TECHNISCHE UNIVERSITÄT DRESDEN

Faculty of Science Department of Psychology

Institute of Clinical Psychology and Psychotherapy

Technische Universität Dresden, 01062 Dresden

World Journal of Psychiatry

Lian-Sheng Ma, President and Company Editor-in-Chief

Prof. Dr.
Hans-Ulrich Wittchen
Head of Institute

Contact:	Susanne Knappe, PhD.
	Research assistant
Telephone:	+49 351 463-39727
Telefax:	+49 351 463-36984
E-mail:	susanne.knappe@tu-dresden.de

Submission to the World Journal of Psychiatry (ID: 02507159)

2016-07-13

Dear Editors

First of all we would like to thank the reviewer for his/her detailed and thoughtful comments. The manuscript has been revised according to the reviewer's suggestions. In the following, we will describe how we dealt with these suggestions. We proceed along the reviewer's comments.

Reviewer 02445209: Dear authors, I do not have any substantial criticism related to your article. Perhaps, the abbreviation "PAF" is used but not explained in the section "Procedure and Participants". I appreciate the comprehensive research methodology in your article. The reviewer

Response: Many thanks for this complaisant feedback. The abbreviation "PAF" is fully explained now in the section "Procedure and Participants".

> Indication of TA based on the PAF (Prüfungsangstfragebogen, [Test Anxiety Questionnaire], see below) was not required as a necessary condition (i.e., no skip rule when no TA was reported) because Knappe et al. [6] observed that fear of taking tests occurred independently from SA but not vice versa.

Reviewer 02445294: Nice little study with a defined focus. Please clearly describe which scales or subscales have been applied twice - once related to oral and once to written exams. What was the rationale for reporting Spearman's and Pearson's correlation coefficients? Please omit those that are not convenient. Please report correlation coefficients with two decimals (1st paragraph of results). I would like to read the exact p values for each of your test scores and not the inconsistent presentation of $p < .05$ or $> .05$ versus $p = .007$. What was the rationale for not using multivariate statistics (e.g., ANOVA models)?

Response: We greatly appreciate the reviewer's feedback and have modified the manuscript as follows:

_a)_ Please clearly describe which scales or subscales have been applied twice - once

related to oral and once to written exams.

Response: Some scales were presented twice to respondents. For social anxiety assessment, the Liebowitz Social Anxiety Scale was modified, and this modification has been added to the text as well.

> **Social anxiety.** The 24-item German self-report Version of the "Liebowitz Social Anxiety Scale" (LSAS)[12] was administered. Item 17 ("taking a test") was splitted for detailed assessment of anxiety and avoidance during written or oral exam situations (namely: taking a written test, taking an oral examination). (…)
> The PAF, TAB TAC, BSPS-G und CDS-9 were administered twice for separate assessments in oral and written TA situations.

b) What was the rationale for reporting Spearman's and Pearson's correlation coefficients? Please omit those that are not convenient.

Response: The rational was added to the statistical analyses section. Both parameters are presented to the reader for the sake of comprehensiveness.

> Pearson correlations were additionally computed since they are necessary for the computation of Steiger´s Z. The correlations were similar and therefore the Pearson correlations were compared within a sample using Steiger´s Z.

c) Please report correlation coefficients with two decimals (1st paragraph of results). I would like to read the exact p values for each of your test scores and not the inconsistent presentation of p < .05 or > .05 versus p = .007.

Response: In line with the reviewer's request, correlation coefficients were cut to two decimals. Please note that p-values were reported as decimal fractions with 3 digits when results are significant, and as p > .05 when results were non-significant according to international conventions and in line with the APA 6th style. After careful checks, two typos were corrected in the results section.

> As expected, SA was unrelated to written TA (Spearman´s r=.17, P>.05; Pearson´s r=.22, P >.05), but positively related to oral TA (Spearman´s r=.343, P <.001; Pearson´s r=.38, P <.001). Correlations of SA with written vs. oral TA however did not differ (Steiger´s Z=-1.18, P >.05).
> (…)
> SA, based on the LSAS, was positively related to safety behaviors (Spearman´s r=.64, P <.001; Pearson´s r=.70, P <.001). Both forms of TA were moderately associated with safety behaviors (written: Spearman´s r=.31, P <.001; Pearson´s r=.28 , P <.001; oral: Spearman´s r=.42, P <.001; Pearson´s r=.453, P <.001); these correlations were however lower than the correlation between LSAS and safety behaviors and did not differ (written TA: Steiger´s Z=3.22, P <.001; oral TA:

Steiger´s Z=2.23, P <.005; written compared to oral TA: Steiger´s Z=-0.93, P >.05).

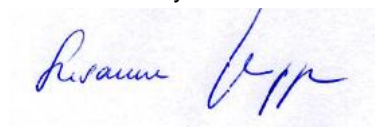e) What was the rationale for not using multivariate statistics (e.g., ANOVA models)?

Response: Given the limited sample size and the absence of normally distributed data for some variables, multivariate statistics were not applied. Statistical analyses based on correlation analyses and dependent t-tests for paired samples were sufficient for testing hypotheses and for the exploration of similarities and differences in written and oral exams, as well as of similarities and differences between both forms of test anxiety and social anxiety. For larger samples with independent assessments for written and oral tests, ANOVA models may be more adequate. We added this statement to the discussion section.

Given the limited sample size and the absence of normally distributed data for some variables, multivariate statistics were not applied. Statistical analyses based on correlation analyses and dependent t-tests for paired samples were sufficient for testing hypotheses and for the exploration of similarities and differences in written and oral exams, as well as of similarities and differences between both forms of test anxiety and social anxiety. For larger samples with independent assessments for written and oral tests, ANOVA models may be more adequate.

In addition, edits of the editor have been addressed.
All edits were made using the track changes function.

We are confident that the quality of the paper has improved significantly, and that this article meets the interest of the editorial team and the readers of the World Journal of Psychiatry. We are looking forward to hearing about your decision by soon,

Yours sincerely

Susanne Knappe, Ph.D.