

Here today, gone tomorrow? Revisiting the stability of teachers' achievement goals

Anna-Katharina Praetorius ^{a,*}, Sebastian Nitsche ^b, Stefan Janke ^b, Oliver Dickhäuser ^b, Katharina Drexler ^a, Michaela Fasching ^a, Markus Dresel ^a

^a Department of Psychology, University of Augsburg, Augsburg, Germany

^b Department of Educational Psychology, University of Mannheim, Mannheim, Germany

1. Introduction

Motivational characteristics are a powerful explanation for how and why people think and act as they do. Many researchers conceptualize motivational characteristics as dispositional characteristics of persons. One famous example is research on motive dispositions, i.e., habitual preferences for dealing with certain kinds of incentives (Heckhausen & Heckhausen, 2010). Several theories and models, however, also point to the influence of occasion-specific characteristics on actual motivation (e.g., model of adaptable learning, Boekaerts & Niemimirta, 2000; cognitive-motivational process model, Vollmeyer & Rheinberg, 1998; for a general overview of the relationship between traits and states, see latent-state-trait theory, Steyer, Schmitt, & Eid, 1999).

In research on teacher motivation, achievement goals are seen as an important explanation for teachers' perceptions of the environment and for their actions (Butler, 2012; Nitsche, Dickhäuser, Fasching, & Dresel, 2011; Retelsdorf & Günther, 2011). It is usually assumed (e.g., Butler, 2007; Dresel, Fasching, Steuer, Nitsche, & Dickhäuser, 2013; Malmberg, 2008; Retelsdorf, Butler, Strebblow, & Schiefele, 2010) that these achievement goals can be seen as traits (as "goal orientations") and, therefore, are only influenced by occasion-specific characteristics to a limited degree. Teachers teach-

ing the same class in the same school should thus set themselves similar goals across several occasions, largely independent of situational circumstances. However, few investigations have tested this assumption empirically. Taking a contrary position regarding the stability of achievement goals, Elliot (2005) stated that the main difference between the achievement goal approach and the classical achievement motive is that the former has a more specific and contextual focus. Increasing our knowledge about the actual stability of teachers' achievement goals will facilitate the development of both an appropriate theoretical understanding of teachers' achievement goals and, subsequently, an adequate model of the construct. More concretely, knowledge about the stability of teachers' achievement goals is important for the following reasons: (a) It helps insure that investigations will capture the characteristics of interest in an appropriate way (e.g., choosing a cross-sectional versus a longitudinal design). (b) It points out how to construct adequate measurement instruments (e.g., general versus situation-based measurements). (c) It helps in the selection of appropriate research questions regarding the level of operationalization of the independent and dependent variables (see Ajzen & Fishbein, 1977). If achievement goals are only stable to a small degree, effects on rather stable characteristics (e.g., teachers' content knowledge) are unlikely. However, if teachers' achievement goals are stable to a large degree, investigating the effect on variable characteristics (e.g., instructional behavior in specific situations) does not seem to be straightforward. (d) Additionally, knowledge about the stability of teachers' achievement goals is useful in deriving appropriate implications based on the results of investigations (see also Murphy

* Corresponding author. Department of Psychology, University of Augsburg, Universitätsstraße 10, 86159 Augsburg, Germany. Fax: +49 821 598 5289.
E-mail address: anna.praetorius@phil.uni-augsburg.de (A.-K. Praetorius).

& Alexander, 2000; Pintrich, 2000). The purpose of the study at hand is to shed light on this topic by investigating the stability of teachers' achievement goals as well as the number of measurement points necessary to reliably measure these goals across occasions.

1.1. Teachers' achievement goals: definition and relevance

Achievement goals explain how and why people behave the way they do in achievement settings (Dweck & Leggett, 1988; Murayama, Elliot, & Friedman, 2012). Achievement goal theory differentiates between various goals. The achievement goals that are commonly distinguished when describing and explaining characteristics of teacher motivation are (a) learning goals (the teacher aims to increase his or her own competencies), (b) performance approach goals (the teacher aims to demonstrate high competencies), (c) performance avoidance goals (the teacher aims to avoid the impression of low competencies), and (d) work avoidance goals (the teacher aims to reduce his or her workload). Several studies have pointed out the relevance of teachers' achievement goals for the teaching profession, as they have revealed relationships between teachers' achievement goals and various teacher and teaching characteristics. Associations have been found, for example, between achievement goals and occupational burden or burn-out (Nitsche, Dickhäuser, Fasching, & Dresel, 2013; Retelsdorf et al., 2010; Tönjes & Dickhäuser, 2009), the perception of help-seeking as beneficial or threatening (Butler, 2007; Nitsche et al., 2011), participation in vocational training programs (Nitsche et al., 2013), aspects of instructional quality (Butler, 2012; Butler & Shibaz, 2008; Retelsdorf & Günther, 2011), and the goal structures teachers realize in their classrooms, i.e., the extent to which pursuing learning vs. performance goals for students is reinforced by the classroom environment (Butler, 2012; Dresel et al., 2013; Retelsdorf et al., 2010). In all of these studies, learning goals were positively correlated with variables that are considered to be beneficial (e.g., attending vocational training programs) and negatively correlated with variables that are regarded as adverse (e.g., burn-out). For performance avoidance goals as well as work avoidance goals the relationship pattern was, in most

cases, reversed. The results regarding the effects of performance approach goals were mixed (e.g., positive effects on teacher self-efficacy in a study by Nitsche et al., 2011; positive effects on the social reference norm in a study by Retelsdorf & Günther, 2011). Most of the relationships identified between teachers' achievement goals and other variables were small, some were moderate.

1.2. Teachers' achievement goals: stable characteristics?

Regarding the conceptualization of achievement goals, large differences can be found (for an overview, see Pintrich, 2000, and Maehr & Zusho, 2009; see also Button, Mathieu, & Zajac, 1996; DeShon & Gillespie, 2005). According to some conceptions, achievement goals are assumed to be rather stable (e.g., Silva & Nicholls, 1993) whereas for other conceptions, they are assumed to be rather unstable (e.g., Elliott & Dweck, 1988). The differences regarding the assumed stability of achievement goals are important: The theoretical conception of achievement goals influences (a) how investigations concerning these goals are conducted (e.g., how many measurement points are used), (b) how they are measured (e.g., whether achievement goals are assessed with respect to specific situations), (c) what research questions are appropriate (e.g., whether it makes sense to investigate effects of achievement goals on stable outcomes), and (d) what implications are derived from the results of the investigations (Murphy & Alexander, 2000; Pintrich, 2000; for a similar argumentation regarding intrinsic and extrinsic motivation, see Harter & Jackson, 1992).

Based on the considerations of Fryer and Elliot (2007) and Pintrich (2000), we developed a conceptual framework to explain why achievement goals, on the one hand, can be assumed to be stable but, on the other hand, are also assumed to be unstable (see Fig. 1). Fryer and Elliot (2007) and Pintrich (2000) do not differentiate between different achievement goals in their argumentation. This implies that differences in the stability of the goals are not expected. Additionally, no information is given regarding the expected magnitude of the stable and the unstable components of achievement goals.

Empirical investigations regarding the stability of achievement goals exist, first and foremost, for students' achievement goals (for an overview, see Fryer & Elliot, 2007; Pintrich, 2000). According to

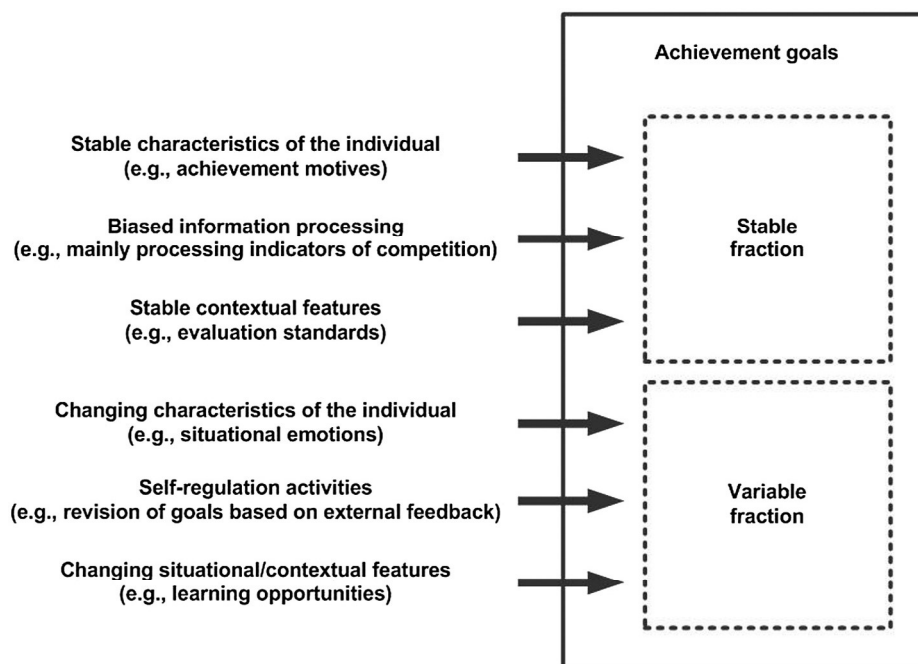


Fig. 1. Conceptual framework for the stability of achievement goals.

Senko, Hulleman, and Harackiewicz (2011), retest correlations range between $r = .40$ and $r = .70$ for both learning and performance goals among students.

Regarding the stability of teachers' achievement goals, very few considerations and empirical investigations exist. The assumptions and empirical results regarding students' achievement goals cannot, however, simply be transferred to teachers, as the achievement settings of students and teachers in the school context are completely different.

Teachers' achievement goals have been conceptualized, to date, as rather stable tendencies to adopt certain goals. This is especially obvious in publications which use the term "goal orientation" (e.g., Butler, 2007; Dresel et al., 2013; Fasching, Dresel, Dickhäuser, & Nitsche, 2010; Malmberg, 2008; Nitsche et al., 2011; Retelsdorf et al., 2010), meaning a general orientation towards goals that includes beliefs about purposes, competence, success, ability, effort, errors, and standards (Pintrich, 2000). As a consequence, teachers' achievement goals are usually assessed using one measurement point, with self-report measures that survey general achievement goals without considering the specific situations teachers are in when reporting these goals. Interpretations of the results of the investigations, as a rule, deem that teachers have a certain stable and dominant achievement goal.

Empirical investigations regarding the degree to which occasions influence measures of teachers' achievement goals are rare. Most of them focus on student teachers (i.e., students carrying out their university studies) or teacher trainees (i.e., students in a practice-orientated training phase following university graduation) rather than on teachers. The rank-order stability findings of these studies are summarized in Table 1. The large variations reported for retest correlations ($.26 \leq r \leq .71$) indicate that the stability of student teachers', teacher trainees', and teachers' achievement goals is not clear. Furthermore, the varying correlations point to the fact that different kinds of achievement goals might be differently stable for teachers. One observed tendency is that learning goals are less stable than performance goals. Possible reasons for these differences are not discussed in the studies considered.

In the study conducted by Fasching et al. (2010), mean-level and intra-individual stabilities of teacher trainees were investigated in addition to rank-order stabilities. Variance analyses revealed that the mean levels of learning goals, performance approach, and performance avoidance goals decreased over the course of the two investigated years, whereas there was no significant variability with regard to work avoidance goals. ICCs were calculated as an estimate of intra-individual variability. Learning goals showed a considerably higher intra-individual variation ($1 - ICC = .72$) compared to performance approach goals ($1 - ICC = .53$), performance avoidance ($1 - ICC = .58$), and work avoidance ($1 - ICC = .52$). As the sample consisted solely of teacher trainees, it remains unclear as to what extent these results can be expected with regard to teachers: The contexts of the two groups differ considerably. Teacher trainees are, for example, assessed several times within the training period

whereas teachers are assessed only very rarely in their daily school experiences.

1.3. Determining the stability of achievement goals using generalizability theory

In most investigations, the stability of achievement goals has been calculated using retest correlations. This method has, however, been criticized because rank-order stabilities cannot be interpreted unambiguously (see Fryer & Elliot, 2007). One disadvantage of analyzing retest correlations is that they confound true state variance (i.e., true score variance that is specific for a certain measurement point) and measurement error. Retest correlations are therefore of restricted utility when investigating the stability of measures, as it is not possible to obtain information on the magnitude of occasion-specific goals.

A solution for this disadvantage is generalizability theory (G theory), a statistical framework in which different aspects of stability (e.g., mean level changes, variation in the values of achievement goals across occasions) can be estimated simultaneously. With G theory, a comparison between the magnitudes of stable and unstable components of the measurement is also possible. This, in turn, gives an impression of how significant the instability of teachers' achievement goals is for the measurement. Another advantage of G theory is that one can estimate the number of measurement points necessary to assess teachers' achievement goals reliably across occasions (see Shavelson & Webb, 1991). To date, teachers' achievement goals are assessed using one measurement point without knowing whether this is sufficient to capture them reliably.

Beyond G theory, there are several other alternatives one can use to investigate the occasion specificity of measures, for example latent state-trait theory (see Steyer et al., 1999). We chose G theory for this investigation as it not only allows for the separation of trait and state variance, but also helps determine the number of measurement points necessary for a reliable measure.

1.4. Research questions and hypotheses

Up to now, we have been able to reveal very little concerning the variability of teachers' achievement goals. The few existing studies have mainly investigated student teachers and teacher trainees.

As the existing studies on the stability of student teachers', teacher trainees', and teachers' achievement goals have shown a large variability in the stability of achievement goals, we did not derive a hypothesis but pursued the following research question instead: To what degree do teachers' achievement goals vary across measurement points over the course of one school year (research question 1)? To answer this research question, we used G theory to take into consideration different aspects of stability.

Based on the assumption that teachers' achievement goals are stable characteristics, these goals, as a rule, were collected at one

Table 1
Overview of studies analyzing the stability of teachers' achievement goals using retest-correlations.

Literature	Sample	Number of time points	Retest interval	Achievement goal			
				Learning	Performance approach	Performance avoidance	Work avoidance
Fasching et al. (2010)	German teacher trainees	5	6 months	.26–.42	.41–.66	.46–.51	.48–.57
Tönjes and Dickhäuser (2009)	German teachers	2	3 months	.55	.61	.63	–
Tönjes and Dickhäuser (2009)	German teacher trainees	2	12 months	.58	.71	.63	–
Malmberg (2008)	Finnish teacher trainees	5	12 months	.37 ^a	.41 ^a	.35 ^a	–

Note: "–" = This goal was not investigated in the study.

^a These correlations are the mean correlations across the four measurement points. None of the reported studies investigated the stability of relational goals.

measurement point in prior research and were then related to other variables (e.g., instructional variables). However, no previous empirical studies have tested whether one measurement point is sufficient to reliably measure teachers' achievement goals. We therefore added the research question: How many measurement points are needed for a reliable measure of teachers' achievement goals (research question 2)?

2. Method

2.1. Sample and procedure

We invited 288 academic-track secondary schools ("Gymnasien") in the German federal state of Baden-Württemberg to participate in the study in the school year 2011/12 via postal letters; 57 of these schools (46 public schools; 11 private schools) decided to take part in the study. Of the participating schools, 13 were located in urban and 44 in rural areas. The participation rate per school varied between one and five teachers as only mathematics teachers teaching in 5th grade classrooms were to participate. We restricted the study to 5th grade classrooms to ensure comparability. For the analyses, we used data from all teachers participating in the study; in total 166 German mathematics teachers (55% female) were assessed. The mean age of the teachers was 41 years ($SD = 13$) at the first measurement point. The teaching experience of the teachers ranged from 0 to 40 years ($M = 13$; $SD = 12$).

The teachers completed a questionnaire at three points over the course of one school year (directly after the summer break in September 2011, December 2011, and March 2012).

2.2. Instruments

In order to measure the achievement goals of teachers, a self-report questionnaire developed by Nitsche et al. (2011) was used. This measure was developed based on existing measures (e.g., the Achievement Goal Questionnaire by Elliot & Murayama, 2008; Goal Orientations for Teaching by Butler, 2007). The instrument formulated by Nitsche et al. (2011) is specifically tailored to the population of teachers and explicitly considers the different aspects of teachers' professional knowledge (e.g., pedagogic content knowledge) to which teachers' learning goals can be directed, and the different relevant others (e.g., school principal) to which teachers can address their performance goals. The instrument is reliable and has been proven to possess factorial, convergent/divergent, and predictive validity (Nitsche et al., 2011). It has since been used in a number of studies (e.g., Dresel et al., 2013; Fasching et al., 2010; Nitsche et al., 2013).

The instrument consists of four achievement goal scales: (a) one scale assessing learning goals with three subscales focusing on different domains of teacher knowledge (pedagogic knowledge, content knowledge, pedagogic content knowledge), each comprising three items (e.g., "In my vocation, I aspire to improve my pedagogic knowledge and competence"), (b) a scale for performance approach and (c) a scale for performance avoidance goals, each with four subscales focusing on different significant addressees (colleagues, principal, students, self) and each with three items (e.g., "In my vocation, I aspire for my students to realize that I teach better than other teachers" [approach]; "In my vocation, I aspire to not show my students when I have more trouble meeting the job demands than other teachers" [avoidance]), and (d) a scale for work avoidance goals with six items (e.g., "In my vocation, I aspire to get through the day with little effort"). For the work avoidance scale, all items conceived by Nitsche et al. (2011) in their first version of the instrument were used. All items were presented alongside 5-point Likert-type scales ranging from 1 (strongly disagree) to 5 (strongly agree).

2.3. Analyses

To answer research question 1 regarding the degree of instability in teachers' achievement goal measures, G theory (Shavelson & Webb, 1991) was used. This method allows for a separation between trait and state true score variance via variance component analysis (Brennan, 2001; Shavelson & Webb, 1991). The objects of measurement in the G analyses were the persons and their achievement goals. Occasions and item parcels were added as facets (i.e., sources of error). As each person was assessed using all item parcels on all occasions, the design was fully crossed. Persons, item parcels, and occasions were treated as random in all of the analyses. Thus, a two-facet, fully crossed random effects design ($p \times i \times o$ design) was applied: persons (p) crossed with item parcels (i) crossed with occasions (o).

We used item parcels for the following reason: It cannot be assumed that the subscales of the instruments (e.g., knowledge facets addressed with learning goals) are interchangeable as they focus on different aspects of achievement goals. To be able to use random effects, we used systematically combined item parcels instead of subscales (Kishton & Widaman, 1994; Little, Cunningham, Shahar, & Widaman, 2002). One item of each subscale of the respective achievement goal was integrated into a parcel. Therefore, the number of items per parcel was two for the work avoidance goals, three for the performance approach and performance avoidance goals, and four for the learning goals.

G theory designs can be illustrated using Venn diagrams. As can be seen in Fig. 2, the design allowed us to separate seven sources of variance in the data: variance resulting from (a) the persons (σ^2_p), (b) the item parcels (σ^2_i), (c) the occasions (σ^2_o), (d) the interaction between persons and item parcels (σ^2_{pi}), (e) the interaction between persons and occasions (σ^2_{po}), (f) the interaction between item parcels and occasions (σ^2_{io}), and (g) the interaction between persons, item parcels, and occasions ($\sigma^2_{pio,e}$). The interaction mentioned last is confounded with an unspecific error component as it is the highest order interaction. The variance components can be grouped into: (a) stable inter-individual differences regarding achievement goals (= trait variance; σ^2_p , σ^2_{pi}), (b) intra-individual differences regarding achievement goals across measurement points due to specific occasions (= state variance; σ^2_o , σ^2_{po} , σ^2_{io}), and (c) residual variance ($\sigma^2_{pio,e}$) as well as further variance components which are not relevant for the topic in question (σ^2_i). In estimating the dependability (i.e., reliability) of the measurement in question, two G coefficients are available in G theory. Both coefficients can be interpreted as being analogous to classic reliability coefficients. Following the rule of thumb in classic test theory, .70 is set as a minimum value for a reliable measure in the present investigation. The absolute G coefficient (ϕ) is

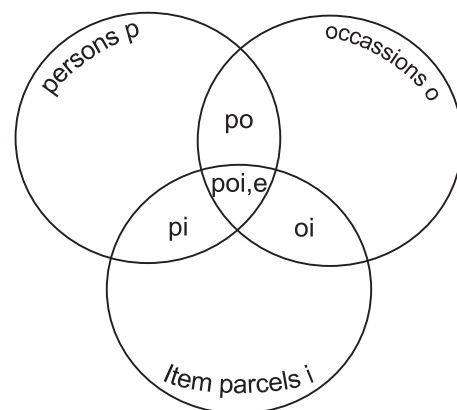


Fig. 2. Venn diagram for the persons \times item parcels \times occasions design.

Table 2
Descriptive statistics of teachers' achievement goals.

Achievement goals	<i>M</i>	<i>SD</i>	α	r_{12}	r_{13}	r_{23}
Learning				.55	.67	.67
Measurement point 1	4.14	0.44	0.82			
Measurement point 2	4.16	0.51	0.88			
Measurement point 3	4.15	0.48	0.87			
Performance approach				.74	.70	.77
Measurement point 1	1.96	0.74	0.93			
Measurement point 2	2.09	0.80	0.95			
Measurement point 3	2.18	0.80	0.95			
Performance avoidance				.75	.71	.81
Measurement point 1	2.49	0.81	0.91			
Measurement point 2	2.47	0.84	0.93			
Measurement point 3	2.58	0.85	0.94			
Work avoidance				.61	.64	.78
Measurement point 1	2.09	0.71	0.82			
Measurement point 2	1.89	0.73	0.87			
Measurement point 3	1.82	0.70	0.88			

Note: Retest intervals were 3 months respectively. *N* = 150–163 teachers.

adequate if one is interested in absolute outcomes (e.g., competencies in large-scale assessment studies). The relative G coefficient (ρ^2) is adequate if one focuses on the relative position of persons or variables and not on the absolute values. As this is the case in the present investigation, the relative G coefficient is reported in the following.

The G analyses were conducted with the urGENOVA program (Brennan, 2001), version 2.1. In urGENOVA, the implemented estimator is the analogous ANOVA procedure. A large advantage of this estimator is that normality assumptions are not required (Brennan, 2001; see also the simulation study of Shumate, Surles, Johnson, & Penny, 2007). Missing data are handled in urGENOVA by adding an additional facet to the data (see Brennan, 2001) and thus are explicitly considered.

To determine the number of measurement points necessary for a reliable measure of teachers' achievement goals (research question 2), D analyses were conducted. These analyses enable researchers to estimate reliability under multiple measurement conditions (e.g., differing numbers of measurement points), and thus provide evidence regarding how many observations of a certain variable are necessary to obtain a sufficient reliability (Brennan, 2001). These analyses are based on the estimated variance components of the G analyses; thus, the information from the G analyses is used to estimate the number of necessary measurement points. The estimation provided by the D analyses works analogously to the

Spearman–Brown formula in classical test theory (Webb, Shavelson, & Haertel, 2006).

The D analyses were conducted using the GENOVA software (Crick & Brennan, 1983). The implemented estimator was the ANOVA procedure.

3. Results

3.1. Descriptive analyses

Table 2 separately presents the mean scores, standard deviations, and internal consistencies for the three investigated measurement points. The internal consistencies were admissible. The retest correlations for the five achievement goals can also be found in Table 2. The correlations were, on a descriptive level, lowest for learning goals ($.55 \leq r \leq .67$) and highest for performance avoidance goals ($.71 \leq r \leq .81$).

3.2. Stability of teachers' achievement goals over time

The estimated variance components (see Table 3) can be grouped into components measuring achievement goals that are occasion unspecific and, thus, stable, and into components that are occasion specific and, thus, unstable. The stable components (i.e., stable differences between persons (σ_p^2) as well as interactions between persons and item parcels (σ_{pi}^2)) captured between 47% and 64% of the total variance. The unstable components (i.e., differences between the occasions (σ_o^2), interactions between persons and occasions (σ_{po}^2), and interactions between item parcels and occasions (σ_{io}^2)) captured between 17 and 27% of the variance. The proportion of stable and unstable components was then compared for every achievement goal (see the ratios in Table 3). The stable component of the measure was, for all achievement goals, larger than the unstable component; however, the amount of unstable variance varied to a large degree for the achievement goals investigated.

To understand the reasons for the rather high proportion of occasion-specific variance of learning goals, it is useful to take a closer look at the relative variance components in Table 3. This reveals that the occasion specificity of this goal type was neither due to changes in the whole group of teachers across occasions (σ_o^2) nor due to differences in the item difficulties across occasions (σ_{oi}^2). In fact, nearly all occasion-specific variance was due to an interaction between persons and occasions (σ_{po}^2). This interaction means that persons had different values for their achievement goals on different occasions, leading to a changing sequence of the persons between

Table 3
G analyses.

	Learning		Performance approach		Performance avoidance		Work avoidance	
Variance component (σ^2)	VC	VC%	VC	VC%	VC	VC%	VC	VC%
Stable components	0.16	47%	0.51	64%	0.60	64%	0.43	54%
Person <i>p</i>	0.14	41%	0.50	63%	0.58	62%	0.37	47%
Person \times item parcel <i>pi</i>	0.02	6%	0.01	1%	0.02	2%	0.06	8%
Occasion-specific components	0.09	27%	0.17	21%	0.16	17%	0.16	19%
Occasion <i>o</i>	0 ^a	0%	0.01	1%	0	0%	0.03	3%
Person \times occasion <i>po</i>	0.09	27%	0.16	20%	0.16	17%	0.13	16%
Occasion \times item parcel <i>oi</i>	0	0%	0	0%	0 ^a	0%	0	0%
Item parcel-specific component								
Item parcel <i>i</i>	0	0%	0	0%	0.04	4%	0	0%
Error component								
Person \times occasion \times item parcel <i>poi,e</i>	0.08	26%	0.12	16%	0.14	15%	0.21	26%
Total variance	0.33		0.79		0.93		0.80	
G coefficient $E\rho^2$	0.75		0.88		0.88		0.81	
Ratio of stable to occasion-specific components	2:1		3:1		4:1		3:1	

Note: VC = absolute variance component. VC% = variance component relative to the total variance. *N* = 166 teachers.

^a A small negative variance component was estimated due to sampling errors. Following the suggestion of Brennan (2001), this negative variance was used for the calculation of the remaining variance components and set to zero afterwards.

occasions and/or to a changing relation of the persons' values between occasions.

The analysis of the variance components also showed striking differences regarding the residual variances between different achievement goals (see $\sigma^2_{poi,e}$ in Table 3): The amount of residual variance (i.e., the amount of variance that could not be explained with the variables included) ranged from 15% to 26%, indicating that different teacher goals could be assessed with different degrees of reliability.

3.3. Number of necessary measurement points

In order to determine how many measurement points are required to assess teachers' achievement goals with sufficient reliability (research question 2), we conducted D analyses with one to ten measurement points. We fixed the number of teachers and item parcels to the actual number in the study at hand. Figure 3 illustrates the results for the D analyses for all achievement goals: The figure shows how reliably the different achievement goals could be measured with a given number of measurement points. For example, in order to obtain a G coefficient (i.e., reliability) larger than .70, one measurement point was sufficient for performance approach and avoidance goals. Two measurement points were needed to assess teachers' work avoidance goals with sufficient precision. Finally, three measurement points were required for learning goals to exceed a reliability of $\rho^2 = .70$.

4. Discussion

According to Pintrich (2000), "goals are not traits in the classic personality sense. They are cognitive representations and may show both intraindividual stability as well as contextual sensitivity" (p. 103). In this study, this assumption was tested with regard to teachers' achievement goals.

4.1. Stability of teachers' achievement goals and its explanation

The G analyses in the present investigation showed that the ratio of teachers' stable to unstable achievement goals ranged between 2:1 and 5:1. A closer look at the results revealed that the unstable proportion of variance could be traced back to the main effect of the occasions as well as the interaction between persons and occasions.

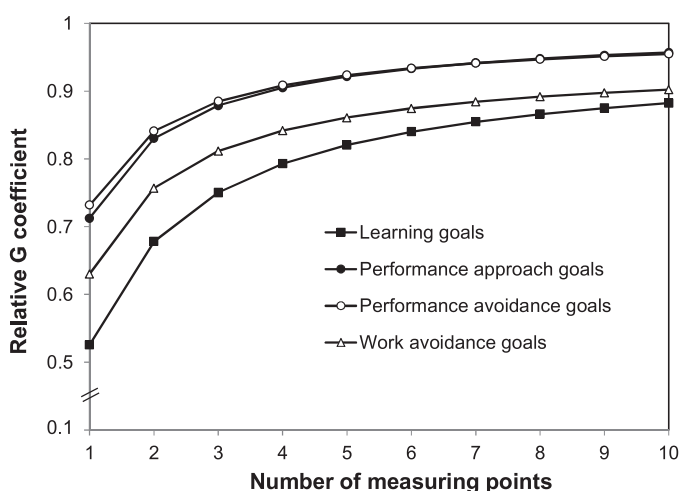


Fig. 3. Decision studies for teachers' achievement goals with 1 to 10 measurement points.

The existing main effect of occasion may be due to different job requirements during the school year. At the beginning of the school year (measurement point 1), teachers have to work through a large number of different tasks (e.g., developing an instruction plan; attending school-year beginning conferences; getting to know new students) and thus aim to avoid additional work. Over the school year the work situation eases, enabling teachers to focus on other goals. However, if we take a look at the size of the main effects of occasions in the G analyses, these main effects are very small in comparison to other existing situation-specific variance components in the data. Mean level differences thus do not seem to be a very important source of variability in teachers' achievement goals.

The interaction between persons and occasions was far more important, accounting for 16 to 27% of the entire variance in the data. Teachers thus had different values for their achievement goals across occasions. This result indicates that measures of teachers' achievement goals contain a large amount of occasion-specific variance – even when these goals are assessed in an occasion-unspecific way. Qualitative investigations would be useful to understand what exactly teachers have in mind when answering achievement goal questionnaires at different measurement points.

How can we know, based on the results of the present study, whether the degree of occasion specificity regarding teachers' achievement goals is large or small? To facilitate the interpretation, one can take a look at the number of measurement points needed to measure teachers' achievement goals. We argue that measurements that require only one measurement point contain a stable proportion that is sufficient enough to characterize the measurement as stable. If, however, more than one measurement point is necessary to capture the characteristic of interest, the characteristic should not be seen as stable. The results of the present study indicate that different kinds of achievement goals differ in the number of measurement points required. Based on our results, one would characterize performance approach and performance avoidance goals as truly stable goals. For work avoidance and learning goals, the trait portions in teachers' achievement goals are not large enough to permit one measurement point to capture them reliably. Although operationalized as dispositions, measures of some achievement goals of teachers thus also measure – at least in their actual operationalization – a considerable part of the specific occasions teachers are in when completing surveys; this is especially true for learning goals (for a similar result for undergraduate students, see Muis & Edwards, 2009). The term "orientation" thus seems only to be appropriate for performance approach and performance avoidance goals, but not for learning and work avoidance goals.

The rather low stability of learning goals is particularly striking. These results are in line with study results found for students' achievement goals (e.g., Schöne, 2008). Several authors have suggested that goals change when the environment changes considerably (e.g., Fryer & Elliot, 2007; Nicholls, 1984). In the present study, teachers were investigated over the course of one school year, teaching the same classes in the same school. Considerable changes in the environment thus cannot be assumed. So why are teachers' learning goals influenced by situational characteristics over short periods of time? One possible explanation is based on methodology: As the learning goal scale showed a high mean at the first measurement point, the instability could be due to central tendencies (Lord, 1963). However, as there were rarely any mean differences between the measurement points (see σ^2_o in the results section), this explanation can be ruled out.

Another possible explanation for the instability of learning goals is related to their definition: Learning goals focus on increasing one's own competencies. The concept of competencies, however, comprises many different aspects. It thus can be assumed that learning goals vary as a function of the specific aspects under

consideration. To operationalize learning goals as validly as possible (see also Groves et al., 2009), it seems useful to distinguish the specific competencies one aims to increase (see Nitsche et al., 2011). However, the more concretely learning goals are assessed, the higher the probability that their stability will recede, as teachers may focus their professional learning on different things at different times (see the aspects “changing characteristics of the individual” and “changing situational/contextual features” in the conceptual framework). For example, a teacher would probably answer the items on the learning goal subscale used in the present study focusing on pedagogic content knowledge differently directly after participating in a vocational training program on didactic aspects than he or she would a few months later when the training content is no longer fresh in his or her mind. Teachers may also adapt their specific learning goals over the course of a school year to the classes they teach: After realizing the specific issues present in the classes he or she is teaching during a specific school year, a teacher may focus on developing certain aspects of his or her competence (e.g., his/her pedagogic knowledge if he/she has several classes with discipline problems).

Referring to the conceptual framework proposed in the present study, there are many more characteristics that could potentially lead to instability in learning goals. Whether the instability is mainly due to changing personal characteristics, to self-regulatory activities, or rather to changing situational/contextual factors is an open question which could be addressed using experimental and/or interview studies in the future.

However, not all achievement goals showed as much variability as learning goals. According to Fryer and Elliot (2007), differences in the stability between different achievement goals are not likely. Why were performance goals so highly stable in the present investigation? One possible explanation is that learning goals refer to specific competencies that vary between occasions (i.e., an interaction between changing characteristics of the individual and changing situational/contextual features in the conceptual framework). Performance goals, in contrast, refer primarily to the individuals or groups the demonstration or avoidance is directed towards (e.g., students or colleagues; see Ziegler, Dresel, & Stoeger, 2008). These significant others rarely vary over the course of one school year for a teacher teaching the same class(es) in the same school (i.e., stable situational/contextual features in the conceptual framework). If we compare the definition and the operationalization of performance goals, there might be a second explanation: Performance approach and avoidance goals focus on demonstrating high competencies or avoiding the impression of low competencies. Thus, there are potentially two issues to be solidified: (a) the significant others to which the demonstration or avoidance is directed (a stable situational/contextual feature) and (b) the specific competencies (analogous to learning goals; a stable or changing characteristic of the person). In the instrument developed by Nitsche et al. (2011), which was used in the present investigation, the significant others are systematically tapped in the item formulations, whereas the competencies are not (e.g., “In my vocation, I aspire to demonstrate to my colleagues that I know more than other teachers”). The reason for not addressing the specific competencies is that investigations with students have shown that learning goals vary between different domains whereas performance goals rarely vary between domains (Nitsche, 2013). However, as this reason is related to the achievement goals of students, it seems to be useful to investigate empirically whether the achievement goals of teachers vary with regard to the specific competencies under consideration. As far as we know, no instrument intending to measure teachers’ achievement goals systematically varies significant others and competencies when measuring performance goals (see e.g., Butler, 2007; Tönjes & Dickhäuser, 2009; Nitsche et al., 2011).

4.2. Implications of the (in-)stability of teachers’ achievement goals

The G analyses indicate that some achievement goals of teachers seem to be influenced by occasional characteristics to a larger degree than commonly assumed. Therefore, researchers should consider the theoretical, methodological, empirical, and practical implications of this instability.

One *theoretical implication* of the results at hand is the need for a deepened understanding of teachers’ achievement goals, how they are generated, and what sorts of occasional characteristics are able to influence them (see also Maehr & Zusho, 2009): “A motivational theory such as goal orientation theory should be able to explicate the core processes that result in the situational construction of a goal orientation and the role of dispositions in this construction” (Kaplan & Maehr, 2007, p. 174). A starting point for investigating these aspects could be the conceptual framework proposed in the present study. To acquire a deepened understanding of teachers’ achievement goals, it is necessary to reconsider their conceptualization as primarily dispositional, as implied by the use of the term “goal orientation” (see also the discussion regarding students’ achievement goals; for an overview, see Maehr & Zusho, 2009). This question has not yet been discussed with regard to teachers’ achievement goals. The present article is a first step in pursuing this question. The results of our study indicate that it might not be appropriate to define and investigate teachers’ achievement goals exclusively as traits. Indeed, our results have revealed that teachers’ achievement goals – even though operationalized as dispositions – are influenced considerably by characteristics of the occasions in which they are assessed. According to Elliot (2005), the partially rather low stability of teachers’ achievement goals is neither surprising nor undesirable; the original aim of the achievement goal approach was to overcome the disadvantages of mere dispositional constructs such as the achievement motive by introducing more contextual information. Elliot claimed that the dispositional focus of researchers regarding achievement goals is thus rather surprising and should be reconsidered.

One important *methodological implication* is that researchers should be clear about the aspects they want to generalize with their measures: When measuring a construct we are often not interested in single performances regarding this construct but in the general construct-related values of persons (Cronbach, Gleser, Nanda, & Rajaratnam, 1972; Lakes & Hoyt, 2008; Shavelson & Webb, 1991). If a measure is largely influenced by occasion characteristics, only statements regarding this specific measurement point can be made. This is unproblematic if researchers are solely interested in characteristics regarding the specific time points at which they are assessed. However, if researchers want to draw conclusions beyond these time points, generalizability across time points has to be taken into account. Therefore, future studies should empirically survey the number of measurement points necessary for a reliable estimate of the construct in question.

An *empirical implication* of the results is that the correlations identified between teachers’ goals (measured at one point in time) and other, stable variables are influenced by the instability of teachers’ achievement goals and are in all likelihood underestimations of the true correlations. This is especially true for learning goals. Additionally, teachers’ achievement goals are differently influenced by occasion characteristics. These differences lead to variation regarding the generalizability of the goals. Comparisons between achievement goals regarding their impact on other variables (e.g., instructional characteristics) are thus not admissible when achievement goals are only measured once.

A *practical implication* of the partially rather low stability of achievement goals is that teachers’ achievement goals are, in principle, modifiable. This is good news as it implies that training programs focusing on achievement goals can be successful (Salas

& Cannon-Bowers, 2001). This is again especially true for learning goals.

4.3. Limitations and further directions

The generalizability of the results across the investigated sample of teachers and the instruments used are crucial for the implications of these results for research on teachers' achievement goals (see also Kaplan & Maehr, 2007). As is the case with many studies on teachers, the present sample was not representative. However, the descriptive results (internal consistencies, means, and standard deviations) were quite similar to those reported by Nitsche et al. (2011). It nevertheless remains unclear as to whether the result patterns of the study at hand apply for the entire teacher population beyond German teachers. Regarding the generalizability across instruments, it has already been mentioned that instruments aiming to assess achievement goals differ considerably. The (in-)stability found in the study at hand therefore might, to a certain degree, be instrument dependent. However, if we take a look at the retest correlations of previous studies, learning goals were, on a descriptive level, less stable when compared to performance goals as well.

Another critical point regarding the results is that it remains unclear as to whether the results can be generalized beyond the three-months retest interval chosen for the present investigation. One can hypothesize that longer retest intervals could lead to lower stability. It can thus be assumed that the time variability found in the present study is a minimum estimation of change occurring over several school years and/or more strongly differing contexts. An indication that this may be true can be derived from the retest correlations of former studies investigating teachers' and teacher trainees' achievement goals: In most of these studies the retest intervals were larger than the one in the present investigation. The retest correlations in these studies were, on a descriptive level, smaller than those found here (see Fasching et al., 2010; Malmberg, 2008; Tönjes & Dickhäuser, 2009). However, further studies directly investigating the effects of retest intervals in the context of teacher achievement goals are necessary. All in all, it seems important to conduct further studies on the stability of teachers' achievement goals to determine whether the results of the present study can be replicated using another sample as well as other instruments.

A third point that can be criticized is the method used for analyzing the data. G theory enables a separation between trait and state variance as well as a determination of the number of measurement points necessary for a reliable assessment of teachers' achievement goals. Applying G theory is therefore very useful for analyzing stability questions. Nevertheless, the variance components provided are subject to sampling variability (Brennan, 2001). To verify the credibility of the results, standard errors and/or confidence intervals of the variance components would be a useful measure (see Hoyt & Melby, 1999). Computing them is, however, not straightforward as this would require distributional assumptions (e.g., normality assumptions) which cannot be assumed in designs such as the one used. A replication of the results at hand is thus also important for this reason. Here this is even more important as teachers are nested within schools in the data. This dependency could, however, not be accounted for in the G analyses as the design would have been too complex and the variance component estimates less trustworthy. Not taking into account the nested structure can have an additional impact on the standard errors. However, as previous research has shown that school effects are rarely relevant for teachers' achievement goals (see e.g., Dickhäuser, Nitsche, Fasching, & Dresel, 2012), this impact can be assumed to be rather small. With regard to the interpretation of the analyses conducted, two additional points can be criticized: First, with the chosen method it is not possible to separate variability and developmental change. Second, time stability and trans-situational

stability may be confounded as situations were not systematically varied or held constant.

In addition to questions regarding the generalizability of the results and the method used for data analysis, further questions remain for future research. One of them is related to factors underlying potential differences between teachers in their stability of goals. One could, for example, hypothesize that more experienced teachers have more stable goals as they may have identified long-term learning goals which do not change within short time periods.

5. Conclusions

Teachers' achievement goals are often assumed to be stable characteristics. The study at hand provides an indication that this is not necessarily the case. Learning goals in particular seem to be considerably influenced by time-variable characteristics. The results show that stability questions are largely relevant for the interpretation of teachers' achievement goals and that theoretical assumptions regarding stability do not suffice. In the short run, it is necessary to include a sufficient number of measurement points in a study in order to capture the constructs in question. In the long run, the results underline the importance of advancing research on teachers' achievement goals both theoretically and methodologically.

Acknowledgments

This research was funded by a grant from the German Federal Ministry of Education and Research to Oliver Dickhäuser (01 HJ 0901) and Markus Dresel (01 HJ 0902).

References

- Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, *84*, 888–918.
- Boekaerts, M., & Niemimirta, M. (2000). Self-regulation in learning: Finding a balance between learning- and ego-protective goals. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 417–450). San Diego, CA: Academic Press.
- Brennan, R. L. (2001). *Manual for urGENOVA*. Iowa City, IA: Iowa Testing Programs, University of Iowa.
- Butler, R. (2007). Teachers' achievement goals and associations with teachers' help seeking: Examination of a novel approach to teacher motivation. *Journal of Educational Psychology*, *99*, 241–252.
- Butler, R. (2012). Striving to connect: Extending an achievement goal approach to teacher motivation to include relational goals for teaching. *Journal of Educational Psychology*, *104*, 726–742. doi:10.1037/a0028613.
- Butler, R., & Shihab, L. (2008). Achievement goals for teaching as predictors of students' perceptions of instructional practices and students' help seeking and cheating. *Learning and Instruction*, *18*, 453–467. doi:10.1016/j.learninstruc.2008.06.004.
- Button, S., Mathieu, J., & Zajac, D. (1996). Goal orientation in organizational research: A conceptual and empirical foundation. *Organizational Behavior and Human Decision Processes*, *67*, 26–48.
- Crick, J. E., & Brennan, R. L. (1983). *Manual for GENOVA: A generalized analyses of variance system [Computer software and manual]*. University of Iowa.
- Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements*. New York: John Wiley.
- DeShon, R. P., & Gillespie, J. Z. (2005). A motivated action theory account of goal orientation. *Journal of Applied Psychology*, *90*, 1096–1127. doi:10.1037/0021-9010.90.6.1096.
- Dickhäuser, O., Nitsche, S., Fasching, M. S., & Dresel, M. (2012). Kommt es auf die Schule an? Zu Bedeutung von wahrgenommenen Kontextmerkmalen für die beruflichen Zielorientierungen von Lehrkräften [Does the school matter? The relevance of perceived context characteristics for teachers' goal orientations]. In A. Philipp & M. Kunter (Chairs), *Macht die Schule den Unterschied? Zur Rolle des Schulkontexts für das Verhalten und Erleben von Lehrkräften*. Symposium conducted at the meeting of the German Psychology Society, Bielefeld, Germany.
- Dresel, M., Fasching, M. S., Steuer, G., Nitsche, S., & Dickhäuser, O. (2013). Relations between teachers' goal orientations, their instructional practices and student motivation. *Psychology (Savannah, Ga.)*, *7*, 572–584. doi:10.4236/psych.2013.47083.
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, *95*, 256–273.
- Elliot, A. J. (2005). A conceptual history of the achievement goal construct. In A. Elliot & C. Dweck (Eds.), *Handbook of competence and motivation* (pp. 52–72). New York: Guilford Press.

- Elliott, E. S., & Dweck, C. S. (1988). Goals: An approach to motivation and achievement. *Journal of Personality and Social Psychology, 54*, 5–12.
- Elliot, A. J., & Murayama, K. (2008). On the measurement of achievement Goals: critique, Illustration, and application. *Journal of Educational Psychology, 100*, 613–628.
- Fasching, M. S., Dresel, M., Dickhäuser, O., & Nitsche, S. (2010). Achievement goals of teacher trainees: Longitudinal analysis of magnitude, change and relevance. *Journal of Educational Research Online, 2*, 9–33.
- Fryer, J. W., & Elliot, A. J. (2007). Stability and change in achievement goals. *Journal of Educational Psychology, 99*, 700–714.
- Groves, R. M., Fowler, F. J., Jr., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2009). *Survey methodology* (2nd ed.). Hoboken, NJ: John Wiley & Sons.
- Harter, S., & Jackson, B. K. (1992). Trait vs. nontrait conceptualizations of intrinsic/extrinsic motivational orientation. *Motivation and Emotion, 16*, 209–230.
- Heckhausen, J., & Heckhausen, H. (2010). Motivation and action: introduction and overview. In J. Heckhausen & H. Heckhausen (Eds.), *Motivation and action* (2nd edition, pp. 1–9). Cambridge: Cambridge University Press.
- Hoyt, W. T., & Melby, J. N. (1999). Dependability of measurement in counseling: An introduction to generalizability theory. *The Counseling Psychologist, 27*, 325–352.
- Kaplan, A., & Maehr, M. L. (2007). The contributions and prospects of goal orientation theory. *Educational Psychology Review, 19*, 141–184. doi:10.1007/s10648-006-9012-5.
- Kishton, J. M., & Widaman, K. F. (1994). Unidimensional versus domain representative parceling of questionnaire items: An empirical example. *Educational and Psychological Measurement, 54*, 757–765.
- Lakes, K. D., & Hoyt, W. T. (2008). What sources contribute to variance in observer ratings? Using generalizability theory to assess construct validity of psychological measures. *Infant and Child Development, 17*, 269–284.
- Little, T. D., Cunningham, W. A., Shahar, G., & Widaman, K. F. (2002). To parcel or not to parcel: Exploring the question, weighing the merits. *Structural Equation Modeling, 9*, 151–173.
- Lord, F. M. (1963). Elementary models for measuring change. In C. W. Harris (Ed.), *Problems in measuring change*. Madison: University of Wisconsin Press.
- Maehr, M. L., & Zusho, A. (2009). Achievement goal theory. The past, present, and future. In K. R. Wentzel & A. Wigfield (Eds.), *Handbook of motivation at school* (pp. 77–104). New York: Routledge.
- Malmberg, L.-E. (2008). Student teachers' achievement goals during teacher studies: Antecedents, correlates and outcomes. *Learning and Instruction, 18*, 438–452.
- Muis, K. R., & Edwards, O. (2009). Examining the stability of achievement goal orientation. *Contemporary Educational Psychology, 34*, 265–277. doi:10.1016/j.cedpsych.2009.06.003.
- Murayama, K., Elliot, A. J., & Friedman, R. (2012). Achievement goals and approach-avoidance motivation. In R. M. Ryan (Ed.), *The Oxford handbook of human motivation* (pp. 191–207). Oxford: Oxford University Press.
- Murphy, P. K., & Alexander, P. A. (2000). A motivated exploration of motivation terminology. *Contemporary Educational Psychology, 25*, 3–53. doi:10.1006/ceps.1999.1019.
- Nicholls, J. G. (1984). Achievement motivation: Conceptions of ability, subjective experience, task choice, and performance. *Psychological Review, 91*, 328–346.
- Nitsche, S. (2013). *Zielorientierungen von Lehrkräften und ihre Bedeutung für das berufliche Lern- und Fortbildungsverhalten*. [Goal orientations among teachers and their significance for occupational training choices]. (Unpublished doctoral dissertation.) University of Mannheim, Germany.
- Nitsche, S., Dickhäuser, O., Fasching, M. S., & Dresel, M. (2011). Rethinking teachers' goal orientations: Conceptual and methodological enhancements. *Learning and Instruction, 21*, 574–586.
- Nitsche, S., Dickhäuser, O., Fasching, M. S., & Dresel, M. (2013). Teachers' professional goal orientations: Importance for further training and sick leave. *Learning and Individual Differences, 23*, 272–278. doi:10.1016/j.lindif.2012.07.017.
- Pintrich, P. R. (2000). Multiple goals, multiple pathways: The role of goal orientation in learning and achievement. *Journal of Educational Psychology, 92*, 544–555.
- Retelsdorf, J., Butler, R., Streblo, L., & Schiefele, U. (2010). Teachers' goal orientations for teaching: Associations with instructional practices, interest in teaching, and burnout. *Learning and Instruction, 20*, 30–46. doi:10.1016/j.learninstruc.2009.01.001.
- Retelsdorf, J., & Günther, C. (2011). Achievement goals for teaching and teachers' reference norms: Relations with instructional practices. *Teaching and Teacher Education, 27*, 1111–1119. doi:10.1016/j.tate.2011.05.007.
- Salas, E., & Cannon-Bowers, J. A. (2001). The science of training: A decade of progress. *Annual Review, 52*, 471–499. doi:0066-4308/01/0201-0471\$14.00.
- Schöne, C. (2008). *Zielorientierung und Bezugsnormpräferenzen in Lern- und Leistungssituationen*. [Goal orientation and reference norm preferences in learning and achievement situations]. (Unpublished dissertation.) University of Gießen, Germany.
- Senko, C., Hulleman, C. S., & Harackiewicz, J. M. (2011). Achievement goal theory at the crossroads: Old controversies, current challenges, and new directions. *Educational Psychologist, 46*, 26–47.
- Shavelson, R. J., & Webb, N. M. (1991). *Generalizability theory: A primer*. Newbury Park, CA: Sage.
- Shumate, S. R., Surlis, J., Johnson, R. L., & Penny, J. (2007). The effects of the number of scale points and non-normality on the generalizability coefficient: A Monte Carlo study. *Applied Measurement in Education, 20*, 357–376.
- Silva, T., & Nicholls, J. (1993). College students as writing theorists: Goals and beliefs about the causes of success. *Contemporary Educational Psychology, 18*, 281–293.
- Steyer, R., Schmitt, M., & Eid, M. (1999). Latent state-trait theory and research in personality and individual differences. *European Journal of Personality, 13*, 389–408.
- Tönjes, B., & Dickhäuser, O. (2009). Längsschnittliche Effekte von Zielorientierungen auf Dimensionen des beruflichen Belastungserlebens im Lehrerberuf. [Longitudinal effects of achievement goal on factors of occupational burden in the teaching profession]. *Zeitschrift für Pädagogische Psychologie, 41*, 79–86.
- Vollmeyer, R., & Rheinberg, F. (1998). Motivationale Einflüsse auf Erwerb und Anwendung von Wissen in einem computersimulierten System. [Motivational influences on the acquisition and application of knowledge in a computer-simulated system]. *Zeitschrift für Pädagogische Psychologie, 12*, 11–24.
- Webb, N. M., Shavelson, R. J., & Haertel, E. H. (2006). Reliability coefficients and generalizability theory. In C. Rao & S. Sinharay (Eds.), *Handbook of statistics* (Vol. 26, pp. 81–124). Amsterdam, The Netherlands: Elsevier.
- Ziegler, A., Dresel, M., & Stoeger, H. (2008). Addressees of performance goals. *Journal of Educational Psychology, 100*, 643–654. doi:10.1037/0022-0663.100.3.643.