

Translation of a scale measuring cognitive test anxiety (G-CTAS) and its psychometric examination among medical students in Germany

Abstract

Aim of the study: Test anxiety expresses itself in a variety of physical and cognitive processes. Due to its influence on test performance, the cognitive component in particular can have a negative impact on those affected. A measuring instrument for this is not yet available in the German-speaking world but does exist in the form of the “Cognitive Test Anxiety Scale” (CTAS), among other languages, in English. The aim of this work was the creation and psychometric review of a German version of the scale (G-CTAS).

Methods: A German translation of the scale was created using a forward-backward procedure. Statistical investigations were then carried out on a cohort of medical students, which included an item analysis with calculation of difficulty, variance and item discrimination as well as the determination of the scale’s internal consistency. The criterion validity was examined using test performance and gender-specific differences.

Results: The final version contains 26 matching items with acceptable item parameters (mean values >1.46 , <3.13 ; variances >0.48 ; part-whole-discrimination-indices >0.37). Cronbach’s alpha was 0.92, the scale was therefore found to be a reliable measuring instrument. The scale validity could be confirmed by significant differences ($p < 0.01$) between total values of female and male participants as well as significant correlations ($p < 0.001$) between total values and test performance in the written and oral part of the first state examination.

Conclusion: With G-CTAS a suitable measuring instrument for cognitive test anxiety within the German-speaking world is available, which can be used, among other things, for studies concerning the relationship between stress, exams and test anxiety among medical students.

Keywords: test anxiety, cognitive test anxiety scale, worry, item analysis, validity check, test performance, gender, medical students, state examination, psychological stress

Alexandra Stefan¹
Christina M. Berchtold²
Matthias Angstwurm³

1 Ludwig-Maximilians-Universität München, Zentrale PJ-Koordination, München, Germany

2 Ludwig-Maximilians-Universität München, LMU-StaR (Staatsexamensrepetitorium), München, Germany

3 Ludwig-Maximilians-Universität München, Klinikum der Universität München, Medizinische Klinik und Poliklinik IV, München, Germany

Introduction

Test anxiety is counted among the anxiety disorders. According to the “Diagnostic and Statistical Manual of Mental Disorders, 5th Edition” (DSM-5 [1]) it is classified as social phobia and according to the “International Statistical Classification of Diseases and Related Health Problems, 10th Revision” (ICD-10 [2]) encoded as a specific phobia [3]. There is no generally accepted and therefore binding definition of test anxiety, but a theory of two components is common. It differentiates the two main components “emotionality” and “worry” [4]. Physiological symptoms such as sweating, increased heart rate and nervousness are attributed to the “emotionality” component. The “worry” component, on the other hand, is attributed with a large number of cognitive processes, originally described by Liebert & Morris as “any cognitive expression of concern about one’s own

performance” [4]. This includes i. a. comparing the own performance with comparison groups, thinking about the consequences of failing and showing low self-confidence as well as loss of self-esteem. Due to the great breadth and complexity of the processes that make up the “worry” component it is also known as cognitive test anxiety.

Test anxiety has multiple effects. A negative impact on the subjective well-being of those affected [5], [6], interrelationships with social and specific phobias [7] and a correlation with depressive symptoms [8] are known. Connections between increased test anxiety and reduced test performance were repeatedly observed [9], [10], [11], [12], [13], [14], [15]. These relationships were commonly noticed for the cognitive component, while the influence of the affective component was mostly rated as neutral or even positive [16], [17].

Various studies show that women report higher values for general and cognitive test anxiety than men [13], [15],

[17], [18], [19], [20], a clear cause for this could not be identified so far.

According to Powell [21], medical students also represent a risk collective for test anxiety. Different studies have shown that aspiring doctors are a group of people with a high psychological burden [22], [23], [24]. Depression and anxiety increase at the start of their studies [25] while mental health decreases [26]. These developments are also relevant because they continue into the working life [27], [28] and can have a negative impact on patient safety through reduced treatment quality [29]. Exams are an issue that should not be underestimated in these developments as they [25] are considered to be a relevant trigger for stress among students due to fear of being overworked and under pressure to perform [30].

Several questionnaires for measuring test anxiety have already been established and some have been translated into other languages. They are pursuing different approaches: Rost & Schermer's "Differential Anxiety Inventory" [31] records triggering and sustaining conditions and thus enables advice and therapy-oriented diagnostics. By contrast, the "Test Anxiety Inventory" by Spielberger [32] and its revised versions [33], [34], [35] are suitable for measuring the components described by Liebert & Morris. The "Cognitive Test Anxiety Scale" (CTAS) by Cassady & Johnson [36] was primarily developed for measuring and examining cognitive test anxiety. The scale and its revised versions have already been used for numerous studies abroad [37], [38], [39], [40], [41], [42], [43], [44], [45], [46] and have been translated into several languages [47], [48], [49], [50].

Problem & objectives

Due to the various negative effects, a more thorough understanding of the causes of test anxiety and the identification of risk groups is important. In the long term, there is also further research needed to develop treatment approaches. This is especially true in the context of medical studies. Precondition for this is the exact measurement using a qualified and reviewed scale. Since existing German scales are not explicitly suitable for measuring cognitive test anxiety and so far no German version of CTAS is available, the possibilities in German-speaking countries are limited. Therefore, the aim of the study was to translate CTAS into German, to statistically examine the items and to investigate the reliability and validity of the scale.

Material & methods

The study project was supervised by the ethics committee of the Ludwig-Maximilians-University of Munich (LMU) and declared as ethically unproblematic (application number 166-15).

Sample

Between autumn 2015 and autumn 2016 an online survey was conducted among medical students of the LMU after the first section of the medical state examination (first state examination). After giving their consent, all students who had been in the 4th or a higher pre-clinical semester at the time of the exam and who had therefore potentially taken part in the first state examination were contacted via e-mail via the university's mailing list. Answering the questionnaire was voluntary, there were no advantages for the students from participating in the study or disadvantages from not participating. Only fully processed questionnaires were included in the analysis. The sample consisted of a total of 291 students, the majority was female (female: n=191, 65.6%; male: n=100, 34.4%), the average age was 22.75 (SD 4.26). The sample size made it possible to carry out an item analysis and determine the internal consistency and criterion validity of the scale, but the size was not sufficient to carry out a factor analysis [51].

Questionnaire

For the online survey, a questionnaire was created consisting of demographic data, an instrument for measuring cognitive test anxiety and information on test performance.

Demographic data

Participants were initially asked about their age and gender.

Cognitive test anxiety

Cognitive Test Anxiety Scale (CTAS)

CTAS contains different facets of cognitive test anxiety, a wide range of symptoms are queried, including task-irrelevant thinking during the test preparation and during the test itself, the comparison with others, invasive thoughts during learning for the test and the examination itself as well as the tendency to skip relevant task details in examinations [36]. All 27 items from the original version of CTAS [36] were used in the study of which nine items are inverse-coded. The answer is given on a four-point Likert scale, which ranges from strong rejection to strong approval. Strong rejection is rated with one point, strong approval with four points. After recoding the inverse-coded items, high scale values are indicative of a high degree of cognitive test anxiety. The 27-point scale has a high internal consistency and high criterion validity [48]. It also proved to be a stable and consistent measure of cognitive test anxiety with high predictive power for test performance. In pilot studies a high degree of agreement was shown on the already established instruments "Test Anxiety Inventory" by Spielberger [32] and "Reactions to Tests" by Sarason [11], [36].

Translation of CTAS into German

All 27 items were translated using a forward-backward method. The authors created a first German translation with the focus less on a literal than on a content-oriented translation. The translation was discussed with a native English speaker and a preliminary version of the scale was created. This was then independently translated back into English by three bilingual persons. The preliminary version was compared with the back translations, deviations in content were identified and discussed. After all, a final version of the scale was created.

Test performance

Participants were asked about their test grades in the written and oral part of the first state examination, the specification of both grades was voluntary.

Analyses

Item analysis

Initially the nine inverse-coded items were recoded and retained for all subsequent analyses. Descriptive statistics of the items, including mean (M), standard deviation (SD) and variance (V), were calculated. Items with a very high level of difficulty of less than 10% ($M < 1.3$) or very low level of difficulty of over 90% ($M > 3.7$) were reviewed in terms of content and removed from the scale if the wording was inadequate. A review of the content also took place if an item showed a relatively small variance or the response format was not fully exhausted. Item discrimination indices were calculated using part-whole corrected discrimination indices [52]. Items with low discrimination indices ($r_{it} < 0.3$) were also reviewed for their content and removed if necessary. After an item was removed, discrimination indices of the remaining items were calculated again.

Scale reliability

To assess the reliability of the scale, its internal consistency was calculated using Cronbach's alpha. The scale was considered reliable if Cronbach's alpha > 0.8 [53].

Criterion validity

To determine the criterion validity of the scale, total values between female and male participants were compared using a t-test. The relationship between scale total values and test performance in the written and oral part of the first state examination was examined using Pearson correlation coefficient (r_p).

Programs

The "Statistical Package for the Social Sciences" (SPSS) 25 and Microsoft Excel 2013 were used for statistical data analysis.

Results

Item analysis

Item difficulty

All 27 items of the original English scale [36] are shown in table 1 together with descriptive statistics of their German version. The difficulty of the items ranged between $M=1.46$ ($SD=0.74$; item 20: "When I take a test that is difficult, I feel defeated before I even start.") and $M=3.13$ ($SD=0.77$; item 13: "I do well in speed tests in which there are time limits.") and therefore within the defined interval of difficulty. Variances of the items ranged from $V=0.48$ (item 19: "During tests, I have the feeling that I am not doing well.") and $V=1.11$ (item 25: "I feel under a lot of pressure to get good grades on tests."). Item 4 ("I tend to freeze up on things like intelligence tests and final exams."), 19 ("During tests, I have the feeling that I am not doing well.") and 20 ("When I take a test that is difficult, I feel defeated before I even start.") showed the comparatively smallest variances, but, like the other items, made full use of the response format. Since there was also no indication of inadequate wording, all three items were retained.

Item discrimination

When calculating the discrimination indices, an insufficient fit from item 25 ("I feel under a lot of pressure to get good grades on tests.") to the rest of the scale was noticed, which was confirmed after a review of the content (see table 2). Item 25 had already been identified as insufficient by Furlan et al. [48] in terms of content and was subsequently removed from the scale for all consequent analyses. After the removal of item 25, the discrimination indices of the other items ranged between $r_{it}=0.37$ (item 3: "I have less difficulty than the average college student in getting test instructions straight.") and $r_{it}=0.68$ (item 14: "During a course examination, I get so nervous that I forget facts I really know.") and thus above the previously determined value.

Scale reliability

Cronbach's alpha of the scale was 0.92.

Criterion validity

G-CTAS total values of the female participants averaged 59.76 ($SD=12.67$), G-CTAS total values of the male participants averaged 54.22 ($SD=13.59$). The values differed significantly from each other in the t-test (mean difference: 5.54, $T=3.46$, $df: 289$, $p < 0.01$). G-CTAS total values correlated significantly ($p < 0.001$) with grades in the written ($r_p=0.44$) and oral part of the first state examination ($r_p=0.40$).

Table 1: Mean (M), standard deviation (SD) and variance (V) of all 27 translated items are shown.

Item	M	SD	V
1.	2.06	0.81	0.65
2.	2.05	0.98	0.96
3. ¹	2.40	0.84	0.70
4.	1.50	0.75	0.56
5. ¹	2.75	0.93	0.86
6.	1.95	0.97	0.93
7.	1.90	0.96	0.92
8. ¹	2.66	0.96	0.93
9. ¹	2.57	0.92	0.84
10. ¹	2.60	0.82	0.68
11.	1.64	0.79	0.62
12.	1.59	0.78	0.60
13. ¹	3.13	0.77	0.60
14.	2.14	0.84	0.71
15.	2.66	0.95	0.90
16.	2.57	1.02	1.04
17. ¹	3.01	0.86	0.74
18. ¹	2.59	0.82	0.67
19.	2.10	0.70	0.48
20.	1.46	0.74	0.55
21. ¹	2.68	0.86	0.74
22.	2.22	1.01	1.01
23.	1.81	0.81	0.66
24.	1.47	0.80	0.64
25.	2.38	1.06	1.11
26.	2.02	0.87	0.76
27.	2.32	0.90	0.81

Comment: Inverse-coded items were initially recoded and are marked with ¹.

Table 2: Discrimination indices of all 27 items [36] as well discrimination indices after removal of item 25 (r_{it} [without item 25]) are shown. Furthermore, Cronbach's alpha is presented for both scales.

Item	r_{it}	r_{it} (without Item 25)
1.	.512	.508
2.	.438	.416
3. ¹	.352	.365
4.	.649	.659
5. ¹	.611	.602
6.	.503	.498
7.	.645	.650
8. ¹	.531	.531
9. ¹	.669	.674
10. ¹	.409	.431
11.	.607	.614
12.	.550	.547
13. ¹	.429	.437
14.	.668	.676
15.	.580	.573
16.	.407	.366
17. ¹	.585	.578
18. ¹	.636	.645
19.	.455	.457
20.	.600	.603
21. ¹	.550	.560
22.	.631	.636
23.	.462	.469
24.	.507	.505
25.	.243	
26.	.451	.469
27.	.660	.665
Cronbach's alpha	0.921	0.923

Comment: Inverse-coded items were initially recoded and are marked with ¹.

Discussion and conclusions

In the present study, a German version of CTAS was created and statistically reviewed and validated on a cohort of medical students. Analyses showed that 26 items of the German version have acceptable difficulty, variance and item discrimination. G-CTAS also has a high internal consistency and construct validity.

Due to its size and gender distribution, the sample is representative of a German medical faculty. These were medical students of similar ages in the same study section and thus a highly selected sample, which is why the results are less suitable for creating standard values or for defining severity of cognitive test anxiety. However, this qualifies the scale particularly for further studies on cognitive test anxiety among medical students. The item formulations enable the use of the scale for researches of cognitive test anxiety during study without any restriction regarding the subject. Further studies at German faculties can therefore follow without adapting the scale. Since a method effect could be observed repeatedly through the use of inverse-coded items, several revised versions of the scale already exist [42], [43], [44]. In the course of this work the translation of the original scale was chosen because the short form of the scale proposed by Cassady & Finch [42] can be easily generated by removing all inverse-coded items.

The items were not selected on the basis of a characteristic value, but in terms of difficulty, variance and item discrimination. The wording of the content played a decisive role in excluding an item. According to Bühner [51], the content should have highest priority when selecting items, but this method has rarely been used in previous analyzes with CTAS and its translations. Nevertheless, the result of the item analysis coincides with the observations by Furlan et al. [48], which also led to the exclusion of item 25. Item 25 probably measures primarily performance pressure and less cognitive test anxiety.

Cronbach's alpha of G-CTAS was also comparable to preliminary investigations [48]. The value indicates a high internal consistency and thus for a homogeneous scale. A detailed analysis of the factor structure could not be conducted because the sample size was not sufficient for an exploratory and subsequent confirmatory factor analysis [51]. This should be carried out as the next step, since precise knowledge of the scale's structure is not only the basis for the formation of standard values but could also contribute to a more thorough understanding of the causes and effects of test anxiety.

Differences between female and male participants, as well as significant, positive correlations with test performance were able to prove the criterion validity of the German scale. The study design did not address gender-specific differences in cognitive test anxiety, which is why their causes still remain unclear. In contrast, the study design made it possible to compare the connections between cognitive test anxiety with written and oral exams. Here, a significant correlation with both test modalities was observed at a similarly high level of correlation.

The connection of cognitive test anxiety with both, written and oral exams, once again illustrates the possible extent for those affected. Connections with the performance in a relevant test such as the first state examination emphasize how decisive test anxiety can be in relation to academic achievement. Finding the causes of test anxiety and treatment approaches is not only important from the perspective of those affected but should also be of high priority for universities. Medical students represent a risk collective regarding mental stress and test anxiety and should be supported in this regard. A more precise understanding of the dynamics between stress, exams and test anxiety is essential in order to identify people at risk and to be able to intervene in these negative developments at an early stage. The basis for this is a suitable and most accurate measuring instrument of cognitive test anxiety. With G-CTAS this is now also available in German-speaking countries.

Acknowledgements

Our thanks go to Dr. Amanda Tufman, Mrs. Sarah Garcia, Mrs. Franziska Enders and Mrs. Miruh Lee for their assistance in translating the scale.

Competing interests

The authors declare that they have no competing interests.

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). 5th Edition. Washington, DC: American Psychiatric Association; 2013.
2. Deutsches Institut für Medizinische Dokumentation und Information (DIMDI). CD-10-GM Version 2019, Systematisches Verzeichnis, Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme, 10. Revision. Köln: Deutsches Institut für Medizinische Dokumentation und Information (DIMDI); 2018.
3. Dachs L. Prüfungsängste. In: Schneider S, Margraf J, editors. Lehrbuch der Verhaltenstherapie, Band 3: Psychologische Therapie bei Indikationen im Kindes- und Jugendalter. Berlin, Heidelberg: Springer Berlin Heidelberg; 2019. p.569-586. DOI: 10.1007/978-3-662-57369-3_32
4. Liebert R, Morris L. Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychol Report*. 1967;20:975-978. DOI: 10.2466/pr0.1967.20.3.975
5. Steinmayr R, Crede J, McElvany N, Wirthwein L. Subjective Well-Being, Test Anxiety, Academic Achievement: Testing for Reciprocal Effects. *Front Psychol*. 2016;6:1994. DOI: 10.3389/fpsyg.2015.01994
6. Pixner S, Kaufmann L. Prüfungsangst, Schulleistung und Lebensqualität bei Schülern. *Lern Lernstörung*. 2013;2(2):111-124. DOI: 10.1024/2235-0977/a000034

7. Schaefer A, Matthes H, Pfitzer G, Köhle K. Mental health and performance of medical students with high and low test anxiety. *Psychother Psychosom Med Psychol.* 2007;57(7):289-297. DOI: 10.1055/s-2006-951974
8. Akinsola EF, Nwajeri AD. Test anxiety, depression and academic performance: assessment and management using relaxation and cognitive restructuring techniques. *Psychol.* 2013;4(06):18. DOI: 10.4236/psych.2013.46A1003
9. Morris LW, Liebert RM. Effects of anxiety on timed and untimed intelligence tests: Another look. *J Cons Clin Psychol.* 1969;33(2):240-244. DOI: 10.1037/h0027164
10. Morris LW, Davis MA, Hutchings CH. Cognitive and emotional components of anxiety: Literature review and a revised worry-emotionality scale. *J Educ Psychol.* 1981;74(4):541-555. DOI: 10.1037/0022-0663.73.4.541
11. Sarason IG. Stress, anxiety, and cognitive interference: Reactions to tests. *J Person Soc Psychol.* 1984;46(4):929-938. DOI: 10.1037/0022-3514.46.4.929
12. Deffenbacher JL, Hazaleus SL. Cognitive, emotional, and physiological components of Test Anxiety. *Cogn Ther Res.* 1985;9(2):169-180. DOI: 10.1007/BF01204848
13. Hembree R. Correlates, causes, and treatment of test anxiety. *Rev Educ Res.* 1988;58:47-77. DOI: 10.3102/00346543058001047
14. Thomas CL, Cassady JC, Heller ML. The influence of emotional intelligence, cognitive test anxiety, and coping strategies on undergraduate academic performance. *Learn Indiv Diff.* 2017;55:40-48. DOI: 10.1016/j.lindif.2017.03.001
15. Putwain DW. Test anxiety and GCSE performance: The effect of gender and socio-economic background. *Educ Psychol Pract.* 2008;24(4):319-334. DOI: 10.1080/02667360802488765
16. Spielberger CD, Anton WD, Bedell J. The nature and treatment of test anxiety. In: Zuckerman M, Spielberger CD, editors. *Emotions and Anxiety (PLE: Emotion)*. London: Psychology Press; 2015. p.317-344.
17. Lowe PA. Should Test Anxiety be Measured Differently for Males and Females? Examination of Measurement Bias Across Gender on Measures of Test Anxiety for Middle and High School, and College Students. *J Psychoeduc Ass.* 2015;33(3):238-246. DOI: 10.1177/0734282914549428
18. Zeidner M. Does test anxiety bias scholastic aptitude test performance by gender and sociocultural group? *J Pers Assess.* 1990;55(1-2):145-160. DOI: 10.1080/00223891.1990.9674054
19. Chapell MS, Blanding ZB, Silverstein ME, Takahashi M, Newman B, Gubi A, Mccann N. Test Anxiety and Academic Performance in Undergraduate and Graduate Students. *J Educ Psychol.* 2005;97(2):268-274. DOI: 10.1037/0022-0663.97.2.268
20. Nilofer Farooqi Y, Ghani R, D. Spielberger C. Gender Differences in Test Anxiety and Academic Performance of Medical Students. *Intern J Psychol Behav Sci.* 2012;2(2):38-43. DOI: 10.5923/j.ijpbs.20120202.06
21. Powell DH. Behavioral treatment of debilitating test anxiety among medical students. *J Clin Psychol.* 2004;60(8):853-865. DOI: 10.1002/jclp.20043
22. Dyrbye LN, Thomas MR, Shanafelt TD. Systematic review of depression, anxiety, and other indicators of psychological distress among U.S. and Canadian medical students. *Acad Med.* 2006;81(4):354-373. DOI: 10.1097/00001888-200604000-00009
23. Dahlin M, Joneborg N, Runeson B. Stress and depression among medical students: A cross-sectional study. *Med Educ.* 2005;39(6):594-604. DOI: 10.1111/j.1365-2929.2005.02176.x
24. Iqbal S, Gupta S, Venkatarao E. Stress, anxiety & depression among medical undergraduate students & their socio-demographic correlates. *Indian J Med Res.* 2015;141(3):354. DOI: 10.4103/0971-5916.156571
25. Burger PH, Tektas OY, Paulsen F, Scholz M. Vom Studienstart bis zum ersten Staatsexamen - Zunahme von Depressivität bei gleichzeitigem Verlust des Kohärenzgefühls und der psychischen Lebensqualität in höheren Semestern Humanmedizin. *Psychother Psych Med.* 2014;64(08):322-327. DOI: 10.1055/s-0034-1374593
26. Moffat KJ, McConnachie A, Ross S, Morrison JM. First year medical student stress and coping in a problem-based learning medical curriculum. *Med Educ.* 2004;38(5):482-491. DOI: 10.1046/j.1365-2929.2004.01814.x
27. Voltmer E, Kötter T, Westermann J. Prävention: Gesund durchs Medizinstudium. *Dtsch Ärztebl.* 2015;112(35-36):A-1414/B-1192/c-1164.
28. Dyrbye L, Shanafelt T. A narrative review on burnout experienced by medical students and residents. *Med Educ.* 2016;50(1):132-149. DOI: 10.1111/medu.12927
29. Panagioti M, Geraghty K, Johnson J, Zhou A, Panagopoulou E, Chew-Graham C, Peters D, Hodkinson A, Riley R, Esmail A. Association Between Physician Burnout and Patient Safety, Professionalism, and Patient Satisfaction: A Systematic Review and Meta-analysis. *JAMA Intern Med.* 2018;178(10):1317-1330. DOI: 10.1001/jamainternmed.2018.3713
30. Herbst U, Voeth M, Eidhoff AT, Müller M, Stief S. *Studierendenstress in Deutschland - eine empirische Untersuchung*. Berlin: AOK-Bundesverband; 2016.
31. Rost D, Schermer F. *Differentielles Leistungsangst-Inventar [Differential Achievement Anxiety Inventory]*. Frankfurt: Pearson; 2007.
32. Spielberger CD. Test anxiety inventory. *Psychol.* 1980;31:503-509. DOI: 10.1146/annurev.ps.31.020180.002443
33. Hodapp V. Das Prüfungängstlichkeitsinventar TAI-G: Eine erweiterte und modifizierte Version mit vier Komponenten [The Test Anxiety Inventory TAI-G: An expanded and modified version with four components]. *Z Päd Psychol.* 1991;5(2):121-130.
34. Hodapp V. The TAI-G: A multidimensional approach to the assessment of test anxiety. In: Schwarzer C, Zeidner M, editors. *Stress, anxiety, and coping in academic settings*. Tübingen: Francke; 1996. p.95-130.
35. Hodapp V, Rohrmann S, Ringeisen T. *Prüfungsangstfragebogen (PAF) [German Test Anxiety Questionnaire]*. Göttingen: Hogrefe; 2011.
36. Cassady JC, Johnson RE. Cognitive Test Anxiety and Academic Performance. *Contemp Educ Psychol.* 2002;27(2):270-295. DOI: 10.1006/ceps.2001.1094
37. Ramirez G, Beilock SL. Writing About Testing Worries Boosts Exam Performance in the Classroom. *Science.* 2011;331:211-213. DOI: 10.1126/science.1199427
38. Cassady JC. The impact of cognitive test anxiety on text comprehension and recall in the absence of external evaluative pressure. *Appl Cogn Psychol.* 2004;18(3):311-325. DOI: 10.1002/acp.968
39. Cassady JC. The stability of undergraduate students' cognitive test anxiety levels. *Pract Ass ResEval.* 2001;7(20):1-5.
40. Tsianos N, Lekkas Z, Germanakos P, Mourlas C, Samaras G. An experimental assessment of the use of cognitive and affective factors in adaptive educational hypermedia. *IEEE Trans Learn Technol.* 2009;2(3):249-58. DOI: 10.1109/TLT.2009.29
41. Kapetanaki EM, editor. *Test anxiety in education: Policies and practices*. Presentation at the 31st World Conference on Stress and Anxiety Research. August 4-6; Galway, Ireland. 2010.

42. Cassady JC, Finch WH. Confirming the Factor Structure of the Cognitive Test Anxiety Scale: Comparing the Utility of Three Solutions. *Educ Ass.* 2014;19(3):229-242. DOI: 10.1080/10627197.2014.934604
43. Cassady JC, Finch WH. Using factor mixture modeling to identify dimensions of cognitive test anxiety. *Learn Individ Diff.* 2015;41:14-20. DOI: 10.1016/j.lindif.2015.06.002
44. Thomas CL, Cassady JC, Finch WH. Identifying Severity Standards on the Cognitive Test Anxiety Scale: Cut Score Determination Using Latent Class and Cluster Analysis. *J Psychoeduc Ass.* 2018;36(5):492-508. DOI: 10.1177/0734282916686004
45. Cassady JC. The influence of cognitive test anxiety across the learning-testing cycle. *Learn Instruct.* 2004;14(6):569-592. DOI: 10.1016/j.learninstruc.2004.09.002
46. Zheng Y. Chinese University Students' Motivation, Anxiety, Global Awareness, Linguistic Confidence, and English Test Performance: A Correlational and Causal Investigation. 2010. Zugänglich unter/available from: <https://qspace.library.queensu.ca/handle/1974/5378>
47. Cassady JC, Mohammed A, Mathieu L. Cross-Cultural Differences in Test Perceptions. *J Cross Cult Psychol.* 2004;35(6):713-718. DOI: 10.1177/0022022104270113
48. Furlan LA, Cassady JC, Pérez ER. Adapting the Cognitive Test Anxiety Scale for use with Argentinean University Students. *Intern J Test.* 2009;9(1):3-19. DOI: 10.1080/15305050902733448
49. Baghaei P, Cassady J. Validation of the Persian Translation of the Cognitive Test Anxiety Scale. *SAGE Open.* 2014;4(4):215824401455511. DOI: 10.1177/2158244014555113
50. Bozkurt S, Ekitli GB, Thomas CL, Cassady JC. Validation of the Turkish Version of the Cognitive Test Anxiety Scale-Revised. *SAGE Open.* 2017;7(1):215824401666954. DOI: 10.1177/2158244016669549
51. Bühner M. Einführung in die Test- und Fragebogenkonstruktion. München: Pearson Studium; 2011. p.640.
52. Rotthoff T, Ostapczuk MS, De Bruin J, Decking U, Schneider M, Ritz-Timme S. Assessing the learning environment of a faculty: psychometric validation of the German version of the Dundee Ready Education Environment Measure with students and teachers. *Med Teach.* 2011;33(11):e624-36. DOI: 10.3109/0142159X.2011.610841
53. Bortz J, Döring N. *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler.* Heidelberg: Springer; 2006. DOI: 10.1007/978-3-540-33306-7

Corresponding author:

Alexandra Stefan

Ludwig-Maximilians-Universität München, Zentrale PJ-Koordination, Pettenkoferstr. 8a, D-80336 München, Germany

alexandra.stefan@med.uni-muenchen.de**Please cite as**

Stefan A, Berchtold CM, Angstwurm M. Translation of a scale measuring cognitive test anxiety (G-CTAS) and its psychometric examination among medical students in Germany. *GMS J Med Educ.* 2020;37(5):Doc50. DOI: 10.3205/zma001343, URN: urn:nbn:de:0183-zma0013436

This article is freely available from

<https://www.egms.de/en/journals/zma/2020-37/zma001343.shtml>

Received: 2019-08-15**Revised:** 2020-01-15**Accepted:** 2020-05-29**Published:** 2020-09-15**Copyright**

©2020 Stefan et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 License. See license information at <http://creativecommons.org/licenses/by/4.0/>.

Übersetzung einer Skala zur Messung kognitiver Prüfungsangst (G-CTAS) und deren psychometrische Überprüfung unter Humanmedizinstudierenden in Deutschland

Zusammenfassung

Zielsetzung: Prüfungsangst äußert sich in einer Vielzahl körperlicher und kognitiver Prozesse. Durch ihren Einfluss auf Prüfungsleistungen kann sich besonders die kognitive Prüfungsangstkomponente negativ auf Betroffene auswirken. Ein Messinstrument hierfür ist im deutschen Sprachraum bisher nicht verfügbar, existiert aber in Form der „Cognitive Test Anxiety Scale“ (CTAS) unter anderem in englischer Sprache. Ziel dieser Arbeit war die Erstellung und psychometrische Überprüfung einer deutschen Version der Skala (G-CTAS).

Methodik: Mithilfe eines Vorwärts-Rückwärts-Verfahrens wurde eine deutsche Übersetzung der Skala erstellt. An einer Kohorte Humanmedizinstudierender wurden anschließend statistische Untersuchungen durchgeführt, welche eine Itemanalyse mit Berechnung von Schwierigkeit, Varianz und Trennschärfe sowie die Ermittlung der internen Skalenkonsistenz beinhalteten. Eine Überprüfung der Kriteriumsvalidität erfolgte über Prüfungsleistungen sowie geschlechtsspezifische Unterschiede.

Ergebnisse: Die finale Version enthält 26 inhaltlich passende Items mit akzeptablen Itemkennwerten (Mittelwerte $>1,46$, $<3,13$; Varianzen $>0,48$; Trennschärfeindices $>0,37$). Cronbach's alpha betrug 0,92, die Skala stellt somit ein reliables Messinstrument dar. Die Skalensvalidität konnte durch signifikante Unterschiede ($p<0,01$) zwischen Summenwerten weiblicher und männlicher Teilnehmer sowie signifikante Korrelationen ($p<0,001$) zwischen Summenwerten und Prüfungsleistungen im schriftlichen und mündlichen Physikum bestätigt werden.

Schlussfolgerung: Mit G-CTAS steht ein geeignetes Messinstrument kognitiver Prüfungsangst auch für den deutschen Sprachraum zur Verfügung, welches sich unter anderem für Untersuchungen des Zusammenhangs zwischen Stress, Prüfungen und Prüfungsangst unter Medizinstudierenden anwenden lässt.

Schlüsselwörter: Prüfungsangst, Cognitive Test Anxiety Scale, worry, Itemanalyse, Validitätsprüfung, Prüfungsleistung, Gender, Medizinstudierende, Staatsexamen, psychische Belastung

Einleitung

Prüfungsangst wird zu den Angststörungen gezählt. Nach dem „Diagnostic and Statistical Manual of Mental Disorders, 5th Edition“ (DSM-5 [1]) wird sie als soziale, nach der „International Statistical Classification of Diseases and Related Health Problems, 10th Revision“ (ICD-10 [2]) als spezifische Phobie kodiert [3]. Eine allgemein akzeptierte und somit verbindliche Definition von Prüfungsangst liegt nicht vor, weit verbreitet ist jedoch die „Zwei-Komponenten-Theorie“, welche die beiden Hauptkomponenten,

„emotionality“ und „worry“ unterscheidet [4]. Physiologische Symptome wie Schwitzen, Anstieg der Herzfrequenz und Nervosität werden der „emotionality“-Komponente zugeschrieben. Der „worry“-Komponente wird dagegen eine Vielzahl an kognitiven Prozessen zugerechnet, ursprünglich von Liebert & Morris beschrieben als „jeder kognitive Ausdruck von Besorgnis über die eigene (Prüfungs-)Leistung“ [4]. Dazu zählen unter anderem der Vergleich der eigenen Leistung mit Vergleichsgruppen, Nachdenken über die Konsequenzen des eigenen Versagens und geringes Selbstbewusstsein sowie Verlust des Selbstwertgefühls. Aufgrund der großen Breite und Komplexität der Prozesse, welche die „worry“-Komponen-

Alexandra Stefan¹
Christina M. Berchtold²
Matthias Angstwurm³

1 Ludwig-Maximilians-Universität München, Zentrale PJ-Koordination, München, Deutschland

2 Ludwig-Maximilians-Universität München, LMU-StaR (Staatsexamensrepetitorium), München, Deutschland

3 Ludwig-Maximilians-Universität München, Klinikum der Universität München, Medizinische Klinik und Poliklinik IV, München, Deutschland

te ausmachen, wird diese auch als kognitive Prüfungsangst bezeichnet.

Prüfungsangst hat vielfältige Auswirkungen. Bekannt sind ein negativer Einfluss auf das subjektive Wohlbefinden von Betroffenen [5], [6], Zusammenhänge mit sozialen und spezifischen Phobien [7] sowie eine Korrelation mit depressiven Symptomen [8]. Ebenfalls konnten wiederholt Zusammenhänge zwischen erhöhter Prüfungsangst und reduzierten Prüfungsleistungen festgestellt werden [9], [10], [11], [12], [13], [14], [15]. Diese Zusammenhänge konnten vermehrt für die kognitive Komponente beobachtet werden, während der Einfluss der affektiven Komponente meist als neutral oder sogar positiv bewertet wurde [16], [17].

Verschiedene Studien zeigen, dass Frauen höhere Werte für allgemeine und kognitive Prüfungsangst angeben als Männer [13], [15], [17], [18], [19], [20], eine eindeutige Ursache hierfür konnte bisher nicht identifiziert werden. Laut Powell [21] stellen auch Humanmedizinstudierende ein Risikokollektiv hinsichtlich Prüfungsangst dar. Unterschiedliche Studien zeigten, dass es sich bei angehenden Ärzten um eine psychisch stark belastete Personengruppe handelt [22], [23], [24]. So nehmen Depressivität und Ängstlichkeit bereits zu Beginn des Studiums zu [25], die psychische Gesundheit ab [26]. Relevant sind diese Entwicklungen auch, da sie sich bis in die Berufstätigkeit fortsetzen [27], [28] und durch eine reduzierte Behandlungsqualität negativ auf die Patientensicherheit auswirken können [29]. Eine nicht zu unterschätzende Rolle in diesen Entwicklungen stellen Prüfungen dar, da sie durch Angst vor Überforderung und Leistungsdruck als relevanter Auslöser für Stress unter Studierenden gelten [30]. Mehrere Fragebögen zur Messung von Prüfungsangst sind bereits etabliert und zum Teil in mehrere Sprachen übersetzt worden. Sie verfolgen dabei unterschiedliche Ansätze: Das „Differentielle Leistungsangst Inventar“ von Rost & Schermer [31] erfasst unter anderem auslösende sowie aufrechterhaltende Bedingungen und ermöglicht hierdurch eine beratungs- und therapieorientierte Diagnostik. Zur Messung der von Liebert & Morris beschriebenen Komponenten eignen sich dagegen unter anderem das „Prüfungsangstinventar“ von Spielberger [32] und dessen überarbeitete Versionen [33], [34], [35]. Die „Cognitive Test Anxiety Scale“ (CTAS) von Cassidy & Johnson [36] wurde primär zur Messung und Untersuchung kognitiver Prüfungsangst entwickelt. Die Skala und deren überarbeitete Versionen wurden bereits für zahlreiche Studien im Ausland verwendet [37], [38], [39], [40], [41], [42], [43], [44], [45], [46] und in mehrere Fremdsprachen übersetzt [47], [48], [49], [50].

Problemstellung & Zielsetzung

Aufgrund der vielfältigen negativen Auswirkungen ist ein genaueres Verständnis von Prüfungsangstursachen sowie die Identifikation von Risikogruppen wichtig. Langfristig ergibt sich außerdem weiterer Forschungsbedarf zur Entwicklung von Behandlungsansätzen. Dies gilt insbe-

sondere im Kontext des Medizinstudiums. Voraussetzung hierfür ist die exakte Messung mithilfe einer geeigneten und überprüften Skala. Da bestehende, deutsche Skalen nicht explizit für die Messung kognitiver Prüfungsangst geeignet sind und bislang keine deutsche Version von CTAS zur Verfügung steht, sind die Möglichkeiten im deutschsprachigen Raum eingeschränkt. Ziel war es daher, CTAS ins Deutsche zu übersetzen, die Items statistisch zu überprüfen sowie eine Untersuchung der Reliabilität und Validität der Skala vorzunehmen.

Material & Methoden

Das Studienvorhaben wurde von der Ethikkommission der Ludwig-Maximilians-Universität München (LMU) geprüft und für unbedenklich erklärt (Antragsnummer 166-15).

Stichprobe

Zwischen Herbst 2015 und Herbst 2016 wurde an der LMU eine Onlinebefragung unter Medizinstudierenden nach dem ersten Abschnitt der ärztlichen Prüfung (Physikum) durchgeführt. Nach deren Einwilligung wurden mittels Verteiler der Universität alle Studierenden via E-Mail kontaktiert, welche sich zum Prüfungszeitpunkt im 4. oder höheren vorklinischen Semester befunden hatten und somit potenziell am Physikum teilgenommen hatten. Die Beantwortung des Fragebogens war freiwillig, es resultierten für die Studierenden keine Vorteile aus einer Studienteilnahme oder Nachteile aus einer Nicht-Teilnahme. Nur vollständig bearbeitete Fragebögen wurden in die Auswertung miteinbezogen. Die Stichprobe setzte sich aus insgesamt 291 Studierenden zusammen, der überwiegende Anteil war weiblich (weiblich: N=191, 65,6%; männlich: N=100, 34,4%), das durchschnittliche Alter lag bei 22,75 (SD 4,26). Die Stichprobengröße ermöglichte die Durchführung einer Itemanalyse sowie die Bestimmung der internen Konsistenz und Kriteriumsvalidität der Skala, die Größe war jedoch nicht ausreichend für die Durchführung einer Faktorenanalyse [51].

Fragebogen

Für die Onlinebefragung wurde ein Fragebogen, bestehend aus demographischen Daten, einem Instrument zur Messung kognitiver Prüfungsangst sowie Angaben zu Prüfungsleistungen, erstellt.

Demographische Daten

Teilnehmer wurden zu Beginn nach Alter und Geschlecht befragt.

Kognitive Prüfungsangst

Cognitive Test Anxiety Scale (CTAS)

CTAS beinhaltet unterschiedliche Facetten kognitiver Prüfungsangst, abgefragt wird ein breites Spektrum an

Symptomen, darunter Aufgabenirrelevantes Denken während der Prüfungsvorbereitung und in der Prüfung selbst, der Vergleich mit anderen, aufdrängende Gedanken während dem Lernen und in der Prüfung sowie die Neigung, relevante Aufgabedetails in Prüfungen zu überlesen [36]. In der Studie wurden alle 27 Items der ursprünglichen Version von CTAS [36] verwendet, wovon neun Items invers gepolt sind. Die Beantwortung erfolgt auf einer vierstufigen Likertskala, welche von starker Ablehnung bis zu starker Zustimmung reicht. Starke Ablehnung wird mit einem Punkt bewertet, starke Zustimmung mit vier Punkten. Nach Umkodierung der invers gepolten Items sind hohe Skalenwerte indikativ für ein hohes Ausmaß kognitiver Prüfungsangst. Die 27-stufige Skala verfügt über eine hohe interne Konsistenz und hohe Kriteriumsvalidität [46]. Sie erwies sich außerdem als stabiles und konsistentes Messinstrument kognitiver Prüfungsangst mit einer hohen Vorhersagekraft für Prüfungsleistungen. Eine hohe Übereinstimmung konnte in Pilotstudien zu den bereits etablierten Instrumenten „Prüfungsangstinventar“ von Spielberger [32] und „Reactions to Tests“ von Sarason [11] gezeigt werden [36].

Übersetzung von CTAS ins Deutsche

Alle 27 Items wurden mittels Vorwärts-Rückwärts-Verfahren übersetzt. Eine erste deutsche Übersetzung wurde von den Autoren erstellt, wobei der Fokus weniger auf einer wörtlichen als auf einer inhaltsgetreuen Übersetzung lag. Die Übersetzung wurde mit einer englischen Muttersprachlerin diskutiert und eine vorläufige Version der Skala erstellt. Diese wurde anschließend von drei zweisprachig aufgewachsenen Personen unabhängig voneinander zurück ins Englische übersetzt. Die vorläufige Version wurde mit den Rückübersetzungen verglichen, dabei wurden inhaltliche Abweichungen identifiziert und diskutiert. Schließlich wurde eine endgültige Version der Skala erstellt.

Prüfungsleistungen

Teilnehmer wurden zu ihren Prüfungsnoten im schriftlichen und mündlichen Physikum befragt, die Angabe beider Noten war freiwillig.

Analysen

Itemanalyse

Zu Beginn wurde eine Umkodierung der neun invers gepolten Items durchgeführt und für alle nachfolgenden Analysen beibehalten. Deskriptive Statistiken der Items, darunter Mittelwert (M), Standardabweichung (SD) und Varianz (V), wurden berechnet. Items mit einer sehr hohen Schwierigkeit von unter 10% ($M < 1,3$) oder sehr geringen Schwierigkeit von über 90% ($M > 3,7$) wurden inhaltlich überprüft und bei mangelhafter inhaltlicher Formulierung von der Skala entfernt. Eine inhaltliche Überprüfung fand außerdem statt, wenn ein Item eine relativ geringe Varianz

aufwies oder das Antwortformat unvollständig ausschöpfte. Die Trennschärfe der Items berechnete sich über part-whole korrigierte Diskriminationsindices [51]. Items mit niedrigen Trennschärfeindices ($r_{it} < 0,3$) wurden ebenfalls bezüglich ihres Inhalts überprüft und falls nötig entfernt. Nach Entfernung eines Items wurde erneut eine Berechnung der Trennschärfen der verbliebenen Items vorgenommen.

Skalenreliabilität

Zur Realitätseinschätzung der Skala wurde deren interne Konsistenz mittels Cronbach's alpha berechnet. Die Skala galt als reliabel, wenn Cronbach's alpha $> 0,8$ [52].

Kriteriumsvalidität

Zur Ermittlung der Kriteriumsvalidität wurden Skalensummenwerte zwischen weiblichen und männlichen Teilnehmern mittels T-Test verglichen. Der Zusammenhang von Skalensummenwerten mit Prüfungsleistungen im schriftlichen und mündlichen Physikum mithilfe des Korrelationskoeffizienten nach Pearson (r_p) untersucht.

Programme

Für die statistische Datenauswertung wurden das „Statistical Package for the Social Sciences“ (SPSS) 25 und Microsoft Excel 2013 verwendet.

Ergebnisse

Itemanalyse

Schwierigkeit

Die deutsche Übersetzung aller 27 Items ist gemeinsam mit deren deskriptiven Statistik in Tabelle 1 abgebildet. Die Schwierigkeit der Items lag zwischen $M=1,46$ ($SD=0,74$; Item 20: „Wenn ich an einer anspruchsvollen Prüfung teilnehme, fühle ich mich schon als Versager, bevor ich überhaupt angefangen habe.“) und $M=3,13$ ($SD=0,77$; Item 13: „Mir fallen Tests unter Zeitdruck besonders leicht.“) und somit innerhalb des festgelegten Schwierigkeitsintervalls. Varianzen der Items lagen zwischen $V=0,48$ (Item 19: „Während einer Prüfung habe ich das Gefühl, dass ich mich nicht besonders gut schlage.“) und $V=1,11$ (Item 25: „Ich fühle mich ziemlich unter Druck, gute Noten zu schreiben.“). Item 4 („Bei wichtigen Klausuren oder Intelligenztests tendiere ich dazu, plötzlich wie gelähmt zu sein.“), 19 („Während einer Prüfung habe ich das Gefühl, dass ich mich nicht besonders gut schlage.“) und 20 („Wenn ich an einer anspruchsvollen Prüfung teilnehme, fühle ich mich schon als Versager, bevor ich überhaupt angefangen habe.“) wiesen die vergleichsweise geringsten Varianzen auf, schöpften aber wie auch die übrigen Items das Antwortformat voll aus. Da sich zudem kein Anhalt für inhaltlich problematische

Tabelle 1: Dargestellt ist die deutsche Übersetzung aller 27 Items mit Mittelwert (M), Standardabweichung (SD) und Varianz (V).

Item	M	SD	V
1. Bevorstehende Prüfungen bereiten mir schlaflose Nächte.	2,06	± 0,81	0,65
2. Während einer wichtigen Prüfung denke ich darüber nach, ob die anderen Studenten vielleicht besser abschneiden als ich.	2,05	± 0,98	0,96
3. Mir fällt es weniger schwer als dem durchschnittlichen Studenten, Prüfungsanweisungen präzise aufzufassen. ¹	2,40	± 0,84	0,70
4. Bei wichtigen Klausuren oder Intelligenztests tendiere ich dazu, plötzlich wie gelähmt zu sein.	1,50	± 0,75	0,56
5. Ich bin vor Tests weniger aufgeregt als der durchschnittliche Student. ¹	2,75	± 0,93	0,86
6. Während einer Prüfung denke ich darüber nach, welche Konsequenzen es hätte, wenn ich versage.	1,95	± 0,97	0,93
7. Zu Beginn einer Prüfung bin ich so nervös, dass ich oft nicht klar denken kann.	1,90	± 0,96	0,92
8. Die Aussicht, in einem meiner Kurse einen Test schreiben zu müssen, bereitet mir keine Sorgen. ¹	2,66	± 0,96	0,93
9. Ich bin ruhiger in Testsituationen als der durchschnittliche Student. ¹	2,57	± 0,92	0,84
10. Ich habe weniger Schwierigkeiten als der durchschnittliche Student, Lehrbuchinhalte zu lernen. ¹	2,60	± 0,82	0,68
11. Ich bekomme einen Blackout, wenn ich mich unter Druck für eine Antwort im Test entscheiden muss.	1,64	± 0,79	0,62
12. Während einer Prüfung kommt mir öfter der Gedanke, dass ich vielleicht nicht sehr klug sein könnte.	1,59	± 0,78	0,60
13. Mir fallen Tests unter Zeitdruck besonders leicht. ¹	3,13	± 0,77	0,60
14. Während einer Klausur werde ich so nervös, dass ich Dinge vergesse, die ich eigentlich weiß.	2,14	± 0,84	0,71
15. Nach einer Prüfung habe ich das Gefühl, dass ich es eigentlich hätte besser machen können.	2,66	± 0,95	0,90
16. Ich mache mir größere Gedanken darüber, in Tests gut abzuschneiden, als ich eigentlich sollte.	2,57	± 1,02	1,04
17. Bevor ich zu einer Prüfung antrete, fühle ich mich selbstsicher und entspannt. ¹	3,01	± 0,86	0,74
18. Während ich an einer Prüfung teilnehme, fühle ich mich selbstsicher und entspannt. ¹	2,59	± 0,82	0,67
19. Während einer Prüfung habe ich das Gefühl, dass ich mich nicht besonders gut schlage.	2,10	± 0,70	0,48
20. Wenn ich an einer anspruchsvollen Prüfung teilnehme, fühle ich mich schon als Versager, bevor ich überhaupt angefangen habe.	1,46	± 0,74	0,55
21. Wenn ich im Test auf eine unerwartete Frage stoße, fühle ich mich eher herausgefordert, als panisch. ¹	2,68	± 0,86	0,74
22. Ich bin kein Mensch für Prüfungen, weil meine Prüfungsleistung nicht zeigt, wie viel ich eigentlich über das Thema weiß.	2,22	± 1,01	1,01
23. Ich bin nicht gut im Schreiben von Prüfungen.	1,81	± 0,81	0,66
24. Wenn ich mein Prüfungsexemplar bekomme, dauert es erstmal eine Weile, bis ich mich soweit beruhigt habe, dass ich klar denken kann.	1,47	± 0,80	0,64
25. Ich fühle mich ziemlich unter Druck, gute Noten zu schreiben.	2,38	± 1,06	1,11
26. Ich schneide nicht besonders gut in Prüfungen ab.	2,02	± 0,87	0,76
27. Ich mache oft Leichtsinnsfehler in Prüfungen, weil ich so nervös bin.	2,32	± 0,90	0,81

Anmerkung. Invers gepolte Items wurden zu Beginn der Analysen umkodiert und sind mit ¹ gekennzeichnet.

Formulierungen ergab, wurden alle drei Items beibehalten.

Trennschärfe

Bei der Berechnung der Trennschärfeindizes fiel eine ungenügende Passung von Item 25 („*Ich fühle mich ziemlich unter Druck, gute Noten zu schreiben.*“) zur übrigen Skala auf, welche sich nach inhaltlicher Überprüfung bestätigte (siehe Tabelle 2). Item 25 war bereits von Furlan et al. [48] als inhaltlich unzureichend identifiziert worden und war daraufhin für alle nachfolgenden Analysen von der Skala entfernt worden. Nach der Entfernung von Item 25 lagen die Trennschärfeindizes der übrigen Items zwischen $r_{it}=0,37$ (Item 3: „*Mir fällt es weniger schwer als dem durchschnittlichen Studenten, Prüfungsanweisungen präzise aufzufassen.*“) und $r_{it}=0,68$ (Item 14: „*Während einer Klausur werde ich so nervös, dass ich Dinge vergesse, die ich eigentlich weiß.*“) und somit oberhalb des zuvor festgelegten Werts.

Skalenreliabilität

Cronbach's alpha der Skala betrug 0,92.

Kriteriumsvalidität

G-CTAS-Summenwerte der weiblichen Teilnehmer lagen im Durchschnitt bei 59,76 (SD=12,67), G-CTAS-Summenwerte der männlichen Teilnehmer im Durchschnitt bei 54,22 (SD=13,59). Die Werte unterschieden sich im T-Test signifikant voneinander (mittlere Differenz: 5,54, $T=3,46$, $df:289$, $p<0,01$). G-CTAS-Summenwerte korrelierten signifikant ($p<0,001$) mit Noten im schriftlichen ($r_p=0,44$) und mündlichen Physikum ($r_p=0,40$).

Diskussion und Schlussfolgerungen

In der vorliegenden Untersuchung wurde eine deutsche Version von CTAS erstellt sowie an einer Kohorte Humanmedizinstudierender statistisch überprüft und validiert. Die Untersuchungen zeigten, dass 26 Items der deutschen Version akzeptable Schwierigkeit, Varianz und Trennschärfe besitzen. G-CTAS verfügt außerdem über eine hohe interne Konsistenz und Konstruktvalidität. Die Stichprobe ist durch ihre Größe und Geschlechterverteilung als repräsentativ für eine deutsche medizinische Fakultät zu bewerten. Es handelte sich um Medizinstudierende ähnlichen Alters im gleichen Studienabschnitt und somit um ein stark selektiertes Studienkollektiv, weshalb sich die Ergebnisse weniger zur Erstellung von Normwerten oder zur Definition von Schweregraden kognitiver Prüfungsangst eignen. Allerdings qualifiziert dies die Skala besonders für weitere Studien zu kognitiver Prüfungsangst unter Medizinstudierenden. Die Itemformulierungen ermöglichen die Anwendung der Skala für Untersuchungen zur kognitiven Prüfungsangst im Studium ohne Einschränkung in Bezug auf den Studiengang. Wei-

Tabelle 2: Dargestellt sind die Trennschärfeindizes aller 27 Items [36] sowie nach Entfernung von Item 25 (r_{it} [ohne Item 25]). Angegeben ist auch Cronbach's alpha der jeweiligen Skala.

Item	r_{it}	r_{it} (ohne Item 25)
1.	,512	,508
2.	,438	,416
3. ¹	,352	,365
4.	,649	,659
5. ¹	,611	,602
6.	,503	,498
7.	,645	,650
8. ¹	,531	,531
9. ¹	,669	,674
10. ¹	,409	,431
11.	,607	,614
12.	,550	,547
13. ¹	,429	,437
14.	,668	,676
15.	,580	,573
16.	,407	,366
17. ¹	,585	,578
18. ¹	,636	,645
19.	,455	,457
20.	,600	,603
21. ¹	,550	,560
22.	,631	,636
23.	,462	,469
24.	,507	,505
25.	,243	
26.	,451	,469
27.	,660	,665
Cronbach's alpha	0,921	0,923

Anmerkung: Invers gepolte Items wurden zu Beginn der Analysen umkodiert und sind mit ¹ gekennzeichnet.

tere Studien an deutschen Fakultäten können somit ohne Adaptation der Skala folgen.

Da wiederholt ein Methodeneffekt durch die Verwendung invers gepolter Items beobachtet werden konnte, existieren bereits mehrere überarbeitete Versionen der Skala

[42], [43], [44]. Im Rahmen dieser Arbeit wurde für die Übersetzung der ursprünglichen Skala entschieden, da sich durch Entfernung aller invers gepolten Items die von Cassidy & Finch vorgeschlagene Kurzform der Skala [42] einfach generieren lässt.

Die Auswahl der Items erfolgte nicht auf Basis eines Kennwerts, sondern in Hinblick auf Schwierigkeit, Varianz und Trennschärfe. Die inhaltliche Formulierung spielte beim Ausschluss eines Items eine entscheidende Rolle. Laut Bühner [51] sollte der Inhalt oberste Priorität bei der Itemselektion haben, in bisherigen Analysen mit CTAS und seiner Übersetzungen wurde auf diese jedoch kaum Bezug genommen. Dennoch deckt sich das Ergebnis der Itemanalyse mit den Beobachtungen von Furlan et al. [48], welche ebenfalls zum Ausschluss von Item 25 führten. Vermutlich misst Item 25 primär Leistungsdruck und weniger kognitive Prüfungsangst.

Auch Cronbach's alpha von G-CTAS war vergleichbar mit Voruntersuchungen [48]. Der Wert spricht für eine hohe interne Konsistenz und somit für eine homogene Skala. Eine ausführliche Analyse der Faktorenstruktur konnte nicht angeschlossen werden, da die Stichprobengröße nicht ausreichend für eine explorative und anschließende konfirmatorische Faktorenanalyse war [51]. Diese sollte als nächster Schritt durchgeführt werden, da eine genaue Kenntnis der Skalenstruktur nicht nur die Grundlage für die Bildung von Normwerten darstellt, sondern außerdem zu einem genaueren Verständnis von Prüfungsangstursachen und -auswirkungen beitragen könnte.

Unterschiede zwischen weiblichen und männlichen Teilnehmern sowie signifikante, positive Korrelationen mit Prüfungsleistungen konnten die Kriteriumsvalidität der deutschen Skala belegen. Das Studiendesign ging nicht weiter auf geschlechtsspezifische Unterschiede kognitiver Prüfungsangst ein, weshalb deren Ursachen weiter unklar bleiben. Dagegen ermöglichte das Studiendesign einen Vergleich der Zusammenhänge von kognitiver Prüfungsangst mit schriftlichen und mündlichen Prüfungen. Hier konnte ein signifikanter Zusammenhang mit beiden Prüfungsmodalitäten auf ähnlich hohem Korrelationsniveau beobachtet werden.

Dass kognitive Prüfungsangst mit schriftlichen als auch mündlichen Prüfungsleistungen in Zusammenhang steht, verdeutlicht einmal mehr die möglichen Ausmaße für Betroffene. Zusammenhänge mit der Leistung in einer relevanten Prüfung wie dem Staatsexamen heben hervor, wie entscheidend Prüfungsangst in Bezug auf den Studienerfolg sein kann. Das Finden von Prüfungsangstursachen sowie von Behandlungsansätzen ist nicht nur aus Sicht der Betroffenen wichtig, sondern sollte auch für Universitäten eine hohe Priorität einnehmen. Medizinstudierende stellen ein Risikokollektiv bezüglich psychischer Belastung und Prüfungsangst dar und sollten dahingehend unterstützt werden. Ein genaueres Verständnis der Dynamiken zwischen Stress, Prüfungen und Prüfungsangst sind wesentlich, um gefährdete Personen identifizieren und frühzeitig in diese negativen Entwicklungen eingreifen zu können. Die Grundlage hierfür stellt ein geeignetes und möglichst exaktes Messinstrument kogni-

tiver Prüfungsangst dar. Mit G-CTAS ist dieses nun auch im deutschsprachigen Raum verfügbar.

Danksagung

Unser Dank gilt Frau Dr. Amanda Tufman, Frau Sarah Garcia, Frau Franziska Enders und Frau Miruh Lee für die Unterstützung bei der Übersetzung der Skala.

Interessenkonflikt

Die Autor*innen erklären, dass sie keinen Interessenkonflikt im Zusammenhang mit diesem Artikel haben.

Literatur

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). 5th Edition. Washington, DC: American Psychiatric Association; 2013.
2. Deutsches Institut für Medizinische Dokumentation und Information (DIMDI). CD-10-GM Version 2019, Systematisches Verzeichnis, Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme, 10. Revision. Köln: Deutsches Institut für Medizinische Dokumentation und Information (DIMDI); 2018.
3. Dachs L. Prüfungsängste. In: Schneider S, Margraf J, editors. Lehrbuch der Verhaltenstherapie, Band 3: Psychologische Therapie bei Indikationen im Kindes- und Jugendalter. Berlin, Heidelberg: Springer Berlin Heidelberg; 2019. p.569-586. DOI: 10.1007/978-3-662-57369-3_32
4. Liebert R, Morris L. Cognitive and emotional components of test anxiety: A distinction and some initial data. *Psychol Report*. 1967;20:975-978. DOI: 10.2466/pr0.1967.20.3.975
5. Steinmayr R, Crede J, McElvany N, Wirthwein L. Subjective Well-Being, Test Anxiety, Academic Achievement: Testing for Reciprocal Effects. *Front Psychol*. 2016;6:1994. DOI: 10.3389/fpsyg.2015.01994
6. Pixner S, Kaufmann L. Prüfungsangst, Schulleistung und Lebensqualität bei Schülern. *Lern Lernstörung*. 2013;2(2):111-124. DOI: 10.1024/2235-0977/a000034
7. Schaefer A, Matthes H, Pfitzer G, Köhle K. Mental health and performance of medical students with high and low test anxiety. *Psychother Psychosom Med Psychol*. 2007;57(7):289-297. DOI: 10.1055/s-2006-951974
8. Akinsola EF, Nwajei AD. Test anxiety, depression and academic performance: assessment and management using relaxation and cognitive restructuring techniques. *Psychol*. 2013;4(06):18. DOI: 10.4236/psych.2013.46A1003
9. Morris LW, Liebert RM. Effects of anxiety on timed and untimed intelligence tests: Another look. *J Cons Clin Psychol*. 1969;33(2):240-244. DOI: 10.1037/h0027164
10. Morris LW, Davis MA, Hutchings CH. Cognitive and emotional components of anxiety: Literature review and a revised worry-emotionality scale. *J Educ Psychol*. 1981;74(4):541-555. DOI: 10.1037/0022-0663.73.4.541
11. Sarason IG. Stress, anxiety, and cognitive interference: Reactions to tests. *J Person Soc Psychol*. 1984;46(4):929-938. DOI: 10.1037/0022-3514.46.4.929

12. Deffenbacher JL, Hazaleus SL. Cognitive, emotional, and physiological components of Test Anxiety. *Cogn Ther Res.* 1985;9(2):169-180. DOI: 10.1007/BF01204848
13. Hembree R. Correlates, causes, and treatment of test anxiety. *Rev Educ Res.* 1988;58:47-77. DOI: 10.3102/00346543058001047
14. Thomas CL, Cassady JC, Heller ML. The influence of emotional intelligence, cognitive test anxiety, and coping strategies on undergraduate academic performance. *Learn Indiv Diff.* 2017;55:40-48. DOI: 10.1016/j.lindif.2017.03.001
15. Putwain DW. Test anxiety and GCSE performance: The effect of gender and socio-economic background. *Educ Psychol Pract.* 2008;24(4):319-334. DOI: 10.1080/02667360802488765
16. Spielberger CD, Anton WD, Bedell J. The nature and treatment of test anxiety. In: Zuckerman M, Spielberger CD, editors. *Emotions and Anxiety (PLE: Emotion)*. London: Psychology Press; 2015. p.317-344.
17. Lowe PA. Should Test Anxiety be Measured Differently for Males and Females? Examination of Measurement Bias Across Gender on Measures of Test Anxiety for Middle and High School, and College Students. *J Psychoeduc Ass.* 2015;33(3):238-246. DOI: 10.1177/0734282914549428
18. Zeidner M. Does test anxiety bias scholastic aptitude test performance by gender and sociocultural group? *J Pers Assess.* 1990;55(1-2):145-160. DOI: 10.1080/00223891.1990.9674054
19. Chapell MS, Blanding ZB, Silverstein ME, Takahashi M, Newman B, Gubi A, Mccann N. Test Anxiety and Academic Performance in Undergraduate and Graduate Students. *J Educ Psychol.* 2005;97(2):268-274. DOI: 10.1037/0022-0663.97.2.268
20. Nilofer Farooqi Y, Ghani R, D. Spielberger C. Gender Differences in Test Anxiety and Academic Performance of Medical Students. *Intern J Psychol Behav Sci.* 2012;2(2):38-43. DOI: 10.5923/j.ijpbs.20120202.06
21. Powell DH. Behavioral treatment of debilitating test anxiety among medical students. *J Clin Psychol.* 2004;60(8):853-865. DOI: 10.1002/jclp.20043
22. Dyrbye LN, Thomas MR, Shanafelt TD. Systematic review of depression, anxiety, and other indicators of psychological distress among U.S. and Canadian medical students. *Acad Med.* 2006;81(4):354-373. DOI: 10.1097/00001888-200604000-00009
23. Dahlin M, Joneborg N, Runeson B. Stress and depression among medical students: A cross-sectional study. *Med Educ.* 2005;39(6):594-604. DOI: 10.1111/j.1365-2929.2005.02176.x
24. Iqbal S, Gupta S, Venkatarao E. Stress, anxiety & depression among medical undergraduate students & their socio-demographic correlates. *Indian J Med Res.* 2015;141(3):354. DOI: 10.4103/0971-5916.156571
25. Burger PH, Tektas OY, Paulsen F, Scholz M. Vom Studienstart bis zum ersten Staatsexamen - Zunahme von Depressivität bei gleichzeitigem Verlust des Kohärenzgefühls und der psychischen Lebensqualität in höheren Semestern Humanmedizin. *Psychother Psych Med.* 2014;64(08):322-327. DOI: 10.1055/s-0034-1374593
26. Moffat KJ, McConnachie A, Ross S, Morrison JM. First year medical student stress and coping in a problem-based learning medical curriculum. *Med Educ.* 2004;38(5):482-491. DOI: 10.1046/j.1365-2929.2004.01814.x
27. Voltmer E, Kötter T, Westermann J. Prävention: Gesund durchs Medizinstudium. *Dtsch Arztebl.* 2015;112(35-36):A-1414/B-1192/c-1164.
28. Dyrbye L, Shanafelt T. A narrative review on burnout experienced by medical students and residents. *Med Educ.* 2016;50(1):132-149. DOI: 10.1111/medu.12927
29. Panagioti M, Geraghty K, Johnson J, Zhou A, Panagopoulou E, Chew-Graham C, Peters D, Hodkinson A, Riley R, Esmail A. Association Between Physician Burnout and Patient Safety, Professionalism, and Patient Satisfaction: A Systematic Review and Meta-analysis. *JAMA Intern Med.* 2018;178(10):1317-1330. DOI: 10.1001/jamainternmed.2018.3713
30. Herbst U, Voeth M, Eidhoff AT, Müller M, Stief S. *Studierendenstress in Deutschland - eine empirische Untersuchung*. Berlin: AOK-Bundesverband; 2016.
31. Rost D, Schermer F. *Differentielles Leistungsangst-Inventar [Differential Achievement Anxiety Inventory]*. Frankfurt: Pearson; 2007.
32. Spielberger CD. Test anxiety inventory. *Psychol.* 1980;31:503-509. DOI: 10.1146/annurev.ps.31.020180.002443
33. Hodapp V. Das Prüfungsängstlichkeitsinventar TAI-G: Eine erweiterte und modifizierte Version mit vier Komponenten [The Test Anxiety Inventory TAI-G: An expanded and modified version with four components]. *Z Päd Psychol.* 1991;5(2):121-130.
34. Hodapp V. The TAI-G: A multidimensional approach to the assessment of test anxiety. In: Schwarzer C, Zeidner M, editors. *Stress, anxiety, and coping in academic settings*. Tübingen: Francke; 1996. p.95-130.
35. Hodapp V, Rohrmann S, Ringeisen T. *Prüfungsangstfragebogen (PAF) [German Test Anxiety Questionnaire]*. Göttingen: Hogrefe; 2011.
36. Cassady JC, Johnson RE. Cognitive Test Anxiety and Academic Performance. *Contemp Educ Psychol.* 2002;27(2):270-295. DOI: 10.1006/ceps.2001.1094
37. Ramirez G, Beilock SL. Writing About Testing Worries Boosts Exam Performance in the Classroom. *Science.* 2011;331:211-213. DOI: 10.1126/science.1199427
38. Cassady JC. The impact of cognitive test anxiety on text comprehension and recall in the absence of external evaluative pressure. *Appl Cogn Psychol.* 2004;18(3):311-325. DOI: 10.1002/acp.968
39. Cassady JC. The stability of undergraduate students' cognitive test anxiety levels. *Pract Ass ResEval.* 2001;7(20):1-5.
40. Tsianos N, Lekkas Z, Germanakos P, Mourlas C, Samaras G. An experimental assessment of the use of cognitive and affective factors in adaptive educational hypermedia. *IEEE Trans Learn Technol.* 2009;2(3):249-58. DOI: 10.1109/TLT.2009.29
41. Kapetanaki EM, editor. *Test anxiety in education: Policies and practices*. Presentation at the 31st World Conference on Stress and Anxiety Research. August 4-6; Galway, Ireland. 2010.
42. Cassady JC, Finch WH. Confirming the Factor Structure of the Cognitive Test Anxiety Scale: Comparing the Utility of Three Solutions. *Educ Ass.* 2014;19(3):229-242. DOI: 10.1080/10627197.2014.934604
43. Cassady JC, Finch WH. Using factor mixture modeling to identify dimensions of cognitive test anxiety. *Learn Individ Diff.* 2015;41:14-20. DOI: 10.1016/j.lindif.2015.06.002
44. Thomas CL, Cassady JC, Finch WH. Identifying Severity Standards on the Cognitive Test Anxiety Scale: Cut Score Determination Using Latent Class and Cluster Analysis. *J Psychoeduc Ass.* 2018;36(5):492-508. DOI: 10.1177/0734282916686004
45. Cassady JC. The influence of cognitive test anxiety across the learning-testing cycle. *Learn Instruct.* 2004;14(6):569-592. DOI: 10.1016/j.learninstruc.2004.09.002

46. Zheng Y. Chinese University Students' Motivation, Anxiety, Global Awareness, Linguistic Confidence, and English Test Performance: A Correlational and Causal Investigation. 2010. Zugänglich unter/available from: <https://qspace.library.queensu.ca/handle/1974/5378>
47. Cassady JC, Mohammed A, Mathieu L. Cross-Cultural Differences in Test Perceptions. *J Cross Cult Psychol.* 2004;35(6):713-718. DOI: 10.1177/0022022104270113
48. Furlan LA, Cassady JC, Pérez ER. Adapting the Cognitive Test Anxiety Scale for use with Argentinean University Students. *Intern J Test.* 2009;9(1):3-19. DOI: 10.1080/15305050902733448
49. Baghaei P, Cassady J. Validation of the Persian Translation of the Cognitive Test Anxiety Scale. *SAGE Open.* 2014;4(4):215824401455511. DOI: 10.1177/2158244014555113
50. Bozkurt S, Ekitli GB, Thomas CL, Cassady JC. Validation of the Turkish Version of the Cognitive Test Anxiety Scale-Revised. *SAGE Open.* 2017;7(1):215824401666954. DOI: 10.1177/2158244016669549
51. Bühner M. Einführung in die Test- und Fragebogenkonstruktion. München: Pearson Studium; 2011. p.640.
52. Rotthoff T, Ostapczuk MS, De Bruin J, Decking U, Schneider M, Ritz-Timme S. Assessing the learning environment of a faculty: psychometric validation of the German version of the Dundee Ready Education Environment Measure with students and teachers. *Med Teach.* 2011;33(11):e624-36. DOI: 10.3109/0142159X.2011.610841
53. Bortz J, Döring N. Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler. Heidelberg: Springer; 2006. DOI: 10.1007/978-3-540-33306-7

Korrespondenzadresse:

Alexandra Stefan
Ludwig-Maximilians-Universität München, Zentrale
PJ-Koordination, Pettenkoferstr. 8a, 80336 München,
Deutschland
alexandra.stefan@med.uni-muenchen.de

Bitte zitieren als

Stefan A, Berchtold CM, Angstwurm M. Translation of a scale measuring cognitive test anxiety (G-CTAS) and its psychometric examination among medical students in Germany. *GMS J Med Educ.* 2020;37(5):Doc50. DOI: 10.3205/zma001343, URN: urn:nbn:de:0183-zma0013436

Artikel online frei zugänglich unter

<https://www.egms.de/en/journals/zma/2020-37/zma001343.shtml>

Eingereicht: 15.08.2019

Überarbeitet: 15.01.2020

Angenommen: 29.05.2020

Veröffentlicht: 15.09.2020

Copyright

©2020 Stefan et al. Dieser Artikel ist ein Open-Access-Artikel und steht unter den Lizenzbedingungen der Creative Commons Attribution 4.0 License (Namensnennung). Lizenz-Angaben siehe <http://creativecommons.org/licenses/by/4.0/>.