

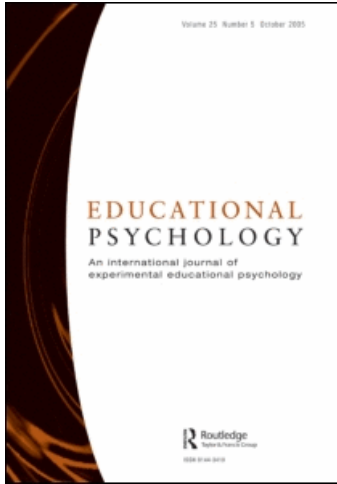
This article was downloaded by: [Krampen, Günter]

On: 19 July 2010

Access details: Access Details: [subscription number 924496685]

Publisher Routledge

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## Educational Psychology

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713415498>

### Improvement of orthography test performance by relaxation exercises: results of a controlled field experiment in basic secondary education

Günter Krampen<sup>a</sup>

<sup>a</sup> Department of Psychology, University of Trier, Trier, Germany

First published on: 11 June 2010

**To cite this Article** Krampen, Günter(2010) 'Improvement of orthography test performance by relaxation exercises: results of a controlled field experiment in basic secondary education', *Educational Psychology*, 30: 5, 533 – 546, First published on: 11 June 2010 (iFirst)

**To link to this Article: DOI:** 10.1080/01443410.2010.486474

**URL:** <http://dx.doi.org/10.1080/01443410.2010.486474>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

## **Improvement of orthography test performance by relaxation exercises: results of a controlled field experiment in basic secondary education**

Günter Krampen\*

*Department of Psychology, University of Trier, Trier, Germany*

*(Received 11 January 2010; final version received 15 April 2010)*

The effects of relaxation exercises on orthography performance in language arts education of fifth to seventh graders were experimentally tested. Participants were 399 basic secondary school students and their language arts teachers from the Hauptschule, a German type of secondary education covering grades five to nine that leads to a basic educational degree. Half of the students were trained in their classrooms in the basic autogenic training (AT) formulas (self-suggestions of passive self-attention, heaviness and warmth) and in the technique of reactivation after relaxation. After completing a regular dictation test students applied the relaxation exercise for four minutes on their own and were then given the opportunity to reread what they had written and to correct any mistakes they found on the test. In comparison to 200 students in the control group (who were not trained in the systematic relaxation exercises), who were also given the opportunity to revise their texts after a four-minute break, the experimental group students made significantly fewer mistakes and received better grades on the test. In addition, the results show that there were significantly fewer correct revisions and significantly more incorrect revisions of the dictation tests made by the students in the control group, thus worsening their academic performance by the possibility of 'self-correction'. Students and teachers evaluated AT exercises positively. Implications of the results for the application of AT in the classroom are discussed.

**Keywords:** academic achievement; orthography; performance tests; relaxation exercises; language arts education; autogenic training

### **Introduction**

Writing skills and basic orthography performance are crucial for success in school and the workplace. A meta-analysis of 123 experimental and quasi-experimental studies conducted by Graham and Perin (2007) identified medium to low average weighted effect sizes for 11 different instructional practices for teaching writing to fourth to twelfth graders. Dependent variables of such evaluation studies on teaching methods are performance tests on orthography and quality of essay writing. However, performance tests may underestimate the writing skills and competence of students because of underachievement influenced by trait factors like test anxiety and low achievement motivation as well as by state factors like performance anxiety, arousal and nervousness, and task irrelevant cognitions.

---

\*Email: [krampen@uni-trier.de](mailto:krampen@uni-trier.de)

The latter factors, i.e. the state factors, are hypothesised to be of special relevance in *low* achievers having had repeated experiences of academic failure resulting in a low self-concept of writing ability and literacy as well as in negative self-attributions – both of which negatively influence achievement motivation and test anxiety as well as academic and career aspiration. Longitudinal results show that subjective ability beliefs about reading performance in the fourth grade and about English in language arts education in the 10th grade are most predictive for the number of language arts courses per year of high school and for language-related career aspirations of high school graduates (see Durik, Vida, & Eccles, 2006).

Writing is a very demanding task, requiring the synthesis of a variety of cognitive resources. Nagy, Berninger, Abbott, Vaughan, and Vermeulen's (2003) results show that fourth graders are still learning to coordinate orthographic, phonological and morphological cues in writing. It can be hypothesised that older students as well – especially low achievers – have not yet mastered important writing processes, orthographical skills and knowledge involved in writing beyond the level of essay writing (see, e.g. De La Paz & Graham, 2002) that includes orthography tests as well.

Performance tests such as dictation tests (i.e. tests in which teachers read a paragraph of information aloud sentence by sentence without any spelling and students are required to write each sentence on their papers) imply – in contrast to essay writing – no possibilities for students to deal with difficult words and sentences strategically. Moreover, feedback on orthography and punctuation mistakes as well as the grades that are assigned are very objective. Therefore, from the student's perspective they are very impressive and strongly affect the self-concept of ability and negative self-attributions. However, researchers suppose that test performance is not only influenced by trait factors like test anxiety (see e.g. Chapell et al., 2005) and achievement motivation, but by state factors like performance anxiety, arousal and nervousness, and task irrelevant cognitions as well, thus impairing particularly the working memory processes involved in orthography and punctuation (see e.g. McCutchen, 1996). Therefore, it is interesting to note that learning goals – at a more general 'trait' level – are negatively related to test anxiety, but performance goals – at a more situational 'state' level – for a specific examination are positively related to test anxiety – at least in college students after mid-term and final examinations in introductory statistic courses (see Bandalos, Finney, & Geske, 2003).

Hence, an *important research question refers to the possibilities to reduce performance deterioration state factors like performance anxiety, arousal and nervousness, and task irrelevant cognitions* in the short run, and, perhaps, trait factors like a negative self-concept of writing ability and literacy, negative self-attributions, test anxiety and low achievement motivation in the long run. Of course, test-taking conditions are crucial and should be favourable, both motivationally and emotionally. However, orthography tests – in schools, job applicant screening, employment tests, personnel selection and on the job as well – are speed tests, causing stress as well as impaired stress and anxiety management. Otherwise, students as well as job applicants and employees often have the opportunity to revise texts they have written on their own, that is, they have the chance to reread what they have written and to check it over for mistakes in orthography and punctuation. These orthography revisions may be either correct and performance enhancing or incorrect and performance worsening. Supposing an individual has an (more or less) adequate orthographic knowledge and competence, we are in need of test-taking conditions that are favourable for the recall

and manifestation of these competences in orthography and punctuation during the stressful performance test situation.

Some types of relaxation exercises have been empirically proven to have a performance-enhancing effect. There are especially positive experimental results for the impact of autogenic training (AT) and similar self-relaxation techniques on convergent and divergent productions (like retention, recall, memory performance, short-term memory, concentration and selective attention, word fluency, music performance quality, typing speed and accuracy) in children, adolescents and adults (see e.g. Budohaska & Orłowska, 1970; Krampen, 1997; Matthews & Quinn, 1987; Siersch, 1989; Sweeney-Burton, 1998). This is explained by the special characteristics of AT, which refer to trained short-term skills to relax and to rest. The autogenic state during AT exercise is a state of autonomic peripheral relaxation (autonomic nervous system), while mental activity (central nervous system) is regular, fixed at the autogenic formulas and slowed down. Mental activity is neither (in comparison with sleep) interrupted nor (in comparison with dozing) associative, fleeting and cursory. Moreover, experimental results on such autonomic parasympathetic effects of AT are well established as is the corresponding impact of AT on electroencephalography, that is, alpha rhythms with very synchronic waves of slow amplitudes and a very regular frequency of cycles (see e.g. Krampen, 1998; Luthe, 1970).

However, the majority of results on the performance-enhancing effects of AT is not only quite old, but it is as well based only on experimental studies with small sample sizes and with limited external validity because neither the treatments nor the performance tests are integrated parts of everyday school education. Therefore, in the following, the effects of basic autogenic exercises on curriculum-based, regular orthography tests in language arts education are analysed in a controlled field experiment with randomisation in a large sample of fifth to seventh graders. One-tailed hypotheses refer to the improvement of examination performance in these students (who were instructed in the techniques of the AT relaxation exercises in their classrooms) after applying the self-relaxation technique for four minutes after completion of the dictation test and before becoming the opportunity to revise their texts. Thus, the research question under study is how low-achieving students improve their writing performance by use of this self-regulation technique applied after writing the text and before rereading and self-corrections. Improvement of performance is evaluated with reference to the actual grades received on the dictation test as well as to the number of correct and incorrect self-corrections made by students during their revision of the text.

## Method

### *Participants*

Participants were an unselected sample of 399 fifth, sixth and seventh graders enrolled in 20 different classes from 10 different *Hauptschulen*, a German type of secondary education covering Grades 5–9 that leads to a basic educational degree. It should be noted that all of these students are low achievers. This is a result of the traditional German ability grouping and educational placement after fourth grade into three types of secondary education: the *Hauptschule*, covering Grades 5–9 and leading to a basic educational degree; the *Realschule*, covering Grades 5–10 and leading to a middle-level educational degree; and the *Gymnasium*, covering Grades 5–12 or 13 (depending on the state school system) and leading to the highest educational degree which

qualifies students for college and university admission. At this point, it should be noted that in the recent past this traditional educational placement system has been the topic of much political and public debate, and various new models of school organisation and administration are currently being developed and implemented in Germany.

De facto the present sample is – in comparison with secondary students attending either the Realschule or Gymnasium – low achieving in all school subjects. This is especially true for German language performance, which is a prerequisite of performance in most other school subjects. There were 219 female and 180 male students who were enrolled in six fifth-grade classes, eight sixth-grade classes and six seventh-grade classes (age:  $M = 11.4$ ,  $SD = 4.32$ , range: 10–13 years). Socioeconomic status of the students is with reference to parents' education and occupational status lower class ( $n = 177$ ; 44%) and lower middle class ( $n = 222$ ; 56%). Ethnicity of most students is German ( $n = 374$ ; 94%), the other students' ethnicities are Italian ( $n = 13$ ; 3%), Turkish ( $n = 11$ ; 3%) and Spanish ( $n = 1$ ; 0.3%) being members of the second migration generation with birth in Southern Germany. There are significant differences between the experimental and the control group neither in age [ $t(398) = 0.73$ ;  $p > .20$ ] nor in socioeconomic status, gender and ethnicity (Chi-square  $< 1.26$ ;  $p > .20$ ).

A total of 10 language arts education teachers (four females, six males) of 10 different schools were involved in the study, with each of these teaching language arts (German) to two parallel courses of the same grade level. Teachers' job experience ranges between 12 and 16 years.

### **Procedure**

A randomised design with a treatment and a control group of students was employed. Beforehand, 10 language arts teachers (self-selection in the recruitment of teachers was achieved by their response to notices that were placed in 12 *Hauptschulen* on teacher involvement in an empirical study on language arts education) completed an introductory course in AT (one group session weekly for 10 weeks) using standard procedures and formulas (see e.g. Schultz & Luthe, 1969). After practicing the 'simple sitting posture' (this was preferred to the horizontal training posture and the reclining chair posture because it is more practicable in everyday life), closing eyes, and passive concentration (implying a casual and functional passivity towards the intended functional changes), and the technique of coming back to normal (psychophysiological reactivation by flexing arms vigorously, breathing deeply and opening eyes), the standard formulas of AT were introduced and rehearsed. One AT formula was introduced in each group session after an introductory discussion of exercise experiences with the participants. The standard mental exercises involve autosuggestions of the standard formulas:

- (1) 'I am at peace – peace', aiming at inner distancing from the current situation and preparing for autogenic concentration;
- (2) 'My right/left arm is heavy – heaviness' (the dominant arm was selected), referring to muscle relaxation in the arms, hands and legs;
- (3) 'My right/left arm is warm – warmth' (the dominant arm was selected), referring to vascular dilation in the arms, hands and legs;
- (4) 'Breathing calm and regular – it breathes me', referring to passive registration of calm, abdominal autonomic breathing;

- (5) 'Heartbeat calm and regular', referring to passive registration of autonomic heartbeat and pulse;
- (6) 'My solar plexus is warm', referring to parasympathetic activity of the solar plexus and warmth in the abdomen; and
- (7) 'My forehead is cool', referring to reactivating and performance.

Participants were instructed to practice passive concentration 'in mental contact with the part of the body indicated by the formula (e.g. the right arm), and maintenance of a steady flow of a film like (verbal, acoustic or visual) representation of the autogenic formula in mind' (Schultz & Luthe, 1969, p. 15). Thus, from a psychophysiological point of view, the state is set for relaxation and self-regulation during autogenic exercises by the reduction of extero- and pro-prioceptive stimulation, and by the verbal content of the formula, which implies that the relevant psychophysiological systems work automatically. Participants practiced the autogenic exercises alone at least twice a day. Training focuses on autogenic exercises AT at home. This was controlled by systematic self-control sheets, which were discussed and reflected in each group session. Compliance of teachers to AT-exercises at home was high. The exercises were scheduled – according to the adaptive indication strategies of relaxation exercises (Krampen & von Eye, 2006) – by taking into account the participants' routines in everyday life and specifying activities after exercises, with rather brief (two to five minutes) autogenic exercises at the beginning. All teachers successfully completed the introductory course in AT and were prepared and still willing to participate in the empirical study, which is described in this text.

The teachers were given special instructions on how to integrate an abbreviated form of basic AT in one of their randomised selected parallel language arts classes. The short form of basic AT consists of (1) a more student- and classroom-adequate relaxed sitting posture with students placing their heads on their arms which are relaxed on the student's desk, (2) closed eyes and passive concentration (implying a casual and functional passivity towards the intended functional changes), (3) the standard formulas of self-suggestions of peace, heaviness and warmth, and (4) the technique of coming back to normal (psychophysiological reactivation by flexing arms vigorously, breathing deeply, and opening eyes as well as standing up and doing some light physical exercises). These short (four to five minutes) autogenic exercises were integrated into the regular language arts education (e.g. after intensive class work, at the end of a teaching unit, but not for student discipline) in the 10 experimental classes ( $n = 199$  students enrolled in three fifth-grade classes, four sixth-grade classes and three seventh-grade classes) and the 10 parallel control classes ( $n = 200$  students enrolled in three fifth-grade classes, four sixth-grade classes and three seventh-grade classes). Each teacher taught language arts to one class of students in the experimental group and one class of students in the control group.

After learning the short autogenic exercises, students in the 10 experimental classes were instructed to apply the AT exercise for a period of four minutes after completing the regular (second) dictation test in the classroom. Self-relaxation of students started after closing the test booklet and stopped with the physical reactivation. After the AT session, students were given the opportunity to reread their text and to revise mistakes in orthography and punctuation with a red pencil. Red pencils were necessary for the *post-hoc* identification of the revisions students made after relaxation and rereading. Five minutes were provided for text revision.

Students in the 10 parallel control classes completed the same dictation test on the same day by the same teacher. After taking the test, students were also instructed to close their test booklet and to rest for four minutes (unspecific resting instruction). Following this and in the same manner as the experimental classes, students in the control classes were given the opportunity to reread their text and to check over for mistakes in orthography and punctuation with a red pencil. Red pencils were necessary for the *post-hoc* identification of the revisions students made after unspecific resting and rereading. Again, text revision was possible for five minutes.

Teachers' dictation procedures were observed in the classrooms systematically with reference to time duration and not spelling by two independent observers each. In addition, classroom observers checked teachers' instructions for relaxation exercise (unspecific resting, respectively), relaxation time (time for unspecific resting, respectively), and time given for rereading and self-corrections for treatment fidelity.

The original study planned to extend this procedure to another dictation test in both the experimental and the control classes. However, the first results motivated all teachers to instruct all of their students – that is, the students in the 10 control classes, too – in the short autogenic exercises. Therefore, all students in the 20 classes were instructed to apply such an exercise for a period of four minutes after completing the next regular (third) dictation test. Then, all students were given the opportunity to reread their texts and to correct mistakes for five minutes. Thus, in the experimental groups, the third dictation test is a short-term maintenance probe administered one month after the second dictation test, in the former control groups, the third dictation test with self-relaxation after completing the test and before rereading is novel and experimental (albeit without controls).

## **Measures**

### *Independent test scoring*

To control for expectancy biases of the teachers involved in the study not only were the assigned grades gathered; in addition, the number of mistakes on all dictation tests as well as the number of right and of wrong 'self-revisions' undertaken by the students after rereading (made by the students with the red pencil) were gathered by independent test scorers. These independent scorers were blind to the tests of the experimental and the control classes. It is worth noting that the scores of these independent scorers are in very high agreement with the scores of the teacher themselves ( $r > .94$ ;  $p < .01$ ), thus not indicating any expectancy biases in the teachers who participated in the study. It is important to note that all dictation tests used in these analyses were regular curriculum-based assessments with direct significance for the school reports of the students.

### *Indicators of test performance*

For the control of the randomisation procedure applied at class-level baseline, information on language arts performance of all students was gathered by considering the grades students had on their previous report cards as well as their grades and number of mistakes they made on the first dictation test in the new school year. Grades and number of mistakes as well as the number of right and of wrong 'self-corrections'

students had after rereading were gathered for the second and the third dictation test. In addition, data on the language arts grades students achieved on their next report card were collected. The German educational grading system is based on numbers indicating: 1 = 'excellent performance' (A), 2 = 'good' (B), 3 = 'satisfactory' (C), 4 = 'fair' (D), 5 = 'unsatisfactory' and 6 = 'fail' (F). Both of the grades 'unsatisfactory' and 'fail' (F) indicate school failure, which – on school report cards – leads to students repeating a grade or to no educational degree and qualification at school leaving, respectively. Of course, number of mistakes and grades are significantly correlated for each dictation test ( $r > .82$ ;  $p < .01$ ).

### *Evaluation of relaxation exercise by students*

At the end of the first semester, students were asked to evaluate the self-relaxation technique: The first item referred to a general evaluation ['In my opinion, the relaxation technique is ... stupid' (= 1) vs. '... great' (= 7)], the second to the frequency of applications of the self-relaxation technique outside school ['Outside school I use the relaxation technique ... daily' (= 5), '... several times a week' (= 4), 'several times a month' (= 3), '... less frequently' (= 2), '... never at all' (= 1)].

## **Results**

### ***Baseline data and randomisation check***

Baseline data on average grades on the last school report cards in language arts education as well as the grades and number of mistakes on the first dictation test of the new school year confirm that randomisation led successfully to comparable experimental and control groups (see upper part of Table 1): Average grade is a 'C-' (i.e. somewhat less than satisfactorily performance) in both groups; average number of mistakes in the curriculum-based dictation tests (with a length of approximately 1–1.5 handwritten pages) is comparably high in both groups as well. Standard deviations (SD; see Table 1) show that all students made very much mistakes in the dictation tests.

### ***Check of treatment fidelity***

The two independent classroom observers agreed in their observations (1) of not spelling during teachers' dictations ( $\kappa = .79$ ;  $p < .01$ ), (2) dictation time duration ( $r = .98$ ;  $p < .01$ ), (3) teachers' correct instructions for relaxation exercise (unspecific resting, respectively) ( $\kappa = .81$ ;  $p < .01$ ), (4) relaxation time (time for unspecific resting, respectively) given by the teachers ( $r = .96$ ;  $p < .01$ ), and (5) time given for rereading and self-corrections for treatment fidelity ( $r = .95$ ;  $p < .01$ ). Thus, treatment fidelity of teachers' instructions was high. In addition, there were no significant differences between the teachers of the experimental and the control group: (1) spelling during dictation did not occur at all, (2) dictation time between the groups does not differ significantly [ $t(19) = 0.30$ ;  $p > .20$ ], (3) all teachers instructed correctly for AT-exercise or unspecific resting of students, respectively, (4) relaxation time and unspecific resting time do not differ significantly between the groups [ $t(19) = 0.26$ ;  $p > .20$ ], and (5) there are no significant group differences in time given by the teachers for students' rereading and self-corrections [ $t(19) = .18$ ;  $p > .20$ ].



Table 1. Academic test performance in language arts education (orthography) of fifth to seventh graders with versus without relaxation exercises after taking a dictation test and before having the opportunity to correct their mistakes.

Dependent variable	Group A (N = 199)		Group B (N = 200)		Bartlett-Box test	t-test (df = 397)	Effect size <i>d</i>
	M	SD	M	SD			
July (at the end of the last school year): without relaxation exercises							
Grade on last school report card <sup>a</sup>	3.4	1.86	3.3	1.90	0.21	0.78	0.05
October: first dictation test without relaxation exercises in Groups A and B							
Grade on first test <sup>a</sup>	3.1	1.11	3.2	1.08	0.05	0.49	0.09
Number of mistakes in first test	12.7	2.78	13.0	2.91	0.30	0.81	0.01
November: second dictation test in Group A with and in Group B without relaxation exercises							
Grade on second test <sup>a</sup>	2.6	0.73	3.1	0.97	0.77	1.98**	0.58
Number of mistakes in second test	7.9	2.01	11.7	2.53	1.02	3.91**	1.46
Number of correct revisions	5.9	1.38	2.0	0.99	0.77	3.13**	2.79
Number of incorrect revisions	1.2	0.69	4.3	0.81	0.38	2.88**	2.58
December: third dictation test in Group A and in Group B with relaxation exercises							
Grade on third test <sup>a</sup>	2.9	0.88	3.0	0.79	0.21	0.37	0.11
Number of mistakes in third test	8.2	1.97	9.1	2.05	0.66	1.04	0.42
Number of correct revisions	5.7	1.05	3.6	1.07	0.03	1.83*	1.91
Number of incorrect revisions	1.8	0.90	2.1	1.00	0.49	0.81	0.27
February: mid-term school report							
Grade on next school report card <sup>a</sup>	3.1	1.77	3.0	1.68	0.33	0.65	0.06

\*\**p* < .01; \**p* < .05.

Notes: <sup>a</sup>The German educational grading is numerically based with grades ranging from 1 = 'excellent performance' (A), 2 = 'good' (B), 3 = 'satisfactory' (C), 4 = 'fair' (D), to 5 = 'unsatisfactory', and 6 = 'fail' (F).

### *Effects of AT-exercises in second dictation test*

In the second dictation test of the school year, Group A (self-relaxation after completing the dictation test and before rereading their texts) in comparison with Group B (controls) attained on average significantly better grades and made significantly fewer mistakes (see Table 1). Average grade assigned is a 'B-' (somewhat less than good performance) and the average number of mistakes is approximately eight in the experimental group, whereas the controls were assigned on average a 'C' and made 12 mistakes on average. In terms of Cohen's (1988) guidelines, effect size of these group differences is medium for the grades and large for the number of mistakes.

Even more impressive are the results on the numbers of right and of wrong 'corrections' that students made while rereading and text revision. Students in the experimental Group A performed significantly more right and significantly less wrong 'corrections' of mistakes than students in the control Group B. In terms of Cohen's (1988) guidelines, effect sizes of these group differences are large (see Table 1). It is worth noting that the controls – on average – performed more wrong than right corrections in rereading, thus worsening their test performance (on average they made 2.3 mistakes more than in the original). In comparison, the experimental subjects enhanced their test performance in the revision (on average making 4.7 mistakes less than in the original).

In addition, analyses of variance with the independent factors 'Group (A, B)' and 'Grade Level (fifth, sixth, seventh)' were computed for the four dependent performance variables. All main effects for 'Group' are significant [ $F(1/375) > 4.01$ ;  $p < .05$ ], confirming the results of the  $t$ -tests. In contrast, neither the main effects for 'Grade Level' [ $F(2/375) < 2.51$ ;  $p > .10$ ] nor the interaction effect of 'Group' and 'Grade Level' [ $F(1/375) < 1.80$ ;  $p < .10$ ] exceed the level of significance. Therefore, the improvement of dictation test performance by self-relaxation exercises before text revision is independent of grade level. In addition, analyses of variance gave no hints for differential effects of AT-exercises in low- versus very-low-achieving students (defined with reference to the grade on last school report), because there are no significant aptitude-treatment interactions neither for the number of mistakes nor for the number of 'right' and the number of 'wrong' self-corrections.

### *Effects of AT-exercises in third dictation test*

After taking the third curriculum-based dictation test, all students applied self-relaxation and reread their texts in search of mistakes. Neither average grades nor number of mistakes and number of incorrect revisions after relaxation discriminates between the two groups under study (see Table 1). Only the number of right revisions is still significantly higher in Group A, which has been utilising the self-relaxation exercises longer, that is, for approximately 8–10 weeks in contrast to only four weeks of relaxation exercises in Group B. Thus, the short-term non-experimental maintenance probe confirms the positive effects of self-relaxation on test performance in Group A; in the former control Group B, the third dictation test with self-relaxation after testing and before rereading is novel and experimental (but without controls), thus indicating, in the time comparisons between the first, second and third dictation tests, improvements of performance as well. However, this improvement refers only to the number of mistakes and correct revisions after rereading, not to the grades themselves. In comparison with the second dictation test, the average grade point on the third test

remains below a 'C' (satisfactory) in Group B and falls back to a 'C' in Group A. This effect was a topic of lengthy and very engaged discussions with the teachers involved in the study. As a result, the teachers (who are legally responsible for educational grading) insisted on their grading strategy, which obviously referred to an adaptation of the grades given to the fewer mistakes students made in the examinations.

This upward adaptation of grades to (fewer) mistakes in the dictation tests is reflected by the average grades the students received in language arts on the next (semester) school report card as well (see bottom row in Table 1). In both groups of students, average grades on their new – the next – school report cards tend towards only an improvement in comparison with the average grades they received on their earlier – the last – school report card; in this case, improvement refers to an average of 'C-' changing to an average of 'C' without statistical significance ( $p > .10$ ).

### *Effects in school classes as unit of research*

Because the randomisation procedure referred to the unit of classes and not to the unit of individual students, it must be tested whether the statistical significant group differences survive when not the individual, but rather the school class is used as the unit of research (see e.g. Cohen, 1988; Cook & Campbell, 1979). The thereupon related loss of degrees of freedom and of variance in the dependent variables under study lead to somewhat lower effect sizes ( $.88 < d < 1.04$ ) – yet in terms of Cohen's (1988) guidelines these effect sizes are still in the large range. Importantly, there is not any loss in the statistical significance of the results presented [ $t(18) > 1.75; p < .05$ ].

### *Students' evaluations of the relaxation exercises*

Last but not least, it should be mentioned that the students evaluated the self-relaxation exercises they learned in their language arts classes at the end of the first half of the school year positively ( $M = 5.1$ ,  $SD = 0.99$  on a seven-point scale). There were neither significant differences between Group A and Group B [ $t(392) = 0.41; p > .10$ ] nor between female and male students [ $t(392) = 0.61; p > .10$ ] and between fifth, sixth and seventh graders [ $F(2/387) = 1.70; p > .10$ ] in these ratings. However, students rarely applied the self-relaxation technique outside of school: 51% never applied the short autogenic exercise outside of school, 39% did so only rarely but sometimes less than 'several times a month', 10% did so several times a month, and no student performed the exercises several times a week or daily. Thus, the self-relaxation technique that was learned at school remains at school. Nonetheless, it is positively evaluated by the students within the school environment, but it is not transferred to other domains of life.

### **Discussion and conclusion**

To sum up, the one-tailed hypotheses on the improvement of examination performance by systematic self-relaxation exercises for four minutes after taking a dictation test and before text revision is confirmed in low-achieving fifth, sixth and seventh graders. Improvement of performance is significant with reference to the assigned grades as well as to the numbers of right and of wrong 'self-corrections' performed in the revision of the texts that were written by the students themselves. In the first part of the study, this is confirmed in a field setting with a randomised, controlled design.

In terms of Cohen's (1988) guidelines, effect sizes of the group differences are large, pointing at a strong impact of the systematic self-relaxation exercises on the improvement of dictation test performance in low achievers.

Most impressive is the result that the controls – on average – in rereading performed more incorrect than correct revisions of their dictation texts, thus worsening their test performance. In a developmental study of how orthographic revisers (fifth graders to undergraduate level) check agreements in written French texts, Largy, Dedeyan, and Hupet (2004) ascertained experimentally that fifth to seventh graders apply a time-consuming, slow algorithmic procedure to verify grammatical agreements and that there was no difference in their response latency between their correct and incorrect responses. Older students used a fast decision strategy based on co-occurrence of formal indices. Their correct answers, however, were slower than their incorrect ones, which led Largy et al. (2004) to the suggestion that they also sometimes use a time-consuming, controlled procedure. Thus, along with the acquisition of writing expertise, the revision procedure itself is progressively facilitated and gradually automatised by substituting a fast decision strategy for a slow and laborious use of revision rules. It is hypothesised that in orthographical low achievers this acquisition process is delayed; revision processes are slow and confused resulting in longer decision times for correct as well as incorrect revisions. It seems as if self-relaxation before text revision improves this decision process because experimental subjects enhanced their test performance in the revision on average by attaining approximately five mistakes fewer as compared to the controls.

Improvement of performance in orthography and punctuation in the experimental group can be explained by the relaxation exercises providing the students with some personal distance to the test situation, thus optimising their stress management and reducing performance anxiety, arousal and nervousness. Therewith, a state of consciousness is created which allows students to more effectively recall their own skill and knowledge in orthography and punctuation. However, what happened to the controls during their four-minute rest period after taking the dictation test? Rather unsystematic surveys of these students on the strategies they applied during this waiting time point towards a broad spectrum of reported dozing, nervousness, task irrelevant cognitions and task relevant, but misleading cognitions on mistakes eventually or probably made on the test. Indeed, relatively many students reported experiencing persevering but ambiguous thoughts on rather specific orthographic questions, which may have made them increasingly unsure of their own orthographic knowledge. This hypothesis of negative effects of simple resting instructions after a performance test on upcoming revisions – at least in low achievers and/or students with a negative self-concept of own ability – should be systematically tested in future research.

In the short run, improvements of test performance by systematic self-relaxation techniques – by applying the basic formulas of AT – were confirmed experimentally in low achievers without any hints for aptitude–treatment interactions, i.e. no differences in the improvements of low versus very low achievers. This finding may provide a small piece of the educational puzzle to improve test performance in low achievers, leading – perhaps – in the long run to improvements in students' self-concept of own ability, self-attributions and achievement motivation. However, it must be stated that such systematic relaxation exercises as well as other stress and anxiety management training programmes will only work in both the short and the long run if students' knowledge and skills are developed by effective teaching methods (for writing instruction see the comprehensive meta-analysis presented

recently by Graham & Perin, 2007). Thus, self-relaxation exercises are 'a humanistic technique to increase achievement' (Matthews & Quinn, 1987, p. 112) in students. However, they are only complementary techniques improving educational test taking by reducing performance anxiety, high arousal and nervousness as well as by the optimising a consciousness state that is favourable for the recall of one's own knowledge, for task relevant (but clear and unambiguous) cognitions, and for self-confidence in dictation and test taking. In addition, self-relaxation techniques enhance the self-regulation skills of students.

The internal validity of the results presented here is good, at least for the first part of the study employing a randomised control design in field settings in 10 different schools. Nonetheless, there was a self-selection of the teachers involved in the study (but not on the part of the students). This threat to internal validity is reduced (1) by the fact that there was no self-selection of the teachers into or out of experimental conditions because all teachers taught experimental and control students, (2) by the independent classroom observation data on high fidelity of treatment in teachers' instructions, (3) by employing independent scorers to record the mistakes students made on the dictation tests, a strategy which controls for expectancy biases in the teachers, and (4) by the empirical proof that randomisation led to groups of students which are comparable in the crucial performance variables (baseline at pre-test). It should be added that simultaneous multilevel statistical analyses were not possible because of the low number of units of analyses at Level 2 (see, e.g. Goldstein, 1995; Snijders & Bosker, 1999). Instead, a successive two-level statistical analysis strategy (first on the level of students, second on the level of school classes) was administered, which resulted in highly comparable results. However, it should be noted that teacher involvement in the study is clear and significant in the second part of the study: All teachers decided to instruct all of their students in the self-relaxation technique, changing the randomised experimental to a pre-experimental, descriptive design. But again, the independent scoring of mistakes students made on the dictation test protects the results from expectancy biases.

Such expectancy biases – on the contrary – may be supposed in teachers' educational grading. However, instead of 'up-grading' after students' fewer mistakes as a consequence of their performing self-relaxation exercises after completing the dictation test, teachers actually implemented medium-term 'down-grading', that is, they adapted their grading to the fewer mistakes students made on average after the introduction of the self-relaxation exercises. This is a very good 'theoretical' example for the educational psychology concept of socially oriented educational grading, which is oriented towards social norms within a reference group (i.e. the single classroom) and referring to a social reference norm orientation of teachers (see, e.g. Krampen, 1987). It is a bad example for teaching and grading practices, however, because this strategy has negative impacts on students' achievement motivation. This can be changed only if teachers are trained in and convinced by grading strategies that are oriented towards objective, transparent achievement criteria (criterion referenced tests) or oriented towards intra-individual achievement comparisons for the single student (individual reference norm orientation see, e.g. Krampen, 1987). This was not an aim of the present study and could therefore not be achieved in the present context.

The results presented here actually extend the up to now quite old and methodologically rather weak literature on experimental results about the performance-improving effects of systematic relaxation exercises. They fit rather well into the results of these older research, but they expand the existing results of experimental

studies with small sample sizes and with limited external validity (because neither the treatment nor the performance tests are integrated parts of everyday life; see, e.g. Budohaska & Orłowska, 1970; Krampen, 1997; Matthews & Quinn, 1987; Siersch, 1989; Sweeney-Burton, 1998) by employing a randomised, controlled design in the field setting of language arts education and curriculum-based, regular orthography tests in a large sample of fifth, sixth and seventh graders from 10 different schools. Low achievers improve their dictation test performance significantly by the possibility to relax systematically and to cope with test stress after dictation and before rereading and self-correction. Our results confirm this for all low-achieving students without any significant differences between low and very low achievers, thus, with high generalisability for all low-achieving students. However, it must be noted that the results presented refer only to such low achievers. Thus, generalisability of results to high and medium achievers is not assured and has to be tested empirically in further research. The same is valid for primary school students below the fifth grade, for students above the seventh grade as well as for low-achieving adults.

## References

- Bandalos, D.L., Finney, S.J., & Geske, J.A. (2003). A model of statistics performance based on achievement goal theory. *Journal of Educational Psychology, 95*, 604–616.
- Budohaska, W., & Orłowska, A. (1970). Effect of autogenic training on learning and retention. *Studia Psychologiczne, 10*, 55–60.
- Chapell, M.S., Blanding, Z.B., Silverstein, M.E., Takahashi, M., Newman, B., Gubi, A., et al. (2005). Test anxiety and academic performance in undergraduate and graduate students. *Journal of Educational Psychology, 97*, 268–274.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.
- Cook, T.D., & Campbell, D.T. (1979). *Quasi-experimentation: Design and analysis for field settings*. Chicago: Rand McNally.
- De La Paz, S., & Graham, S. (2002). Explicitly teaching strategies, skills, and knowledge: Writing instruction in middle school classrooms. *Journal of Educational Psychology, 94*, 687–698.
- Durik, A.M., Vida, M., & Eccles, J.S. (2006). Task values and ability beliefs of high school literacy choices: A developmental analysis. *Journal of Educational Psychology, 98*, 382–393.
- Goldstein, H. (1995). *Multilevel statistical analysis* (2nd ed.). London: Arnold.
- Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99*, 445–476.
- Krampen, G. (1987). Differential effects of teacher comments. *Journal of Educational Psychology, 79*, 137–146.
- Krampen, G. (1997). Promotion of creativity (divergent productions) and convergent productions by systematic relaxation exercises: Empirical evidence from five experimental studies with children, young adults, and elderly. *European Journal of Personality, 11*, 83–99.
- Krampen, G. (1998). *Einführungskurse zum Autogenen Training* [Introductory courses on autogenic training] (2nd ed.). Göttingen: Hogrefe.
- Krampen, G., & von Eye, A. (2006). Treatment motives as predictors of acquisition and transfer of relaxation methods to everyday life. *Journal of Clinical Psychology, 62*, 83–96.
- Largy, P., Dedeyan, A., & Hupet, M. (2004). Orthographic revision: A developmental study of how revisors check verbal agreement in written texts. *British Journal of Educational Psychology, 74*, 533–550.
- Luthe, W. (1970). *Research and theory: Autogenic therapy*. New York: Grune & Stratton.
- Matthews, D.B., & Quinn, J.L. (1987). Relaxation training: A humanistic technique to increase achievement. *Journal of Humanistic Counseling, Education and Development, 25*, 112–121.

- McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review*, 8, 299–324.
- Nagy, W., Berninger, V., Abbott, R., Vaughan, K., & Vermeulen, K. (2003). Relationship of morphology and other language skills to literacy skills in at-risk second-grade readers and at-risk fourth-grade writers. *Journal of Educational Psychology*, 95, 730–742.
- Schultz, J.H., & Luthe, W. (1969). *Autogenic methods*. New York: Grune & Stratton.
- Siersch, K. (1989). Zum Einfluß des Autogenen Trainings auf die Konzentrationsleistung und auf subjektive Beanspruchungsfolgen [Impact of autogenic training on concentration performance and subjective stress]. *Ärztliche Praxis und Psychotherapie*, 11, 5–10.
- Snijders, T., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modelling*. London: Sage.
- Sweeney-Burton, C. (1998). Effects of self-relaxation techniques training on performance anxiety and on performance quality in a music performance condition. *Dissertation Abstracts International*, 58(7-A), 2581.