Published as

Lang, J. W. B., & Fries, S. (2006). A revised 10-item version of the Achievement Motives Scale: Psychometric properties in German-speaking samples. *European Journal of Psychological Assessment, 22*, 216-224. doi: 10.1027/1015-5759.22.3.216

Running Head: REVISED ACHIEVEMENT MOTIVES SCALE

A Revised 10-Item Version of The Achievement Motives Scale: Psychometric Properties in German-Speaking Samples

Jonas W. B. Lang

RWTH Aachen University

Stefan Fries

University of Mannheim

Jonas W. B. Lang, Institut für Psychologie, RWTH Aachen, Jägerstr. 17-19, D-52056 Aachen, Germany, E-mail: Jonas.Lang@psych.rwth-aachen.de, Tel.: +49(0)241-80-96133

Dr. Stefan Fries, Fakultät für Sozialwissenschaften, Lehrstuhl für Erziehungswissenschaft II (Pädagogische Psychologie), Universität Mannheim, Seminargebäude A 5, D-68131 Mannheim, Germany, E-mail: Stefan.Fries@phil.uni-mannheim.de, Tel.: +49(0)621-181-2212

Acknowledgements:
We thank Jessica Ippolito and Tobias Dörfler for their helpful comments on an earlier version of this paper.

Abstract

The Achievement Motives Scale (AMS) is a well established and frequently used scale to assess hope of sucess and fear of failure. In 3 studies with German-speaking samples (*N*=3523, *N*=132, *N*=126), the authors developed a revised form of the AMS using confirmatory factor analysis. As found in previous research, the original 30-item set of the AMS did not provide an acceptable fit to a 2-factor model. In contrast, a revised 10-item version (AMS-R) provided an adequate fit to the theoretically intended 2-factor model. The adequate fit could be validated in cross-validation procedures. Furthermore, the revised scales provided adequate reliability, lower inter-scale correlations and criterion-related validity with respect to typical criteria of achievement-related behavior.

Keywords: Achievement Motives Scale, motive measurement, self-attributed motive

A Revised 10-Item Version of The Achievement Motives Scale: Psychometric Properties in German-Speaking Samples

Past researchers developed a variety of measures to assess the achievement motive. One of the best established and most frequently used questionnaire-measures in achievement motivation research is the Achievement Motives Scale (AMS; Gjesme & Nygard, 1970). The AMS contains 30 items to measure hope of success and fear of failure. Previous research generally found supporting evidence for the reliability and predictive validitiy of the AMS (e.g., Dahme, Jungnickel, & Rathje, 1993; Gjesme, 1971; Rand, 1978). However, studies investigating the factor structure using exploratory factor analyses (EFA) did not clearly support the theoretically intended structure of the measure (Christophersen & Rand, 1982; Dahme et al., 1993; Hagtvet & Zuo, 2000; Man, Nygard, & Gjesme, 1994). In order to account for the unsatisfying approximation of the theoretically assumed two-factor structure and also to simply increase the economy of the scales, various authors developed shortened versions of the AMS (Baumert, Gruehn, Heyn, Köller, & Schnabel, 1997; Engeser, in preparation, cited after Rheinberg, 2004; Hagtvet & Zuo, 2000; Halvari & Kjormo, 1999) by trying to eliminate non-fitting items. The composition of these scales considerably differs due to differences in psychometric goals and methodological procedures considered by the respective authors.

What may be problematic for the quality of many revised AMS versions is that their factor structure was not cross-validated with independent samples. However, cross-validation procedures of factor structures are crucial for scale development based on empirical criteria to ensure that the developed scale is not a result of singular characteristics of the respective sample (Floyd & Widaman, 1995). Another problem of past revised AMS versions is that a validation with respect to typical criteria of achievement motivation is missing for the scales. Recently, Smith, McCarthy, and Anderson (2000) claimed that the literature on the development of shortened scales is characterized by overly optimistic views. Often, developers of shortened scales simply assume that the validity of an original measure transmits to a shortened instrument. Smith et al. point out that this assumption is not justified on psychometric principles. Simply, one can think of a short form as an alternate form of a measure with reduced coverage of the target domain. Thus, Smith et al. advise developers of shortened scales to conduct a thorough examination of the new form’s reliability and validity in independent samples. Given the variability of the proposals for revised AMS-forms and the frequently missing evidence for the validity of the revised scales, the aim of the present studies was to develop a well validated revised AMS version.

Theoretical Background: The Achievement Motive

Human beings sometimes engage in activities just for the sake of finding out whether or not they are capable to perform an activity, successfully. Such achievement-motivated behavior can be characterized by a concern with a standard of excellence that is important to the individual (McClelland, Atkinson, Clark, & Lowell, 1953). The underlying achievement motive is understood as a generalized evaluative and behavioral tendency in situations in which a standard of excellence can be applied.

Scales to measure the achievement motive are commonly seperated according to two distinctions. First, there is a consensus in the literature that the achievement motive splits into an approach and an avoidance tendency (e.g., Atkinson, 1957; Heckhausen, 1963, 1991). Typically, the approach tendency is labelled hope of success (HS) and the avoidance tendency is called fear of failure (FF). The approach/avoidance-distinction has its roots in experimental studies, which suggested that individuals do not only differ in their tendency to pursue success but also can be separated by their tendency to avoid a possible failure (Heckhausen, 1991). Furthermore, the distinction was popularized as it is a key assumption of the prominent risk-taking model (Atkinson, 1957)—a theoretical model, which was very fruitful for achievement motivation research.

Second, researchers seperate implicit and self-attributed achievement motives (McClelland, Koestner, & Weinberger, 1989; Spangler, 1992). Implicit and self-attributed motives are only slightly correlated (Spangler, 1992, found $r = .09$ in 36 meta-analyzed studies), relate to different important classes of behavior, and are activated by different types of incentives. Both types of motives are related to a specific method of motive measurement. Chiefly, implicit motives predict spontaneous behavioral trends over time and they are activated by so called activity incentives. Typically, activity incentives are an inherent feature of a task and they are experienced through a task itself. Implicit motives are typically assessed through fantasy-based measures like thematic apperception tests (TAT). Self-attributed motives predict immediate specific responses to specific situations or choice behavior. They are primarily facilitated by social incentives like rewards, expectations, demands, and norms that come from outside the task. A typical example of a social incentive is an achievementoriented instruction before the start of an experiment. Self-attributed motives are assessed using questionnaires like the AMS (see Heckhausen, 1991; Rheinberg, 2004, for overviews).

Psychometric Properties of the Original AMS

From a conceptual perspective, Heckhausen (1991) designated the AMS as the best available questionnaire-measure of the achievement motive for two reasons. First, its items ask persons to rate their positive or negative affect towards an achievement activity and hence reflect the description of achievement motivation given by McClelland et al. (1953). Second, the AMS replicates the theoretical structure of the self-attributed achievement motive by distinguishing two different dimensions. In contrast, many other questionnaires assess only one dimension or split HS and FF into further dimensions without an explicit theoretical background (see Heckhausen, 1991).

The AMS consists of 15 items to measure HS and 15 items to assess FF. The questionnaire was originally developed in Norwegian language (Gjesme & Nygard, 1970) and was later translated into a variety of languages (see Dahme et al., 1993; Göttert & Kuhl, 1980; Hagtvet & Zuo, 2000; Man et al., 1994). As outlined, empirical investigations using EFAs did not clearly support the factor structure of the AMS (Christophersen & Rand, 1982; Dahme et al., 1993; Hagtvet & Zuo, 2000; Man et al., 1994). In almost every factor analysis of the AMS, several items showed only small loadings on the intended factor or high loadings on the other factor of the two-factor solution. Furthermore, in some investigations, the scree and Kaiser criterion did not suggest the extraction of two factors. Concerning the reliability of the scales, it is typically found that the two scales for HS and FF show acceptable internal consistencies (Cronbach's α : .71 to .83 for HS and.81 to .89 for FF; Christophersen & Rand, 1982; Dahme et al., 1993). Previous studies found moderate correlations between the two scales ranging from -.17 to -.36 in samples of adolescents (Christophersen & Rand, 1982; Dahme et al., 1993; Man et al., 1994; Rand, 1978) and students (Hagtvet & Zuo, 2000). The validity of the AMS is confirmed for a wide range of criteria. For example, the AMS predicts adolescents' participation in a school research contest (Dahme et al., 1993), school grades (Gjesme, 1971) and achievement in verbal as well as numerical tasks (Rand, 1978).

Present Studies

The objectives of the present studies were threefold. Our first objective was to conduct a confirmatory factor analysis (CFA) on the complete item set of the German AMS. To the authors' knowledge, there is no study that examined CFAs on the complete item set of the AMS. Hagtvet and Zuo (2000) performed seperate CFAs on reduced item sets of each AMS-subscale to further investigate factor structures they found in EFAs. However, they did not conduct CFAs on the full AMS. Since previous studies relied on EFAs, the present study also includes EFAs for comparability reasons.

The second objective was to compose a revised AMS (AMS-R) with (a) an adequate fit to the theoretical two-factor model of the AMS through CFAs and EFAs, (b) a low HS-FF correlation, (c) acceptable internal consistencies, (d) a considerable smaller number of items in order to increase the economy of the measure, and (e) good stability in cross-validation procedures. Note that our goal was not to create scales having a maximum overlap with the original measures but instead to create revised scales with an improved factor structure.

Finally, our third objective was to thoroughly validate the newly developed AMS-R with respect to typical criteria of achievement motivated behavior. In line with achievement motivation research, we expected that individuals with a high motive to pursue success should (a) perform better (Heckhausen, 1991; Spangler, 1992), (b) be more likely to persist (Spangler, 1992), (c) enjoy working on achievement related tasks more (Puca & Schmalt, 1999), (d) evaluate themselves more positively after working on achievement related tasks (Heckhausen, 1991), (e) be more likely to experience flow during tasks (Puca & Schmalt, 1999), and (f) be more likely to set themselves realistic and challenging goals than people with a low motive to pursue success (Birney, Burdick, & Teevan, 1969; Heckhausen, 1963, 1991). In contrast, we expected that individuals with a high motive to avoid failure should (a) worry more while working on achievement-related tasks (Elliot & McGregor, 1999), (b) evaluate themselves more negatively after working on achievement-related tasks (Heckhausen, 1991), (c) be less likely to experience flow while working on achievement-related tasks (Puca & Schmalt, 1999), (d) be less likely to set realistic and challenging goals (Birney et al., 1969; Heckhausen, 1963, 1991), and (e) have more test anxiety (Elliot & McGregor, 1999) than individuals with a low motive to avoid failure. With respect to performance, Atkinson (1957) initially conceptualized FF as a strictly inhibitory tendency that debilitates performance and prevents persistence in achievement situations. In contrast, Birney et al. (1969) argued that FF inhibits performance in complex and new tasks, whereas it faciliates performance in easier tasks. Given these contradictory theoretical propositions, we included both easy and complex tasks as criteria in the present studies without generating a hypothesis concerning the relationship between FF and performance.

To achieve the outlined objectives, we conducted three studies. Study 1 accomplished the first two objectives, whereas Study 2 and Study 3 were designed to validate the AMS-R.

Method

Study 1

Participants and procedure. The data of Study 1 were obtained on the Internet. Web-based studies are a new and promising way to gain large sample sizes that are more similar to the general population than samples from university settings (e.g., Gosling, Vazire, Srivastava, & John, 2004). A total of 3523 German-speaking web-users (1567 male, 1956 female) with a mean age of 28.08 years (*SD*=10.22) volunteered to participate in the study. In terms of education, 19% of participants held a college degree, 38% had finished high school (9 to 10 years of education in Germany, Switzerland and Austria) and 61% had completed the highest degree of schooling possible (i.e., Abitur or Matura, which includes 12 to 13 years of education). Information about the study was distributed by using the website of the local psychology department and mailing lists. As an incentive, every participant had the opportunity to receive an individual feedback and win one of three prizes (50, 30 and 10 Euro).

AMS. The 30 items of the German Version of the AMS by Göttert and Kuhl (1980) were rated on 4-point scales ranging from *strongly agree (4)* to *strongly disagree (1)*.

Study 2

Participants and procedure. Participants were recruited from the campus of a German university. A total of 132 persons (62 male, 70 female) with a mean age of 25.10 years (*SD*=5.79) participated. Participants worked on the measures described subsequently sitting alone in a room of the university instructed by a trained experimenter.

AMS. The AMS was described in Study 1.

Digit-symbol performance and persistence. Participants worked on a digit-symbol substitution task similar to the task included in the Wechsler Adult Intelligence Scale (Wechsler, 1955). As in the original task, we used eight different digit-symbol combinations. Participants worked on six trials each containing 48 digit-symbol substitutions (288 overall) on a separate sheet of paper. Participants were instructed to work as fast and as precisely as possible. The time the person needed to complete the 48 substitutions was stopped by the experimenter for each trial. After each trial, participants had a short recreation period. As a measure of performance, we used the mean number of correct substitutions per second across the six trials. After all trials of the digit-substitution task were completed, the person's persistence was assessed to test the proposed positive relationship between HS and persistence. We therefore asked participants whether they were willing to work on additional digit-symbol trials in order to improve their performance (coded yes = 1, no = 0).

Task enjoyment. To tap task enjoyment, we administered three items after participants completed the digit-substitution task. We used the items "*I enjoyed working on this task*", "*This task was really challenging*", and "*Working on this task was interesting*". Participants responded to these items on 5-point scales ranging from *strongly agree* to *strongly disagree*.

Positive and negative self-evaluation. After participants had completed their work on the digit-symbol substitution task, we administered two items assessing positive self-evaluation ("*How happy do you feel about your performance in the task?*" and "*How proud of your test performance are you after working on these tasks?* ") and two items assessing negative self-evaluation ("*How annoyed do you feel about your performance in the task?* " and "*How ashamed do you feel about your performance in the task?*") Participants responded to these items on 5-point scales ranging from *strongly* to *not at all*.

Pauli task performance and goal-setting. As previously outlined, HS should contribute to moderately positive goal setting behavior, whereas FF should entice individuals to set negative, equal, or very high positive goals (Birney et al., 1969; Heckhausen, 1963, 1991). To test these predictions, we used the Pauli task, which has been previously used by Heckhausen (1963) to validate motive measures. In this task, participants work on rows of simple addition and substraction tasks. Each row contains 30 tasks and participants are given 20 s to complete each row. Participants first worked on one practice trial and then were asked to mark their goal prior to working on each of six additional trials. To evaluate participants' goal-setting behavior, we first calculated the average difference score (new goal - performance in the last trial) across the six trials for each participant. Next, we coded whether participants' average difference scores were moderately positive or not. Difference scores were coded as low positive when they were larger than 0 and smaller than 5. A cut-off value of 0 as the upper cut-off value is typically used in research on goal-setting (e.g., Birney et al., 1969). We chose 5 as the upper cut-off value as improvments above 4 were rare in the task (less than 85% of all trials) and thus reflected unrealistically high goal-setting.

Study 3

Participants and procedure. Participants were 126 (79 male, 47 female) 9th and 10th-grade students (mean age: 15.12 years, SD=0.84) from two German schools. Participants worked on the measures described subsequently in groups of 10 to 20 persons.

AMS. The AMS was described in Study 1.

Test anxiety. Elliot and McGregor (1999) considered test anxiety as a domain-specific adaptation of FF. Therefore, test anxiety should be strongly but not perfectly correlated with FF. To assess test anxiety, we used the 20 items of the German adaptation of the test anxiety inventory (Hodapp, Laux, & Spielberger, 1982). Items were rated on a 4-point scale ranging from *almost never* to *almost always*.

Spatial reasoning performance. Participants worked on the unwinding subtest (20 items; administration time: 9 min) of the Wilde Intelligence Test-2 (Kersting, Althoff, & Jäger, in press), a frequently used German intelligence test.

Worry. Worry was measured using a 3-item scale developed by Rheinberg (2004). Participants worked on these items after finishing their work on the performance task and responded to these items on 7-point scales ranging from *strongly agree* to *strongly disagree*.

Flow. Flow experiences were measured using a 9-item scale developed by Rheinberg (2004). Participants filled out the flow items after working on the performance task. Items were rated on 7-point scales ranging from *strongly agree* to *strongly disagree*.

Results

Factor Analyses of the Original AMS

We conducted both CFAs and EFAs on the full item set. All CFAs were calculated using maximum likelihood estimation. For model evaluation, we used the standardized root-mean-square residual (SRMR) to detect simple model misspecification. The comparative fit index (CFI), the Tucker-Lewis index (TLI) and the root-mean square error of approximation (RMSEA) were used as indicators of complex model misspecification (Hu & Bentler, 1998). Hu and Bentler suggested that cut-off values of $SRMR \leq .08$, $RMSEA \leq .06$, $CFI \geq .95$, and $TLI \geq .95$ indicate a relatively good fit between the hypothesized model and the observed data. Fit values of $SRMR \leq .10$, $RMSEA \leq .08$, $CFI \geq .90$, and $TLI \geq .90$ are typically considered indicators of an acceptable fit. CFA results revealed that the theoretically intended 2-factor model did not provide an acceptable fit to the data of all three studies (see Table 1). Particularly, the CFI and TLI indices pointed towards a serious amount of complex model misspecification in all samples. In Study 1, the two factors were strongly correlated in the CFA model. This was surprising as all previous studies reported lower HS-FF correlations. Howeverthose studies used samples of young adolescents or young and high-achieving persons. Thus, we hypothesized that the previous findings may be a result of specific sample characteristics. To test this hypothesis, we calculated the correlation between the two factors in two subsamples of Study 1. Subsample 1 consisted of persons under 30 years, who received a college degree (n=213). Subsample 2 contained persons under 15 years (n=51). In both subsamples, the factor correlations were comparable to the findings of previous research (Table 1).

To ensure continuity with previous research, we also conducted EFAs using maximum likelihood estimation and promax rotation (Table 1). Eigenvalues yielded a steep decline between the first and the second factor and another marked decline between the second and the third factor suggesting either a one- or a two-factor solution according to the scree criterion. For theoretical reasons, we extracted two factors. Similar to previous findings in the literature, there was a considerable number of missloading items in all samples we examined.

Scale Development

To develop the revised factor structure, we randomly divided the complete sample of

Study 1 into two subsamples so that the revised model developed on the first dataset could be cross-validated on the second dataset. Subsample 3 (n=1762) was used for the item selection procedure, whereas Subsample 4 (n=1761) was used for cross-validation purposes. To select the items we relied on CFA methodology. In a stepwise procedure, we gradually modified the full 30-item set using modification indices and factor loadings. Items were dropped from the analyses if the modification indices suggested a strong loading on the non-intended factor or if the factor loading suggested a weak loading (<.40) on the intended factor. The final item set consisted of five items for HS and five items for FF (Table 2). The reduced item set provided a good fit to the targeted two-factor model in the sample we used for the procedure (Subsample 3). The good fit could be confirmed in the cross-validation sample (Subsample 4) and Study 2, whereas the fit of the model to the data of Study 3 was acceptable (Table 1). The examination of modification indices in the item selection procedure lead to a noteworthy lower correlation between the two factors in all samples than in the original item set. Factor loadings for the AMS-R items in the cross-validation sample are provided in Table 2.

EFA analyses of the AMS-R revealed similar results as the CFAs. Comparable to the original scales, we found steep declines between the first and the second factor and markable declines between the second and the third factor. There were no missloading items in any of the samples. Factor loadings for the cross-validation sample are presented in Table 2.

Internal consistencies (Cronbach's α) were higher than .70 in all samples for both revised scales (Table 1). As a comparison, the original composition of the scales with 15 items each, resulted in slightly higher internal consistencies for both scales.

For exploratory reasons, we determined the amount of overlap between the AMS-R and the full AMS by calculating correlations between the scales of the AMS-R and scales composed of the eliminated items of the original AMS. Note that this is a conservative procedure, which likely underestimates the correspondence between the measures. However, this procedure avoids overestimation of the correlations between the full and the shortened scales through the inclusion of short form items on both sides of the correlation (Smith et al., 2000). We found that the reduced HS-scale (Study 1: *r*=.72, *r*=.73, *r*=.70 and *r*=.77 for Subsamples 1 to 4; Study 2: *r*=.68; Study 3: *r*=.66) and the shortened FF-scale (Study 1: *r*=.86, *r*=.86, *r*=.88 and *r*=.83 for Subsamples 1 to 4; Study 2: *r*=.86; Study 3: *r*=.61) were strongly correlated with the scales composed of the eliminated items suggesting that the AMS and the AMS-R measured the same construct.

Age and Gender Differences

Correlations with age and gender in the demographically diverse sample of Study 1 are presented in Table 3. Older persons tended to score higher on HS and lower on FF. Women (AMS-R: *M*=11.64, *SD*=3.40; AMS: *M*=33.13, *SD*=8.61) had higher scores on FF than men (AMS-R: *M*=10.80, *SD*=3.46; AMS: *M*=31.43, *SD*=8.82). On the HS-scales, women (AMS-R: *M*=17.06, *SD*=2.36; AMS: *M*=47.83, *SD*=6.12) did not have markable higher score than men (AMS-R: *M*=16.93 *SD*=2.41; AMS: *M*=47.93, *SD*=6.51).

Criterion-Related Validity

As expected, significant relationships between the revised HS-scale and the following criteria emerged (see Table 3): (a) performance in the Pauli task, the digit-substitution task and the reasoning task, (b) persistence, (c) task enjoyment, (d) flow, (e) positive self-evaluation, and (f) moderately positive goal-setting. In contrast, the original AMS was not significantly associated with digit-substitution performance, task enjoyment, and positive self-evaluation. For FF, we predicted (a) a positive relationship with negative self-evaluation, (b) a positive influence on worry, (c) a positive association with test anxiety, (d) a negative relationship with flow experiences, and (e) a negative relationship with moderately positive goal-setting. In line with this reasoning, we found significant correlations in the proposed directions between both FF-scales and these variables. As previously outlined, we generated no hypothesis regarding the association of FF with performance due to contradictory findings and theoretical propositions in the literature. For the revised FF-scale, we found a positive correlation with the simple digit-substitution task, no significant relationship with the slightly more complex Pauli task and a negative correlation with the complex spatial reasoning task in the present studies. These findings are in line with the assumption that FF increases performance in simple tasks and impairs performance in complex tasks (Birney et al., 1969).

Discussion

The objectives of the present research were to examine the factor structure of the AMS and to develop a revised form of the instrument. Concerning the factor structure of the full-length AMS, our results replicate previous investigations, which did not find clear support that the items of the full-length AMS are adequately described by a two-factor model. Our findings on the factor structure of the newly developed AMS-R suggest that the AMS-R provides a good solution to this issue as indicated by the better fit and lower HS-FF correlations of the AMS-R. The findings from Study 2 and Study 3 further suggest that the scales of the AMS-R are valid indicators of the two self-attributed motives. Admittedly, the correlations found are not particularly high. However, they are in line with those reported in the meta-analytic literature (Spangler, 1992). With respect to some criteria of achievement motivation, the AMS-R scales seem to be more valid than the original scales. Considering that the scales are also notably shorter, these findings are promising, especially because typically, it is found that shortened scales are less valid than longer scales (Smith et al., 2000).

To conclude, an important limitation of the present investigation is noteworthy. The findings of the current research are restricted to the German translation of the scales and need to be replicated with AMS items in other languages and cultural contexts.

In addition to providing insight into the psychometric characteristics of the AMS instrument, the results of the current investigation also have broader implications for the measurement of the self-attributed achievement motive. From a theoretical perspective, the results suggest that the two-factor conceptualization of the self-attributed achievement motive can be successfully operationalized in valid and reliable scales. From a practical perspective, the AMS-R may be more popular among researchers and respondents because it reduces frustrating responses associated with answering a large number of highly similar items.

References

Atkinson, J. W. (1957). Motivational determinants of risk-taking behavior. *Psychological Review, 64*, 359-372.

Baumert, J., Gruehn, S., Heyn, S., Köller, O., & Schnabel, K.-U. (1997). *Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU). Dokumentation - Band 1. Skalen Längsschnitt I*. [Educational and psychosocial development in adolesence. documentation - volume I, scales, longitudinal data collection I]. Berlin: Max-Planck-Institut für Bildungsforschung.

Birney, R. C., Burdick, H., & Teevan, R. C. (1969). *Fear for failure*. New York: Van Nostrand-Reinhold.

Christophersen, K., & Rand, P. (1982). Factor structure of the achievement motives scale (AMS): Two factors–two samples. *Scandinavian Journal of Educational Research, 26*, 13-28.

Dahme, G., Jungnickel, D., & Rathje, H. (1993). Güteeigenschaften der Achievement Motives Scale (AMS) von Gjesme & Nygard (1970) in der deutschen Übersetzung von Göttert und Kuhl – Vergleich der Kennwerte norwegischer und deutscher Stichproben. [Psychometric properties of a German translation of the achievement motives scale (AMS): Comparison of results from Norwegian and German samples]. *Diagnostica, 39*, 257-270.

Elliot, A. J., & McGregor, H. A. (1999). Test anxiety and the hierarchical model of approach and avoidance achievement motivation. *Journal of Personality and Social Psychology, 76*, 628–644.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis and the development and refinement of clinical assessment instruments. *Psychological Assessment, 7*, 286-299.

Gjesme, T. (1971). Motive to achieve success and motive to avoid failure in relation to school performance for pupils of different ability levels. *Scandinavian Journal of Educational Research*, *15*, 81-99.

Gjesme, T., & Nygard, R. (1970). Achievment-related motives: Theoretical considerations and construction of a measuring instrument. Unpublished report, University of Oslo.

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires. *American Psychologist, 59,* 93-104.

Göttert, R., & Kuhl, J. (1980). *LM-Fragebogen: Deutsche Übersetzung der AMS-Scale von Gjesme und Nygard* [LM-Questionnaire: German translation of the AMS-scale by Gjesme and Nygard]. Unpublished report, Ruhr-Universität Bochum.

Hagtvet, K. A., & Zuo, L. (2000). Conceptual and empirical components of an internal domain study: An illustration in terms of the Achievement Motives Scale. *Scandinavian Journal of Educational Research, 44*, 49-78.

Halvari, H., & Kjormo, O. (1999). A structural model of achievement motives, performance approach and avoidance goals and performance among norwegian olympic athlethes. *Perceptual and Motor Skills, 89*, 997-1022.

Heckhausen, H. (1963). *Hoffnung und Furcht in der Leistungsmotivation* [Hope and fear in achievment motivation]. Meisenheim/Glan, Germany: Hain.

Heckhausen, H. (1991). *Motivation and action*. New York: Springer.

Hodapp, V., Laux, L., & Spielberger, C. D. (1982). Theorie und Messung der emotionalen und kognitiven Komponenten der Prüfungsangst [Theory and measurment of the emotional and cognitive components of test anxiety]. *Zeitschrift für Differentielle und Diagnostische Psychologie, 3*, 169–184.

Hu, L.-T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*, 424-453.

Kersting, M., Althoff, K., & Jäger, A. O. (in press). *WILDE-Intelligenz-Test-2 (WIT)* [WILDE-Intelligence-Test-2 (WIT-2)] (3rd ed.). Göttingen, Germany: Hogrefe.

Man, F., Nygard, R., & Gjesme, T. (1994). The Achievement Motives Scale (AMS): Theoretical basis and results from a first try-out of a Czech form. *Scandinavian Journal of Educational Research, 38*, 209-218.

McClelland, D. C., Atkinson, J. W., Clark, R., & Lowell, E. L. (1953). *The achievement motive*. New York: Appelton-Century-Crofts.

McClelland, D. C., Koestner, R., & Weinberger, J. (1989). How do self-attributed and implicit motives differ? *Psychological Review, 96*, 690-702.

Puca, R. M., & Schmalt, H.-D. (1999). Task enjoyment: A mediator between achievement motives and performance. *Motivation and Emotion*, *23*, 15-29.

Rand, P. (1978). Some validation data for the Achievement Motives Scale (AMS). *Scandinavian Journal of Educational Research, 22*, 155-171.

Rheinberg, F. (2004). *Motivationsdiagnostik* [Motivational assessment]. Göttingen, Germany: Hogrefe.

Smith, G. T., McCarthy, D. M., & Anderson, K. G. (2000). On the sins of short-form development. *Psychological Assessment, 12*, 102-111.

Spangler, W. D. (1992). Validity of questionnaire and TAT measures of need for achievement: Two meta-analyses. *Psychological Bulletin, 112*, 140-154.

Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale - Revised*. New York:

Psychological Corporation.

Table 1

Confirmatory Factor Analyses, Exploratory Factor Analyses, Reliabilities and Inter-Scale-Correlation of the AMS and the AMS-R

		Confirmatory Factor Analysis						Exploratory Factor Analysis: First Four	Exploratory Factor Analysis: Missloading Items		Cronbach's α		HS-FF correlation	
	n	χ^2	*df*	CFI	TLI	RMSEA	SRMR	Eigenvalues	HS	FF	HS	FF	raw	latent
AMS														
Study 1	3523	7318.37	404	.85	.84	.07	.07	6.81, 1.68, 0.85, 0.70	1	1	.86	.92	-.55	-.68
Study 1, Subsample 1	213	858.57	404	.84	.82	.07	.08	6.16, 1.90, 0.86, 0.77	-	-	.83	.93	-.34	-.44
Study 1, Subsample 2	51	678.67	404	.60	.57	.12	.12	5.61, 2.92, 1.45, 1.21	2	-	.84	.90	-.23	-.23
Study 2	132	891.04	404	.73	.71	.10	.13	5.25, 2.20, 1.04, 0.73	3	1	.85	.91	-.41	-.47
Study 3	126	698.56	404	.70	.68	.08	.10	4.40, 2.47, 1.07, 1.05	-	4	.80	.86	-.34	-.38
AMS-R														
Study 1, Subsample 3	1762	123.70	34	.99	.98	.04	.03	2.68, 0.95, 0.41, 0.38	-	-	.78	.85	-.36	-.45
Study 1, Subsample 4	1761	86.77	34	.99	.99	.03	.02	2.62, 0.93, 0.39, 0.37	-	-	.77	.86	-.34	-.43
Study 1, Subsample 1	213	24.35	34	1.00	1.00	.00	.03	2.50, 0.99, 0.44, 0.37	-	-	.71	.86	-.14	-.17
Study 1, Subsample 2	51	33.03	34	1.00	1.00	.00	.08	2.41, 1.06, 0.67, 0.58	-	-	.71	.76	-.14	-.22
Study 2	132	50.29	34	.97	.96	.06	.05	2.24, 1.21, 0.38, 0.35	-	-	.80	.88	-.08	-.11
Study 3	126	55.19	34	.92	.90	.07	.08	1.94, 1.52, 0.59, 0.49	-	-	.71	.79	-.07	-.11

Note. CFI=Comparative Fit Index. TLI=Tucker-Lewis Index. RMSEA=Root-Mean-Square Error of Approximation. SRMR=Standardized Root-Mean Square Residual. Missloading Items=Items loading < .20 on the intended factor and/or higher on the non-intended factor in two factor promax-rotated maximum likelihood analyses.

Table 2

Factor Loadings for the AMS-R Items in the Cross-Validation Sample

	CFA		EFA	
Item	HS	FF	FF	HS
5. I like situations, in which I can find out how capable I am.	.63		.05	.65
7. When I am confronted with a problem, which I can possibly solve, I am enticed to start working on it immediately.	.52		-.04	.50
8. I enjoy situations, in which I can make use of my abilities.	.51		.03	.53
10. I am appealed by situations allowing me to test my abilities.	.79		-.03	.77
14. I am attracted by tasks, in which I can test my abilities.	.72		.00	.72
17. I am afraid of failing in somewhat difficult situations, when a lot depends on me.		.72	.70	-.03
21. I feel uneasy to do something if I am not sure of succeeding.		.72	.70	-.03
22. Even if nobody would notice my failure, I'm afraid of tasks, which I'm not able to solve.		.74	.75	.01
26. Even if nobody is watching, I feel quite anxious in new situations.		.75	.76	.03
28. If I do not understand a problem immediately I start feeling anxious.		.76	.75	-.02

Note. *n*=1761. For the original German versions of the items see Dahme et al. (1993)

Table 3

Correlates of the AMS-R and the AMS

		AMS-R		AMS	
Variable	Cronbach's α	HS	FF	HS	FF
Study 1 (*N*=3523)					
Age	-	.05**	-.21***	.14***	-.19***
Gender	-	.03	.12***	-.01	.10***
Study 2 (*N*=132)					
Digit-symbol task performance	.96	.16…	.26**	.03	.19*
Digit-symbol task persistence	-	.25**	.21*	.24**	.18*
Task enjoyment	.81	.19*	.10	.13	.09
Positive self-evaluation	.78	.18*	.10	.08	.08
Negative self-evaluation	.72	.08	.18*	.05	.19*
Pauli task goal-setting	-	.24**	-.15…	.22*	-.16…
Pauli task performance	.96	.15…	-.05	.14	-.10
Study 3 (*N*=126)					
Test anxiety	.81	.03	.40***	-.10	.46***
Reasoning task performance	.82	.20*	-.23*	.22*	-.24*
Worry	.83	-.10	.28**	-.07	.35***
Flow	.85	.31***	-.20*	.35***	-.26**

…*p*<.05 one-tailed. **p*<.05 two-tailed. ***p*<.01 two-tailed. ****p*<.001 two-tailed.