzen der gemessenen Variablen einen Anteil haben können, kamen hinzu (Situationseinflüsse bei Testwiederholungen usw.). Da sich das hier favorisierte komplexitätsoffene Verfahren beim Aufspüren multipler Varianzquellen bewährt hat, liegt es nahe, eine entsprechende Leistung auch für MTMM-Daten zu erwarten.

**Empirischer Beitrag 9:**

**Wissensprüfung mit Variation der Prüfmethode**

(Daten von Campbell & Fiske)

Zur Differenzierung methodischer Varianzquellen entwickelten Campbell & Fiske (1959) ein Verfahren, das auf eine systematische Inspektion der Interkorrelationsstabellen hinausläuft. Sie demonstrierten ihr Vorgehen anhand zahlreicher Korrelationsdaten, die sie in der Fachliteratur fanden. Ein Datensatz aus Campbell und Fiske, der schon von Cronbach und Vernon für andere Zwecke reanalysiert worden war – ihre Quelle geben sie nicht an – stammt von einem Wissenstest, bei dem im Physik-Unterricht vermittelte Inhalte zur Mechanik und Elektrizität durch rein verbale Fragen sowie durch Fragen mit Hilfe bildlicher Darstellungen geprüft wurden.

Die Ergebnisse unserer Auswertung für Varimin und Varimax zeigt Abbildung 13. Nach dem Eigenwertverlauf (2.67, 0.62, 0.45, 0.27) wäre nach üblichen Kriterien nur  $F_1$ , allenfalls noch  $F_2$  heranzuziehen. Doch Campbell & Fiske erkannten mithilfe ihrer Korrelationsinspektion recht deutlich die Auswirkung auch der beiden Varianzquellen, die in unserer Faktorenanalyse als  $F_2$  und  $F_3$  in Erscheinung treten.<sup>20</sup>

Ergebnis: Varimin- $F_1$  repräsentiert mit hohem Kommunalitätsanteil den Generalfaktor, zu dem wohl nicht nur die unterschiedliche Intelligenz der Probanden, sondern auch ihr unterschiedlicher Fleiß für das

Unterrichtsfach beigetragen haben. Hinzu kommt mit weitaus geringerem Gewicht  $F_2$  als Methodenfaktor (Sprache vs. Bilder) und  $F_3$  als Faktor der Lehrinhalte (Elektrizität vs. Mechanik). Unter den Probanden gab es offenbar solche, die mehr mit verbal ausformulierten Fragen als mit Abbildungen zurecht kamen und solche, denen Probleme mit Abbildungen leichter fielen ( $F_2$ ). Auch hatten einige wohl mehr für die Mechanik als für die Elektrizität gelernt und andere umgekehrt ( $F_3$ ). Wieder erweist sich hier Bipolarität als ein Indikator für zwei konkurrierende Gegebenheiten, nicht etwa nur für das Vorliegen oder Nicht-Vorliegen einer einzigen Gegebenheit.

Die Gegenläufigkeit der Wirkungen bei verbalem vs. bildlichem Prüfmaterial geht in der Varimaxlösung verloren. Ferner fehlt in der Varimaxlösung eine faktorielle Repräsentanz der Elektrizitätsinhalte, nur die Mechanikinhalte werden durch Varimax faktoriell repräsentiert. In einem letzten Beitrag soll die Frage geklärt werden, ob sich durch Varimin auch die Varianzquelle ausdrücken wird, die durch den Wechsel befragter Personenstichproben gegeben ist.

**Empirischer Beitrag 10:**

**Selbst- und Fremdbeurteilung von Kindern**

(Daten von Matson & Nieminen)

In einer Fragebogen-Untersuchung zu Verhaltensstörungen, Depression und Angst bei Kindern von Matson & Nieminen (1987) verwendeten die Autoren sechs Skalen, die den Kindern selbst sowie ihren Lehrern/Lehrerinnen vorgelegt wurden, welche die Kinder zu beurteilen hatten. Das Ergebnis unserer Auswertung zeigt Abbildung 14. Die Eigenwerte betragen 3.98, 2.14, 1.41, 0.99, 0.86..., die Prozentanteile an aufklärter Varianz bei den Initialfaktoren 35.1, 17.8, 11.8, 8.2..., wonach drei Faktoren als substantiell zu bezeichnen sind.

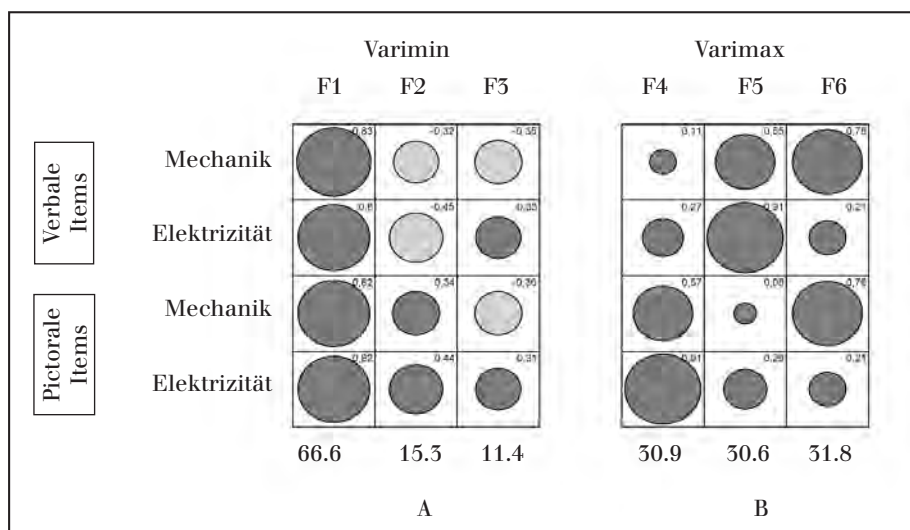


Abbildung 13: Varimin- (A) und Varimaxlösung (B) der Campbell & Fiske-Daten (1959)

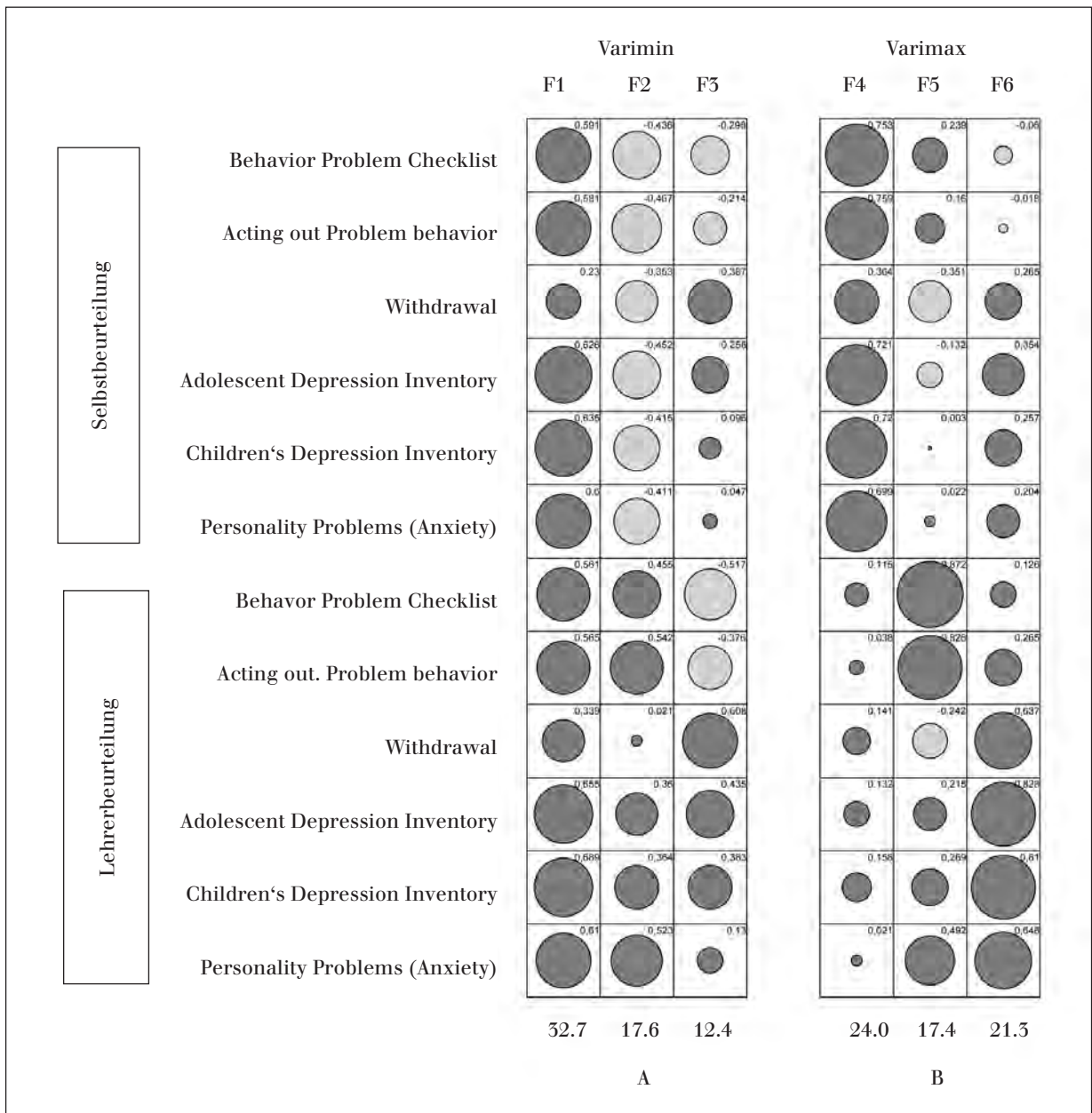


Abbildung 14: Varimin-(A) und Varimaxlösung (B) der Matson & Nieminen-Daten (1987)

Zum Ergebnis der Variminlösung, Abbildung 14A: Die im Generalfaktor  $F_1$  sich ausdrückende Gemeinsamkeit der Kovarianzquelle bei Kindern und Lehrern lässt mit etwas Zurückhaltung darauf schließen, dass in den Urteilen tatsächlich vorliegende dysfunktionale Symptome bei den Kindern mit ihrer interindividuellen Varianz zum Ausdruck gebracht werden. Zurückhaltung ist geboten, weil sich bei den meisten Fragebogenuntersuchungen, so auch wohl in dieser, im ersten Faktor zusätzlich die Akquieszenz als Antwortungsstil geltend machen kann. Dass in dieser Untersuchung die Variablen mit  $F_1$  bei Kindern und Lehrern unterschiedslos geladen sind, ist kein stichhaltiger

Einwand, da Akquieszenz bei Kindern und Lehrern gleichermaßen ausgeprägt sein kann. Zusätzlich kann zu  $F_1$  bei Kindern und Lehrern ein Einfluss der *social desirability* (SD) beigetragen haben der nicht erkennbar ist, da alle Skalen negativ bewertetes Erleben und Verhalten zum Inhalt haben.

Mit gegensätzlichen Vorzeichen beim bipolaren  $F_2$ -Faktor macht sich offenbar der Informanten-Einfluss geltend. Dieser ist als solcher zwar nicht zweifelhaft, seine Herkunft aber bleibt mehrdeutig. Es mag sein, dass die Kinder sich selbst tendenziell als weniger erlebungs- und verhaltensgestört beurteilen als die Lehrer, oder auch als gestörter, oder dass sie zu mehr

oder zu weniger Akquieszenz oder SD neigen als die Lehrer. Auch können sich diese Einflüsse ähnlich wie die Teil-Einflüsse bei  $F_1$  addieren.

Mit dem Vorzeichen-Gegensatz des bipolaren Faktors  $F_3$  werden die dysfunktionalen Symptome inhaltlich differenziert, bei Kindern und Lehrern in gleicher Weise. Auf der einen Seite gruppieren sich die Störungen mit enthemmter Auswirkung (acting out, conduct disorder), auf der anderen Seite die Störungen mit gehemmter Wirkung (withdrawal, depression, anxiety). Auch hier ist eine antagonistische Beziehung anzunehmen, dysfunktionale Dispositionen können sich je nach Veranlagung entweder als Hemmungen ausdrücken oder aber zu einem enthemmenden ‚acting out‘ führen (Aggressionen z. B.), wobei Angst ‚abgeführt‘ wird, was psychoanalytisch so gedeutet werden könnte.

Die Varimaxlösung (Abbildung 14 B) führt zu drei Clustern der Schüler- und Lehrerskalen. In einem Cluster sind die sechs von den Schülern beurteilten Skalen vereint, in einem zweiten die drei von den Lehrern beurteilten Skalen der Enthemmungsstörung, in einem dritten Cluster die zwei von den Lehrern beurteilten Skalen der Hemmungsstörung, wobei eine Reihe von Ausreißern und mehrfaktorielle Ladungen das Bild beeinträchtigen. Mit diesen Clustern lässt sich konzeptuell kaum etwas anfangen.

Die Reanalyse der Daten von Matson & Nieminen hat nochmals deutlich gemacht, dass der Ausdruck ‚latent‘ (‚unsichtbar‘, ‚verborgen‘) zur Kennzeichnung der Faktoren, die den Variablen zugrunde liegen, nicht immer zutrifft. Die Verwendung wechselnder Beurteiler liegt als Bedingungsfaktor offen zutage. Ihr Einfluss, wenn ein solcher vorliegt, sollte sich durch eine Analyse, welche Varianzquellen aufzudecken hat, ebenso zeigen wie der Einfluss zunächst unbekannter Varianzquellen, oder auch nicht zeigen, wenn kein solcher Einfluss vorliegt. Im vorliegenden Fall konnte die Interpretation von Varimin- $F_2$  (Informanten) auf einen transparenten Einfluss zurückgreifen, die Interpretation von  $F_3$  (Enthemmungs- vs. Hemmungsstörung) musste sich mit einer weniger transparenten inhaltlich-psychologischen Unterscheidung und Abstraktion begnügen.

### 5 Diskussion

Fünf Fragen zur Methodologie wurden behandelt, die sich mit der Einführung des Rotationsverfahrens Varimin ergeben. Zehn empirische Anwendungsbeispiele sollten helfen, sie zu klären.

Gegen die Auswahl der zehn Beiträge könnte man vorbringen, dass sie zu wenig Datensätze aus psychologischen Untersuchungen enthielt. Doch sollte die

Güte einer neuen Methode zunächst mit Daten geprüft werden, bei denen zweifelsfrei bestimmte Ergebnisse zu erzielen sind, sofern die Methode funktioniert. Wenn ein zu erwartendes Ergebnis nicht erzielt wird, soll man schlussfolgern dürfen, dass mit der Methode etwas nicht stimmt. Von dieser Prüfstrategie machen Methodiker unter den Psychologen m. E. generell zu wenig Gebrauch. Psychologische Daten, zumal einschlägige sprachliche Daten, deren Ergebnisse oft mehrdeutig bleiben, eignen sich zur Methodenprüfung kaum. Die Sprache hat zwar zur Differenzierung von Erlebens- und Verhaltensweisen eine Fülle an Wörtern und Wendungen hervorgebracht. Doch auf eine transparente sprachliche Semantik wie die der Verwandtschaftsnamen kann sich die Kommunikation über psychologisch relevante Sachverhalte nicht stützen.

Natürlich soll die Faktorenanalyse, wenn sie ihre methodischen Prüfungen bestanden hat, bei der Lösung inhaltlicher Probleme unserer Disziplin eingesetzt werden. Aus den inzwischen vorliegenden Ergebnissen der Methodenforschung lässt sich schon jetzt eine bedeutsame Schlussfolgerung voraussagen: Auch dem manifesten verbalen Urteilsmaterial, das Psychologen zur Analyse einbringen, liegen auf latenter Ebene generell Komponenten- oder Merkmalstrukturen zugrunde. Die Merkmale oder Komponenten psychisch funktionaler Einheiten sind nur viel schwerer zu identifizieren und voneinander zu separieren als die Konzepteinheiten, die die Erfassung der Gegebenheiten unserer anschaulich gegliederten Lebenswelt unterstützen.<sup>21</sup> Deshalb wird man sich bei der Interpretation von Varimin-Faktoren einer psychologischen Domäne nicht, wie gewöhnlich bei der Interpretation von Simple Structure-Faktoren, mit Faktorbenennungen begnügen dürfen, die unter den verwendeten *manifesten Variablen* (Items, Statements) ausgewählt werden oder mit diesen synonym sind. Man wird sich auf die Ebene der latenten Merkmale begeben und sich trennen müssen von Begriffen umgangssprachlich festgeschriebener psychologischer Konzepte, z.B. von Trait-Konzepten wie *conscientious*, *agreeable*, *open-minded* usw., weil nicht sie, sondern deren Varianzquellen gefragt sind. Man wird sich ähnlich verhalten müssen wie bei der Analyse der Verwandtschaftsdomäne, bei der man zur Entschlüsselung der Faktoren auf eine ‚latente‘ Ebene hinabsteigen und sich von den manifesten Begriffen *Mutter*, *Schwester*, *Tante* usw. trennen musste, weil nicht diese Begriffe selbst oder deren Cluster, sondern deren *distinctive features* gefragt waren.

Die empirischen Beispielfälle, die relativ transparente Domänen zur Analyse wählten, haben gezeigt, dass man Varimin-Faktorlösungen durchaus deuten kann, obgleich die den manifesten Variablen zugrunde liegenden Varianzquellen komplexe Strukturen

darstellen. Die Deutungen waren für Variminlösungen weit befriedigender als für Varimaxlösungen. Mit Anwendung des minimalen Paarvergleichs, der bei schwierigen Variminfaktor-Deutungen vorgenommen werden kann, beschränkt man sich auf die Feststellung von Unterschieden zwischen gepaarten Variablen hinsichtlich nur eines Merkmals, was durchweg leichter fällt als das Auffinden von Gemeinsamkeiten bei unanalysierten Merkmalkomplexen, die durch Varimax gebildet werden.

Drei Einschränkungen sind indes nachzutragen. Erstens, die Anwendung des Minimalpaar-Vergleichs setzt voraus, dass im Datensatz geeignete Variablenpaare vorhanden sind. Oft gewinnt man sichere Deutungen eines Faktors erst mit mehreren Minimalpaaren, die verschiedene Merkmalprofile aufweisen. Da man dieses Desiderat bei der Variablenselektion im voraus meist nicht berücksichtigen kann, wird mitunter ein Faktor nicht sicher gedeutet werden können, weil minimale Paare fehlen. Die noch fehlenden Paare können sich erst mit einer repräsentativeren Variablen-Stichprobe einstellen.

Das Vorkommen minimaler Paare ist umso eingeschränkter, je weniger Variablen verwendet wurden und je mehr substantielle Faktoren extrahiert wurden, so dass der Bedarf an Kombinationen gegensätzlicher Faktorgewichte zur Bildung von Minimalpaaren nicht gedeckt ist. Deshalb empfiehlt es sich, bei der Zusammenstellung von Variablenstichproben für eine faktorielle Untersuchung einer Domäne die Zahl der vermuteten Varianzquellen eher etwas zu beschränken, aber innerhalb des gewählten Vermutungsbereichs Bedingungsvariationen großzügig zuzulassen.

Die zweite Einschränkung resultiert aus dem Umstand, dass die Varianzquellen, die mit Varimin gefunden werden, immer einen höheren Grad an Abstraktheit besitzen als die Begriffe, die man gewöhnlich zur Beschreibung von Simple Structure-Faktoren verwendet. Mit den Verwandtschaftswörtern lässt sich das andeuten: Das Merkmal ‚*Generation*‘, mit Varimin gewonnen, ist abstrakter als ein Begriff wie ‚*die Jüngsten und Ältesten in der Abstammungslinie*‘, der bei der Varimax-Analyse notwendig war, um Großeltern und Enkel unter einen faktoriellen Hut zu bringen. Man wird in der psychologischen Domäne mit der Identifizierung abstrakter Merkmale größere konzeptuelle Schwierigkeiten haben als mit den konkret typisierenden Begriffen, wie sie in der Simple Structure-Praxis verwendet werden. Auch unter *Möbeln* z. B. kann man sich nicht so schnell etwas vorstellen als unter *Kleiderschrank* und *Polstersessel*.<sup>22</sup>

Schließlich ist die Minimalpaarbildung in der Regel bei Generalfaktoren nicht anwendbar. Die Generalfaktor-Ladungen der Variablen einer Domäne sind durchweg positiv, und sie weisen meist wenig Varianz

auf. Andererseits repräsentiert ein Generalfaktor in der Regel die für den Datensatz jeweils dominante Varianzquelle, an der ein Forscher in erster Linie interessiert ist und die meist transparenter ist als hinzu kommende schwächere Varianzquellen. Man wird einen Generalfaktor bei der Analyse von z. B. psychophysiologischen Variablen selbstverständlich auf z. B. Aktivationsvarianz zurückführen und nicht etwa auf Intelligenzvarianz.

Vor allem bei Generalfaktoren ist indes immer auch daran zu denken, dass sich mit ihnen mehr als nur eine Varianzquelle ausdrücken kann. Zum Beispiel kann sich in einem Generalfaktor bei Intelligenztests *Intelligenz plus Leistungsehrgeiz* ausdrücken, bei Fragebögen kann sich im Generalfaktor die mit den Items erfragte Traitdisposition zusammen mit einer mehr oder weniger ausgeprägten Akquieszenzneigung ausdrücken. Über die relativen Anteile der an einem Faktor beteiligten ‚Teil-Varianzquellen‘ ist aus den Daten selbst oder aus ihrer formalen Repräsentation keine Information zu gewinnen. Die Komplexität latenter Bedingtheiten kann sich mit Varimin nur in dem Maße differenziert ausdrücken, wie sich die relevanten Varianzquellen mit *unterschiedlichen* Beiträgen an der Varianz der verwendeten Variablen beteiligen.

Dass sich die Interpretierbarkeit initialer komplexer Strukturen, wenn diese mehr als zwei Faktoren aufweisen, durch Varimin-Rotationen verbessert, haben einige Beispiele verdeutlicht. Mir ist bei zahlreichen faktoriellen Reanalysen kein Fall begegnet, bei dem das Ergebnis einer Varimin-Transformation *weniger* gut interpretierbar gewesen wäre als das Ergebnis der vorausgehenden Initiaillösung.<sup>23</sup>

Was die Interpretierbarkeit von *Simple Structure*-Strukturen betrifft, so ließ diese bei den ausgewählten Beispielen viel zu wünschen übrig. In der Praxis würde man bei manchen der hier verwendeten Datensätze auf Faktorenanalysen nicht zurückgreifen, man würde auf andere Verfahren, vornehmlich auf MDS (bzw. NMDS) oder MTMM oder Circumplex-Verfahren ausweichen. Der Vorzug des Komplexität-modellierenden Verfahrens ist unter anderem darin zu sehen, dass es sich zur Analyse der verschiedensten Datensätze eignet, während die Anwendbarkeit des Simple Structure-Verfahrens aufgrund seiner Modelliermängel von vorne herein begrenzt ist.

Analysen mit Simple Structure-Rotationen können indes *auch* nützlich sein. Ihre Bedeutung besteht nur nicht in dem, was man bisher glaubte, dass mit ihnen ‚Dimensionen‘ der untersuchten Domänen aufgedeckt würden, sondern darin, dass sie die Variablen einer Domäne clustern, während gleichzeitig die clusterbildenden Basismerkmale verborgen bleiben. Nur wenn eine Fragestellung zu bearbeiten sein sollte, die das Clustern von kovariierenden Variablen erfordert,



ohne dass die Quellen ihrer Kovarianz interessieren, würden sich Varimax und die obliquen Rotationsverfahren – diese insbesondere – als hilfreich erweisen.

Varimax könnte darüber hinaus nützlich bleiben für den Fall, dass Variablen zur Analyse anstehen, von denen man vorweg weiß, dass ihnen nur *eine* Varianzquelle zugrunde liegt. Das wäre z. B. der Fall, wenn die mit Varimin aufgedeckten latenten Merkmale, z. B. durch Itembildung für Fragebögen, zu manifesten Variablen erhoben werden, wenn diese dann für eine Korrelationsstudie eingesetzt werden, und wenn die so gewonnene Matrix einer Faktorenanalyse übergeben wird.

Komplikationen können für Varimin-Untersuchungen auch dann entstehen, wenn a priori unbekannt ist, ob den zur Analyse ausgewählten Variablen tatsächlich eine größere Zahl von Varianzquellen zugrunde liegt. Bislang wurde davon ausgegangen, dass Variablen in aller Regel multifaktoriell bedingt sind. Dies aber schließt Ausnahmen nicht aus, und man weiß nicht, wie viele Ausnahmen man zulassen darf und wie sie als solche zu erkennen sind.

Von Ungewissheit bei der Deutung Varimin-rotierter Faktoren war schon die Rede bei der Behandlung der Frage, ob und wie *negative* Faktorladungen zu deuten sind. Ob bei den mit Plus- und Minusvorzeichen bipolar gekennzeichneten Strukturen funktional antagonistische Verhältnisse vorliegen oder ob sich in den Vorzeichen lediglich das Vorhandensein bzw. Nichtvorhandensein einer latenten Bedingung ausdrückt, darüber sagt die Analyse selbst nichts aus, darüber entscheidet man mit einschlägigem Kontextwissen und den daraus sich ergebenden Mutmaßungen.

Bei der Verwendung des neuen Analyseverfahrens ist zu beachten – wie zuvor angedeutet –, dass die Variablen-Stichprobe für die Domäne, deren Kovarianzquellen differenziert werden sollen, repräsentativ ist. Die Variablen sollten von vorne herein als Ausdruck einer Mehrzahl latenter Bedingungen verdächtig sein, die unter den Domänvariablen Kovariationen hervorrufen. Zumindest sollte man Variablen nicht aufnehmen, die plausiblerweise nur wenig Aufschluss über Kovariationsquellen liefern. Würde man z. B. altersabhängige Variablen unter die Variablen eines Persönlichkeitsfragebogens mischen („ich werde immer vergesslicher“, „mir geht es gesundheitlich nicht so gut wie früher“), dann würden diese möglicherweise einen eigenen Faktor („Altersanzeichen“) bilden. Diese würden zwar mit Plusladungen einen Beitrag zur Varianz bekunden – wenn ein solcher vorliegt –, sie würden aber bei den Variablen, die mit Alter nichts zu tun haben, vermutlich Minusladungen hervorbringen („Einfluss nicht vorhanden“). Mit aufschlussarmen Minus-Ladungen bei diesen Variablen würde Kommunalität ‚verschwendet‘ werden.<sup>24</sup> Re-Analysen

mit Varimin-Rotation führen mitunter durchaus auch zu unbefriedigenden Ergebnissen. Diese sind jedoch meist auf unzureichende Variablen-Stichproben zurückzuführen. Mithilfe eines Measures of Sampling Adequacy (Kaiser, 1970) ließe sich ggf. die Adäquatheit von Datensätzen für exploratorische Faktorenanalysen routinemäßig ermitteln.

Zur weiteren Agenda gehört die Klärung des Invarianzproblems. Die Frage bleibt zu klären, ob das Varimin-Verfahren zu Lösungen führt, die gegenüber Hinzufügungen zur Variablenstichprobe und gegenüber Streichungen aus dieser oder auch gegenüber einem Wechsel der untersuchten Personenstichproben stabil genug sind. Eine nicht-publizierte Untersuchung (Ertel, b) ergab deutlich mehr Stabilität für Varimin als für Varimax beim Hinzufügen einer artifiziellen Variablen zu einer zuvor analysierten Stichprobe natürlicher Variablen.<sup>25</sup>

Zu den noch zu bearbeitenden Fragen gehört auch die nach der optimalen Faktoren-Extraktionsmethode. Im bisherigen Projekt wurde ausschließlich die Hauptkomponentenmethode (PCA) verwendet, weil sie sich in der faktorenanalytischen Forschung zum Standard entwickelt hat. Es mag sein, dass das Einsetzen von Einsen in die Diagonalfelder der Korrelationsmatrizen, das zur PCA gehört, auf die Endergebnisse u. U. eine leicht verzerrende Wirkung hat, vor allem dann, wenn überwiegend niedrige Interkorrelationen vorliegen. Zu den weiteren Aufgaben der Methodenforschung gehört, die Hauptachsen- und das Maximum-Likelihood-Verfahren vergleichsweise auf ihren Beitrag zur Modelliergüte zu prüfen und ggf. weitere Verbesserungen des Vorgehens vorzunehmen, falls durch eine andere Wahl der Extraktionsmethode am Ergebnis noch etwas zu verbessern sein sollte.<sup>26</sup>

Die vorliegende Arbeit beschränkte sich auf die Entwicklung von Varimin als einem Verfahren, das für die exploratorische Forschung eine neue Richtung zeigt. Nachdem sich das Verfahren in zahlreichen Prüfungen mit verschiedenen Datensätzen bewährt hat, wird sein Einsatz bei der systematischen Bearbeitung psychologischer Sachfragen zum Desiderat. Zwei abgeschlossene Projekte zur Frage der Varianzquellen der Persönlichkeit (Ertel, c) und der Intelligenz (Ertel, b) liegen vor, sie würden grünes Licht vonseiten innovationsbereiter Zeitschriften-Herausgeber und ihrer Manuskript-Begutachter begrüßen.<sup>27</sup>

Abschließend sei ein Einwand bedacht, der sich richten könnte gegen die dem vorliegenden Ansatz zugrunde liegende Strategie, über formale Fragen der Datenverarbeitung (hier: Faktorenanalyse) anhand von Ergebnissen ihrer Anwendung auf inhaltliche Problemfelder zu urteilen. Mit ‚weichen‘ Argumenten des alltäglichen Lebens, wie man sie abwertend nennen

könnte, würde man einem präzisen mathematischen Instrumentarium nichts anhaben können.

Diesem Einwand würde man entgegen halten dürfen, dass es in dieser Arbeit nur darum ging, einen einzigen Fehler der Methodiker zu korrigieren: die Dogmatisierung von *Simple Structure* als Modell für latente Bedingungen bei multivariaten Analysen. Dieses Modell ist ein formales, doch es wurde aus *nicht-formalen* Gründen eingeführt, denn die Behauptung seiner Notwendigkeit (simple structure – Transformation mache Faktorstrukturen erst interpretierbar) ergibt sich, wenn man die Güte der Methode auf die Inhalte der mit ihr gewonnenen Ergebnisse (auf deren ‚Interpretierbarkeit‘) bezieht.<sup>28</sup> Dieser Bezug erfordert psychologisches *Verstehen*, die Mitverwendung von Erkenntnis, in welcher sich neben bisherigen Ergebnissen wissenschaftlicher Forschung, die oft kontrovers bleiben, eine ungleich größere Menge an ‚weicheeren‘ Welt- und Menschen-Erfahrungen aufgebaut haben, von denen die meisten im Leben erfolgreich genutzt werden. Thurstonianer geraten mit ihrer Begründung des faktoriellen Rotationsprinzips, mit der sie dem nicht-formalisierten ‚Verstehen‘ Konzessionen einräumen, in die Nachbarschaft anderer Forscher, die sich auch – vielleicht mehr noch als sie – dem ‚Verstehen‘ verpflichtet fühlen. Auf dem Forum, auf dem sich die Positionen der Modellformalisten mit denen der weniger streng und eng denkenden Forscher verträglich überschneiden, sollten Letztere ein Mitspracherecht geltend machen dürfen. Sie sollten auf mögliche Fehler der Formalisten achten dürfen, soweit sie deren Sprache verstehen, sollten die Fehler prüfen, die sie vermuten, sie durchdenken, wenn sie nicht wegzudenken sind, und, soweit es in ihren Kompetenzbereich fällt, auch helfen dürfen, sie zu korrigieren.

### Anmerkungen

- 1 Die Faktorenextraktion durch PCA führt dazu, dass die Interkorrelationen der Initialfaktoren bzw. die Kongruenzindices, mit denen man ihre Ähnlichkeit untereinander ermittelt, alle exakt Null ergeben. Wenn aber bei suboptimalen Variablen-Stichproben die latenten Komponenten, die die Korrelationen unter den manifesten Variablen herbeiführen, miteinander korreliert auftreten, führt der Zwang zur Null-Korrelation der Faktoren, die die Komponenten manifestieren sollen, zu einer gewissen Verzerrung dieser Manifestationsleistung. Diese soll durch die Faktorenrotation Varimin korrigiert werden, d.h. nach Anwendung von Varimin sind  $\approx 0$ -Inter-Kongruenzen der Faktoren feststellbar.
- 2 Die zusätzlich übliche Unterscheidung hinsichtlich Artikulationsmodus (plosiv, frikativ, nasal usw.) ist spezieller und kann mit der beabsichtigten Phonem-Stichprobe nicht ermittelt werden. Das [m] und das [f] sind beides Phoneme mit längerer Artikulationsdauer (im Unterschied z. B. zu den kurz artikulierten plosiven Lauten [b], [p], [t], [k]), doch [m] gehört nach seinem Artikulationsmodus zu den nasalen Lauten, [f] zu den frikativen Lauten. Da alle plosiven Phoneme kurz sind, hätte man auch eine bipolare Kategorie ‚plosive‘ vs. ‚nicht-plosive‘ Phoneme bilden und diese als Faktor im Beurteilungsversuch erwarten können.
- 3 Die hier bevorzugte Forschungsstrategie bei der faktorenanalytischen Arbeit lässt sich unmittelbar verknüpfen mit theoretischen und empirischen Ansätzen, in denen Begriffe und andere kognitiv repräsentierte Objekte als Bündel oder Strukturen mehr oder weniger latenter Komponenten aufgefasst werden (‚Komponentenmodell‘ von Einstellungsobjekten, Feger, 1971, ‚Feature Pattern Analysis‘, Feger & Brehm, 2001). Auch läuft sie in die gleiche Richtung wie die Bemühungen um das Phänomen des Ähnlichkeitserlebens, zu dessen Erklärung bei Shepard, 1980, ‚hidden structures‘ und bei Tversky, 1977, ‚collections of features‘ analysiert wurden.
- 4 Umgangssprachlich ersetzt oft *one pence* das eigentlich richtige *one penny*.
- 5 Die leicht variierenden Ladungshöhen bei  $F_2$  (am meisten fällt das 2 *Pence*-Stück aus der Reihe) könnten mit münzspezifischen subjektiven Gewichtungen der Merkmale zusammenhängen. Tatsächlich wirkt das 2-*Pence* Stück angesichts seines geringen Wertes übermäßig groß, so dass bei ihm im Vergleich mit anderen Münzen wahrscheinlich der Größeneindruck stärker als der Formeindruck zur Geltung kommt.
- 6 Die Interpretation von  $F_2$  als Faktor der anschaulichen Münzengröße wird gestützt durch Korrelationen der  $F_2$ -Ladungen mit der gemessenen Münzengröße:  $r = .81$  (T. S.),  $r = .85$  (Studenten). Das Gewicht der Münzen in Gramm, das mit ihrer Größe in cm bei  $r = .80$  korreliert, korreliert mit den Faktorladungen nur geringfügig.
- 7 Zimmermann (1955) bemerkte, dass Variablen, die in der unrotierten Lösung mit ähnlichem Profil mehrfaktoriell geladen sind, sich durch Rotation zur Einfachstruktur in einem Faktor vereinigen: „... a test which actually contains variance on two or more factors may appear with all of that variance confined on a single factor“. Diesen nennt er ‚composite factor‘. Der Autor hält also an

- der inhaltlichen Bedeutung der initialen Faktoren (Centroidfaktoren) fest, die er im rotierten Faktor als Faktorenbündel („composed“) vereint sieht. „*It is my feeling that the failure to give composite factors the attention they merit must be considered either a serious oversight or a serious error or omission*“ (Zimmermann, 1953, 389).
- 8 Dass die Rotation zur Einfachstruktur nicht zur Analyse zugrunde liegender Dimensionen, sondern nur zur Clusterung ähnlicher Testvariablen führt, wurde explizit von Overall (1964) nicht nur behauptet, sondern anhand eines Analysebeispiels demonstriert: „*Rotation to simple structure can be understood as an elaborate approach to cluster analysis. It identifies clusters of tests which measure the same things, but there is no assurance that these „same things“ are simple and primary dimensions*“ (p. 271). „... *there is no need to assume that simple structure factors will correspond to any particular set of fundamental dimensions of the objects ...*“ (p. 276).
- 9 Die Probanden hatten 16 Verwandtschaftswörter nach selbst gewählten Kriterien zunächst in zwei beliebig große Teilgruppen zu teilen, dann jede der beiden Teilgruppen wieder in zwei Teile usw., bis jedes Wort an einem unteren Ende in der Hierarchie vereinzelt übrig blieb. Die Ähnlichkeit zweier Worte für einen Probanden wurde definiert als die Anzahl der Schritte, die nötig war, um aus der isolierten untersten Position der beiden zu vergleichenden Wörter in der Gruppierungshierarchie aufsteigend einen ersten gemeinsamen Knoten zu erreichen. Tabelle 2 enthält die über die Probanden summierten ‚Schritte bis zum ersten gemeinsamen Knoten‘.
- 10 Diese lineare Transformation wurde lediglich zum leichteren Verständnis der aus der Tabelle abgelesenen Einzelwerte vorgenommen.
- 11 Eigenwerte: 4.31, 2.34, 1.54, 1.09, 1.05, 0.90, 0.79, 0.75...
- 12 Man könnte  $F_1$  als Ausdruck der Verwandtschaftskategorie deuten. Seine differenzierende Funktion würde faktoriell vermutlich erst in Erscheinung treten, wenn man in dieser Untersuchung zusätzlich Wörter einer anderen Kategorie verwendet hätte.
- 13 Die Interpretationen der Varimin-Faktoren  $F_2$ ,  $F_4$  und  $F_5$  sind auch in einer vergleichbaren MDS-Untersuchung von Romney & D’Andrade (1964) anzutreffen (gender, generation, consanguinity). Die vorliegende Untersuchung differenziert noch innerhalb  $F_2$  (Linealität) zwischen ‚Kernfamilie‘ und ‚Nicht-Kernfamilie‘ (durch  $F_3$ ). Die Untersuchung von Romney & D’Andrade bietet ansonsten
- einen vorzüglichen Einstieg in die begrifflichen und methodischen Grundlagen der ‚componential analysis‘ von Begriffen und ein Verständnis für die Schlussfolgerungen, die man aus der Existenz der ‚discriminative stimuli‘ (der ‚Sememe‘) für den Begriffserwerb ziehen kann.
- 14 Es wäre denkbar, dass Probanden, die viele extreme Ablehnungen abgeben, auch viele Unentschieden-Urteile oder viele gemäßigte Zustimmungen abgeben, was motivational aber nicht nachvollziehbar wäre.
- 15 Aus diesem Grund sind multivariate Kausalmodelle, wie sie mit Methoden der konfirmatorischen Faktorenanalyse und mit Strukturgleichungsverfahren Eingang fanden, als fragwürdig zu betrachten. Die Modellierung durch multivariate Analysen zielt in erster Linie auf multiple Quellen der Kovarianz manifester Variablen ab. Das Zusammenwirken der latenten Quellen lässt sich nicht ohne weiteres als ein kausales verstehen.
- 16 Da die ‚Zuweisung‘ von positiven und negativen Vorzeichen durch das Rotationsprogramm vom zugeordneten Inhalt der Messungen unabhängig erfolgt, ist auch darüber zu entscheiden, ob die initial gegebenen Vorzeichen beibehalten oder der jeweiligen Deutung entsprechend ggf. umgekehrt werden sollten. Die Variablen, bei denen der Faktor funktional hemmenden Einfluss ausdrückt bzw. bei denen das faktoriell repräsentierte Merkmal als nicht vorhanden zu interpretieren ist, können zugunsten eines leichteren Verstehens ein negatives Vorzeichen erhalten. Es hat sich gezeigt (bislang eine Daumenregel), dass das Nichtvorliegen eines Merkmals oder der negative Einfluss durch eine Varianzquelle innerhalb eines faktoriellen Vektors tendenziell eher bei denjenigen Variablen vorkommt, die niedrigere Ladungen haben und deren Vorzeichen (Plus oder Minus) überwiegen.
- 17 Ein aus der Frühzeit der Centroid-Faktorenanalyse stammender Hinweis, der den unrotierten bipolaren Faktoren einen Informationswert einräumt, der damals schon vernachlässigt wurde, findet sich in einer Notiz von Zimmermann (1953): „*It is well-known in dealing with intellectual variables that the first centroid loadings are usually all positive and the second centroid, as well as those that follow, divide positive and negative variables equally. What is apparently overlooked is the tendency for the second centroid to split the most obvious dichotomy, the third centroid to split the next most obvious dichotomy and so on. For example, if the battery contains both linguistic and*

- quantitative tests, the second centroid will most likely separate these two major groups...*“ (p. 389).
- 18 Die Simple Structure-Modellierung psychologischer Daten hat vielfach zur Auflösung bipolarer Konstrukte geführt: Urteile über negativ getönte Gefühle scheinen nach faktorieller Einfachstruktur-Modellierung statistisch unabhängig zu sein von Urteilen über positiv getönte Gefühle und sollen dann entgegen aller Lebenserfahrung generell so erlebt werden (Diener & Emmons, 1985, Watson & Clark, 1988, späte Korrektur nach Modellvergleichen durch Crawford & Henry, 2004). Ähnlich zusammenhangslos erscheinen in Faktorenanalysen mit Einfachstruktur-Rotation Optimismus und Pessimismus, die normalerweise bipolar sich auswirkenden Einstellungen zur Zukunft (Marshall et al., 1992). Auch geht die Polarität der Geschlechter-Typologie durch Faktorenanalysen verloren: Femininität und Muskulinität soll aufgrund faktorieller Orthogonalität nicht polar aufeinander bezogen sein (Bem 1981).
- 19 Man könnte die Tests mit zwei alternativen Zusatzinstruktionen durchführen lassen, eine Speed-Instruktion würde die Probanden generell zu mehr Tempo ermuntern, eine Power-Instruktion zu mehr bedachtsamer Konzentration. Vermutlich wird man je nach Testart differenzielle Steigerungen bzw. Verminderungen ihrer Leistung beobachten.
- 20 Die Eigenwertkriterien der Mindesthöhe (>1) bzw. der Eigenwertknick (zu ermitteln durch den Scree Plot) sind schon von anderer Seite als unzuverlässig erkannt worden (Fabrigar, 1999, p. 287). Der u.a. auch von Fabrigar empfohlene 'Paralleltest' ist mit den vorliegenden Korrelationsdaten nicht durchführbar. Nach eigenen Erfahrungen mit faktoriellen Reanalysen von Datensätzen haben sich hier als informativ erwiesen die Prozentanteile, die die Initialfaktoren an der Summe der Kommunalitäten haben. Ein Faktor, dessen Anteil an der Summe  $h^2$  weniger als 10% beträgt, ist in der Regel nicht mehr interpretierbar. Im vorliegenden Datensatz von Campbel & Fiske betragen die Kommunalitätsprozente der vier Initialfaktoren 66.7, 15.4, 11.5, 6.7. Wenn ein Mindestanteil von 10% Kommunalität als Kriterium angesetzt wird, sind  $F_2$  und  $F_3$  noch zu berücksichtigen.
- 21 „*Characterizing concepts in terms of features works very well for certain types of words ... [those] that are very well structured by physical, social, or biological dimensions, which then serve as the basis of semantic features. Other types of words ... which do not come from well-structured domains, are more difficult to characterize as a set of features. ... No methodology directly reveals the meaning components of a word. Despite the difficulty in confirming what is or is not a feature of a word, it is generally accepted that ... words are comprehended by accessing various features or meaning components*“ (1984, Just & Carpenter, p. 63 f).
- 22 In einem Projekt mit Reanalysen von NEO PI-R-Daten, das mithilfe der Varimin-Transformation von Faktoren zu einer neuen Konzeptualisierung von Persönlichkeitskomponenten führen sollte (Ertel b), wurden die Begriffe ‚Aktivationslevel‘, ‚Aktivationsgradient‘, ‚Regulationsquelle‘, ‚Repräsentationsmodus‘, ‚Funktionsbilanz‘ zur Deutung der Varimin-Faktoren vorgeschlagen. Abstrakte Begriffe dieser Art liegen nicht immer fertig vor, sie sind schwer zu bilden, doch sind sie leichter miteinander verknüpfbar.
- 23 Die Gesamtzahl der Datensätze, an denen das Varimin-Programm erprobt wurde, liegt inzwischen bei über 500.
- 24 Die Faktorladung einer Varianzquelle, die bei einer Variablen nur ein ‚nicht vorhanden‘ aussagt, wird dennoch bei der betreffenden Variablen als Beitrag zu ihrer Kommunalität mit aufgenommen eine Art Informationsverschwendung. Dies ist, zumal wenn es sich um eine Varianzquelle handelt, die nur für eine Minorität von Variablen als ‚vorhanden‘ relevant ist.
- 25 Einen Mangel an Invarianz von Simple Structure-Lösungen beklagte schon Butler (1969): „... *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*“ (p.13). „*Normal varimax factors cannot be regarded as factorially invariant...*“ (p. 24)
- 26 Dies ist indessen nach bislang vorliegenden Vergleichen der Extraktionsverfahren kaum zu erwarten: „*The major conclusion of this article is that there is little basis to prefer component analysis or factor analysis. For practical purposes the choice of method is not a decision that will greatly affect empirical results or substantive conclusions*“ (Velicer & Jackson, 1990, p. 19).
- 27 Der weiteren faktorenanalytischen Forschung mit Varimin durch Interessenten stelle ich auf Anfrage das von Dr. Uwe Engeland in meinem Auftrag erstellte Varimin-Varimax-Programm faktor.exe zur Verfügung, wobei eine kurze Beschreibung des Ziels der jeweiligen Untersuchung, in der das Programm eingesetzt werden soll, möglicherweise hilfreichen Empfehlungen von mir nach sich ziehen. Was seine Nutzerfreundlichkeit angeht, ist faktor.exe überarbeitungsbedürftig. Derzeit sind für seine erstmalige Nutzung noch beträchtliche Instruktionen erforderlich.

28 „... consistent psychological meaning is by far the most important criterion for the success of a factor analysis that is designed to illuminate psychological phenomena. The simple structure criterion was designed only as a means to that end“ (Guilford & Hoepfner, 1969, 6). „The arguments in favor of rotation are not mathematical; and in each research the investigator has to decide its merits on non-mathematical grounds“ (Burrighs & Miller, 1961, p. 35).

## Literatur

- Bem, S. L. (1981). Gender schema theory: A cognitive account of sex typing. *Psychological Review*, 88, 354-364.
- Burrighs, G. E. R. & Miller, H. W. L. (1961). The rotation of principal components. *The British Journal of Statistical Psychology*, 14, 35-49.
- Butler, J. M. (1969). Simple structure reconsidered: Distinguishability and invariance in factor analysis. *Multivariate Behavioral Research*, 4, 5.
- Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Carl, W. (1968). Eine Untersuchung zur Faktorenstruktur von Antworttendenzen („response sets“) bei Antwortskalen verschiedener Stufenzahl. *Zeitschrift für experimentelle und angewandte Psychologie*, 15, 419-434.
- Comrey, A. L. (1978). Common methodological problems in factor analytic studies. *Journal of Consulting and Clinical Psychology*, 46 (4), 648-659.
- Crawford, J. R. & Henry, J. D. (2004). The Positive and Negative Affect Schedule (PANAS): Construct validity, measurement properties and normative data in a large non-clinical sample. *British Journal of Clinical Psychology*, 43, 245-265.
- Diener, E. & Emmons, R. A. (1985). The independence of positive and negative affect. *Journal of Personality and Social Psychology*, 47, 5, 1105-1117.
- Ertel, S. (2009). Kritik am Dogma „Simple Structure“. *Journal Psychologie des Alltagshandelns*, 1, Vol. 2, 36-65.
- Ertel, S. (a) Intelligenz (g) und Lernkapital. Komplexstrukturelle Faktorenanalyse des I-S-T (unveröffentlicht).
- Ertel, S. (b) Faktorielle Komplexstruktur durch Varimin-Rotation. Manifestation simulierter Zusatzvarianz in I-S-T-Daten. (unveröffentlicht).
- Ertel, S. (c) Varianzquellen der Persönlichkeit. Exploratorische Faktorenanalyse mit komplextheoretischer Ausrichtung (unveröffentlicht).
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4, 272-299.
- Feger, H. (1971). Einstellungsstruktur und Einstellungsänderung: Ergebnisse, Probleme und ein Komponentenmodell der Einstellungsobjekte. *Zeitschrift für Sozialpsychologie*, 10, 331-349.
- Feger, H. & Brehm, M. (2001). *New developments in feature pattern analysis*. Lengerich: Pabst Science Publishers.
- Gorsuch, R. L. (1974) *Factor analysis*. Philadelphia: Saunders.
- Guilford, J. P. (1952). When not to factor analyze. *Psychological Bulletin*, 49, 26-37.
- Guilford, J. P. & Hoepfner, R. (1969). Comparison of varimax rotations with rotations to theoretical targets. *Educational and Psychological Measurement*, 29, 3-23.
- Harman, H. H. (1967). *Modern factor analysis*. Chicago: The University of Chicago Press.
- Herrmann, T. (1965). *Psychologie der kognitiven Ordnung*. Phänomenologisch-Psychologische Forschungen. In C. F. Graumann & J. Linschoten (Eds.), (Vol. 6). Berlin: Walter de Gruyter & Co.
- Hinz, A., Brähler, E., Geyer, M., & Körner, A. (2005). Urteileffekte beim NEO-FFI. *Diagnostica*, 49, 157-165.
- Holzinger, K. J. & Swineford, F. (1939). A study of factor analysis: The stability of a bi-factor solution. *Supplementary Educational Monographs* 48. Chicago: University of Chicago, Department of Education.
- Humphreys, L. G. & Davey, T. C. (1988). Continuity in intellectual growth from 12 months to 9 years. *Intelligence*, 12, 183-197.
- Jacobson, R. & Halle, M. (1956). *Fundamentals of Language*. Den Haag: Mouton.
- Jöreskog, K. G. & Sörbom, D. (2005) LISREL 8.52. *Website*: www.ssicentral.com.
- Just M. A. & Carpenter, P. A. (1984). *The psychology of reading and language comprehension*. Boston: Allyn and Bacon.
- Kaiser, H. F. (1970). A second generation Little Jiffy. *Psychometrika*, 35, 401-416.
- Köhler, T. & Troester, U. (1991). Changes in the palmar sweat index during mental arithmetic. *Biological Psychology*, 32, 143-154.
- Marx, A. & Hejj, A. (1989). *Subjektive Strukturen. Ergebnisse aus der Sprach- und Einstellungsfor-*schung. Göttingen: Hogrefe.

- Marshall, G. N., Wortman, C. B., Kusulas, J. W., Hervig, L. K., & Vickers, R. R. (1992). Distinguishing optimism from pessimism: Relations to fundamental dimensions of mood and personality. *Journal of Personality and Social Psychology*, 62 (6), 1067-1074.
- Matson, J. L. & Nieminen, G. S. (1987). Validity of measures of conduct disorder, depression, and anxiety. *Journal of Clinical Child Psychology*, 16 (2), 151-157.
- Overall, J. E. (1964). Notes on the scientific status of factors. *Psychological Bulletin*, 61, 270-276.
- Rasch, D. (1962). Die Faktorenanalyse und ihre Anwendung in der Tierzucht. *Biometrische Zeitschrift*, 4, 15-39.
- Reise, S. P., Waller, N. G., & Comrey, A. L. (2000). Factor analysis and scale revision. *Psychological Assessment*, 12 (5), 287-297.
- Romney, A. K. & D'Andrade R. G. (1964). *Transcultural Study of Cognition. Special publication of the American Anthropologist*. Vol. 66, Part 2.
- Shepard, R. N. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science*, 210, 390-398.
- Tversky, A. (1977). Features of similarity. *Psychological Review*, 84, 6, 327-348.
- Überla, K. (1971). *Faktorenanalyse. Eine systematische Einführung für Psychologen, Mediziner, Wirtschafts- und Sozialwissenschaftler*. 2. Auflage, Berlin: Springer.
- 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.
- Watson, D., & Clark, L. A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS Scales. *Journal of Personality and Social Psychology*, 54 (6), 1063-1070.
- Wellek, A. (1959). *Die Polarität im Aufbau des Charakters*. Bern: Francke.
- Zimmermann, W. S. (1955). A note on the recognition and interpretation of composite factors. *Psychological Bulletin*, 50, 387-389.
- 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

# Instructions to authors

## Kinds of contributions:

The journal *Psychology of Everyday Activity* publishes the following formats:

### *Original contributions*

Original contributions contain results of empirical research, method developments, or theoretical reflections (max. 40,000 characters).

### *Research reviews*

Research reviews encompass the current state of research considering a specific subject (max. 50,000 characters).

### *Research notes*

Research notes represent pilot studies or replications, or inform about new research projects or research programs and their first results (max. 20,000 characters).

### *Discussion*

Discussion contributions take argumentatively position on a discussion-worthy topic with reference to psychological research or practice (max. 20,000 characters).

### *Book reviews*

Reviews refer to a new published work from the psychological research or practice (max. 8,000 characters).

## Submission of manuscripts:

Manuscripts should consider the usual guidelines of manuscript design of the German Society of Psychology (DGPs) or the American Psychological Association (APA). However, in contrast to the guidelines of DGPs or APA we ask the authors to set those passages which should appear in *italics* in the printing version *already* in the manuscript version in italics. The abstract should not exceed 1,000 characters. Contributions can be written in German or English language. If a manuscript is written in German language, both a German *and* an English abstract should be submitted. Following the abstract, up to six Keywords should be listed. In German contributions the keywords should indicated both in German and in English.

In order to ensure an anonymous review, the names of the authors should appear only on the title page.

Tables and figures should be numbered and attached separately at the end of the manuscript. The place in which the respective table or figure shall be inserted should be marked in the manuscript text.

Please submit your manuscripts to the following email-address:

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

We need a version in pdf-format as well as a version in .doc (e.g., Word) or .rtf.

### *Specimen copy*

The first authors receive one issue in which the paper has been published as well as a pdf-copy of their article.

# 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 erschienenen 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.