

Instructions for authors, subscriptions and further details:

<http://ijep.hipatiapress.com>

Education Science Students' Statistics Anxiety: Developing and Analyzing a Scale for Measuring their Worry, Avoidance, and Emotionality Cognitions

Günter Faber¹, Heike Drexler¹, Alexander Stappert² & Joana Eichhorn¹

- 1) Leibniz University Hannover, Germany
- 2) University of Vechta, Germany

Date of publication: October 24th, 2018
Edition period: October 2018 – February 2019

To cite this article: Faber, G., Drexler, H., Stappert, A., & Eichhorn, J. (2018). Education science students' statistics anxiety: Developing and analyzing a scale for measuring their worry, avoidance, and emotionality cognitions. *International Journal of Educational Psychology*, 7(3), 248-285. doi: [10.17583/ijep.2018.2872](https://doi.org/10.17583/ijep.2018.2872)

To link this article: <http://dx.doi.org/10.17583/ijep.2018.2872>

PLEASE SCROLL DOWN FOR ARTICLE

The terms and conditions of use are related to the Open Journal System and to [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/) (CC-BY).

Education Science Students’ Statistics Anxiety: Developing and Analyzing a Scale for Measuring Their Worry, Avoidance, and Emotionality Cognitions

Günter Faber, Heike Drexler, Joana Eichhorn
Leibniz University Hannover

Alexander Stappert
University of Vechta

Abstract

Current instruments for assessing university students’ statistics anxiety prevalingly emphasize the affective construct component. In order to unfold the construct in a more exhaustive and differentiated manner, a scale for measuring university students’ worry, avoidance, and emotionality cognitions was developed. In two samples of education science majors the present pilot study aimed at analyzing the scale’s psychometric properties and at gaining preliminary validation results. Principal component analyses led to the formation of a unidimensional scale which appeared to be sufficiently reliable. Its relations to domain-specific self-belief and background variables turned out as theoretically expected – thus, for the time being the scale should claim criterion validity.

Keywords: statistics anxiety; worry; avoidance, and emotionality cognitions; scale development; psychometric analysis; preliminary validation results.

La Ansiedad con la Estadística en Estudiantes de Ciencias de la Educación: Desarrollo y Análisis de una Escala para Medir su Preocupación, Evasión y Cogniciones Emocionales

Günter Faber, Heike Drexler, Joana Eichhorn
Leibniz University Hannover

Alexander Stappert
University of Vechta

Resumen

Los instrumentos corrientes para evaluar las estadísticas sobre la ansiedad de estudiantes en la universidad enfatizan predominantemente el componente del constructo afectivo. Para revelar el constructo de forma más exhaustiva y diferenciada, se desarrolló una escala para medir la preocupación, evasión y emocionalidad cognitiva de los estudiantes. En dos muestras de carreras de ciencias de la educación, el estudio piloto presente ha tenido por objetivo analizar las propiedades psicométricas de la escala y la obtención de resultados preliminares de validación. Los análisis del componente principal condujeron a la formación de una escala unidimensional que apareció ser suficientemente fiable. Su relación con la seguridad en uno mismo en ámbitos específicos y las variables de fondo resultado siendo la teóricamente esperada - así, por el momento la escala debería reclamar la validez de criterio.

Palabras clave: estadísticas de ansiedad; preocupación; evasión, y emocionalidad cognitiva; desarrollo de una escala; análisis psicométrico; resultados preliminares de validación.

Quite a number of undergraduate and graduate students of the social sciences, education, psychology and business appear to struggle with statistics (Onwuegbuzie & Wilson, 2003; Zeidner, 1991). When dealing with the requirements of quantitative method courses which commonly are compulsory for earning their degree, these students mostly suffer from strong failure expectations and frequently experience feelings of apprehension and personal threat. As a result, they are at risk to develop and maintain a heightened level of anxiety in the face of statistical analyses – in particular, when being confronted with statistical tasks of data gathering, processing, and interpreting (Cruise, Cash, & Bolton, 1985).

Structurally, students' emerging statistics anxiety has to be considered a multidimensional construct reflecting the complex interplay of several cognitive, motivational, and physiological components (Rost & Schermer, 1989). Based on empirical findings of test anxiety research it can be defined as a domain-specific form of performance or evaluation anxiety which manifests as repeatedly occurring worry cognitions, task-irrelevant and interfering thoughts, marked states of emotional tension and physiological arousal (Zeidner, 1991). The worry component of test anxiety refers to the students' mental anticipation of failure and its negative consequences, whereas the emotionality component refers to their feelings of tenseness, nervousness or distress, and the physiological component refers to their perceptions of bodily symptoms. Over the past decades, there has been ample evidence for these components being distinguishable but mutually reinforcing (Deffenbacher, 1980; Hodapp & Benson, 1997; Kieffer & Reese, 2009; Sarason, 1984). In most cases they could be demonstrated to negatively affect the students' learning process and achievement outcomes. However, the worry component generally turned out to most strongly predict academic performance or test results (Cassady & Johnson, 2002; von der Embse, Jester, Roy, & Post, 2018). This debilitating effect of worry cognitions appeared to be mainly caused by their strongly biased and task-irrelevant mode of information processing (Schwarzer, 1996; Zeidner, 1998).

Based on this conclusive body of evidence a theoretically and methodologically sound framework for the statistics anxiety construct

should implicitly consider its cognitive, emotional, and physiological components. In particular, it is essential that it addresses the issue of relevant worry cognitions as they are closely linked to the debilitating effects of statistics anxiety on statistical learning and performance. To date, however, most empirical analyses in the field have emphasized the emotional and physiological components of statistics anxiety (Cruise et al., 1985). They define the construct as feelings of anxiety or as habitual anxiety in the face of statistically loaded situations (Macher, Paechter, Papousek, & Ruggeri, 2012; Onwuegbuzie & Wilson, 2003). Admittedly, the measurement items used in these empirical analyses cover a wide range of statistical tasks that students typically encounter in everyday study or the context of their course. These items can be empirically assigned to various situation- or task-specific dimensions. Thus, for instance, factor analyses of the task- and course-related items of Zeidner's (1991) Statistics Anxiety Inventory led to the development of a content- and a test-specific subscale. Likewise, factor analyses of the widely used Statistics Anxiety Rating Scale (Cruise et al., 1985) and the conceptually related Statistical Anxiety Scale (Vigil-Colet, Lorenzo-Seva, & Condon, 2008) provided separate subcomponents concerning the students' interpretation and test anxiety, their fear of asking for help and fear of statistics teachers. Similarly, the Statistics Anxiety Measure (Earp, 2007) and the Statistics Comprehensive Anxiety Response Evaluation (Griffith et al., 2014) revealed some distinct task- or situation-specific subcomponents referring to statistically relevant course requirements and situations. In summary, these approaches to measuring the construct of statistics anxiety undoubtedly represent typically anxiety-evoking task features and test situations in a most elaborate way. However, with the exception of the Statistics Anxiety Measure (Earp, 2007) and the short research scale Hong and Karstensson (2002) used, both of which include single worry items, it is notable that all other instruments fail to integrate students' cognitive, emotional, and physiological anxiety reactions, in particular with regard to the critical worry component.

In contrast, a concurrently operating research line in the test anxiety field had already decomposed the statistics anxiety construct and assessed the students' worry and emotionality responses separately. However, in most cases a composite score including both components was used because of both components' high interrelation (Benson, Bandalos, & Hutchinson, 1994; Finney & Schraw, 2003; González, Rodríguez, Faílde, & Carrera, 2016;

Hong & Carstenson, 2002). That way, albeit merely having total anxiety scores available, the interpretation of students' responses explicitly allowed for a traceable cognitive perspective. As this research line essentially contributes to refine the statistics anxiety construct with respect to its motivationally operating components, its task- or situation-specific references appear less differentiated. That is, in all studies the items for assessing statistics anxiety referred exclusively to the taking of a statistical test or exam. This contextual limitation, hence, should challenge the representativity or content validity of the worry and emotionality measures, because their scores account only for a particular part of the relevant learning setting (Haynes, Richard, & Kubany, 1995).

Moreover, another issue crucial to the conceptualization of the statistics anxiety construct refers to the role of the students' avoidance tendencies. Already in the very first beginning of empirical test anxiety research, Mandler and Sarason (1952) posited anxiety responses to manifest as "implicit attempts at leaving the test situation" (p. 166). Subsequently, empirical findings lent support for this assumption and yielded sound evidence for students' avoidance cognitions being substantially related to their anxiety responses (Blankenstein, Flett, & Watson, 1992; Galassi, Frierson, & Sharer, 1981; Hagtvet & Benson, 1997; Skaalvik, 1997). In particular, the analyses of Elliot and McGregor (1999), Pekrun, Elliot and Maier (2009), and Putwain and Symes (2012) demonstrated clearly that students with heightened avoidance orientations reported a higher extent of worry cognitions and lower scores on subsequent exam performance. Heretofore, only few conceptualizations of the test anxiety construct had addressed this issue and claimed the students' escape or avoidance cognitions to constitute an essential part of their worries and to represent an important factor to elicit interfering, task-irrelevant thoughts (Pekrun, Goetz, Perry, Kramer, Hochstadt, & Molfenter, 2004; Schwarzer & Quast, 1985). Accordingly, further research should develop appropriate measurements being designed not only to assess the students' worries about threatening failure outcomes, but also to inquire their thoughts to preferably avoid getting involved with threatening tasks or situations. Currently available questionnaires for measuring students' statistics anxiety either do not consider their avoidance cognitions at all (Griffith et al., 2014; Onwuegbuzie

& Wilson, 2003; Vigil-Colet et al., 2008) or include just a single item assessing avoidance behavior (Earp, 2007). Research in the field should refine its conceptualization of students' worry responses and develop instruments that explicitly capture avoidance cognitions with respect to statistically loaded tasks and situations. That way, an important step to refine the substantive and structural stage of construct validation would be done (Benson, 1998).

Approaching refined measurement

To overcome the conceptual limitations of current instruments as a very first attempt a new scale to assess university students' statistics anxiety was developed. Nevertheless, it should adopt the particular strengths of existing instruments. Thus, it was assigned to approach a refined measure of the construct by meeting the following criteria: Its items should (1) specifically take into account the students' anxiety responses in a most differentiable way and, hence, consider their worry and avoidance cognitions as well as their emotional reactions. Thereby, its items should embed the various anxiety reactions (2) into a representative range of statistically loaded task features and course situations the students would typically encounter.

The construction of this scale for measuring university students' "Worry, Avoidance, and Emotionality Cognitions Encountering Statistical Demands" (WAESTA) largely followed the procedure of facet theory using a mapping sentence (Guttman & Greenbaum, 1998). This mapping sentence served as a heuristic device to cover all major facets of the construct and, thus, to achieve sufficient content validity (Edmundson, Koch, & Silverman, 1993). A pool of eligible items was drawn up using this constitutive mapping sentence which included conceptually relevant anxiety components, situational references, and intended response categories (Zeidner, 1998). In particular, each item to represent the statistics anxiety domain was specified with respect to four key facets: a relevant reaction facet referring to the worry, avoidance, and emotionality component, and three contextual facets referring to the (1) outcome in a statistics exam, the (2) individual learning of statistical procedures and handling of statistical demands, and the (3) public mastering of statistical content. Furthermore, an additional range facet defined the response categories to assess the students' perceived magnitude of individual anxiety reactions. As seemingly appropriate response range a

four-point format was decided – in order to avoid artificial complexities in the respondents' decision making but instead to ensure a cognitive-motivationally realistic as well as just manageable number of rating references. This four-facet mapping sentence was used as a cross-classification template to systematically operationalize the statistics anxiety construct as it allowed to operationalize the various elements of facets in a most differentiated manner (Hox, 1997). That way, a final scale version with 17 four-point Likert-type rating items was built (Table 1).

Table 1
WAEFTA items assessing relevant anxiety reactions to statistically loaded task features and course situations

Test Anxiety Component	Course Exam Outcome	Understanding Explanation Application	Oral Task Presentation Explanation
Worry	01, 14, 16	05, 08	10, 11, 12
Avoidance	03	06, 17	04
Emotionality		02, 07, 09, 13	15

Sample item and response range:

I would hardly be able to present a report on statistical research findings adequately.

Does not apply	1	2	3	4	Applies in full
----------------	---	---	---	---	-----------------

All items concerned a mentally imaginable situation the students should easily manage to anticipate. Eight items referred to the students' worries about their potentially expected failure to master the course exam and to cope with several statistical requirements. Four items concerned their cognitions to preferably avoid the statistics course and particular statistical

demands. Five items are related to their emotional tension when being confronted with a certain statistical task (Appendix).

As statistically indicated task requirements for the understanding of course contents, the interpretation of quantitative research results as well as the application of statistical formulas and procedures were considered. Likewise, the oral presentation and explanation of statistical content in the public course situation was included. To warrant conceptual clarity, the avoidance items should neither tap the students' avoidance reactions by suppressing or substituting individually occurring threat cognitions (Williams, 2015), nor should they refer to the students' actual avoiding strategies to cope with disliked or threatening academic events (Onwuegbuzie, 2004). Rather they should operationalize the students' mentally processed avoidance thoughts or even escape illusions before or during statistical task completion – and, thus, might indicate a specific subcomponent of worry cognitions. Moreover, the component of physiological symptoms or bodily tensions was not explicitly addressed. Instead it was thought to be indirectly inferred from the emotionality items. Conceptually, this restriction seemed to be justifiable, since the students' perceived affective state should always reflect their actual physiological arousal (Zeidner, 1998).

Validation framework and objectives

As test anxiety is assumed to be a multifaceted construct (Zeidner, 1998), first of all, the factor structure of the WAESTA scale should be analyzed. As the WAESTA items were theoretically designated to represent each the worry, avoidance, and emotionality component of the statistics anxiety construct, a clear three factor solution appeared to be expectable, at best. However, relevant research findings in the test anxiety field had demonstrated these components being substantially correlated (Deffenbacher, 1980; Cassady & Johnson, 2002; Hodapp & Benson, 1997; Hong & Karstensson, 2002; Sarason, 1984). Furthermore, in certain research contexts dealing with school students' domain- or subject-specific test anxieties, all worry, avoidance, and emotionality items repeatedly loaded on one common anxiety factor (Faber, 1995, 2012b). Therefore, an accurate prediction of the scale's ultimate factor structure seemed difficult. Rather the present study should explore the scale's underlying structure in a most tentative way – and should, thus, take into account three alternatives: a three

factor solution separating the worry, avoidance, and emotionality components, a two factor solution with the worry and avoidance items loading on a first factor and the emotionality items loading on a second factor, and a one factor solution subsuming all items. Presuming the avoidance component to represent a specific worry element, the two factor solution could definitely reveal a reasonable perspective, in particular. With the reservation of this initially performed analysis, the final version of the WAESTA scale should be determined and its psychometric properties examined.

As relevant cognitive-motivational constructs academic competence and control beliefs were assessed (Schunk & Zimmerman, 2006) which are well proven to regulate students' engagement and learning approach in the long term. In particular, they essentially affect the students' anxiety experience. As unfavorable competence beliefs usually come along with increased expectancies of failure, they will provoke a strong sense of personal threat and, thus, lead to an individually heightened level of anxiety. As the students, likewise, are not (or not anymore) able to realize individually feasible perspectives to prevent a certain failure outcome, they will develop reduced control beliefs which all the more strengthen their feelings of threat and anxiety.

There is sound evidence that domain-specific academic competence beliefs or self-concepts substantially predict the individually existing magnitude of test anxiety (Ahmed, Minnaert, Kuyper, & van der Werf, 2012; Goetz, Pekrun, Hall, & Haag, 2006). Correspondingly, in the statistics domain the crucial role of self-concepts had been well established. In most cases, high-anxious students reported a lowered self-concept of own mathematics or statistical competencies (Bandalos et al., 1995; Benson, 1989; González et al., 2016; Macher et al., 2012; Williams, 2014; Zeidner, 1991). Therefore, it should be assumed the WAESTA scale scores to correlate negatively and substantially with the students' mathematics self-concept. As well, to sufficiently clarify the domain-specificity of the WAESTA scale, its relation to the students' verbal self-concept should be concurrently analyzed. According to the multidimensional feature of academic self-beliefs (Green, Martin, & Marsh, 2007), research in the field could consistently demonstrate the students' mathematics anxiety being

substantially related to their performance and motivation in the mathematics but not in the verbal domain (Goetz, Frenzel, Pekrun, Hall, & Lüdtke, 2007; Gogol, Brunner, Preckel, Goetz, & Martin, 2016). From this validation perspective, the WAESTA scale should claim preliminary subject-specificity if its correlation with the verbal self-concept variable would turn out to be distinctly weaker than with the mathematics self-concept variable.

Besides, the students' control beliefs largely manifest as implicit theories or mindsets which may stress an entity view of more or less fixed and unchangeable abilities – or an incremental view of more or less modifiable and changeable abilities (Dweck & Leggett, 1988). They could be demonstrated to significantly affect the students' motivational orientations, learning strategies, and, eventually, their task performance (Blackwell, Trzesniewski, & Dweck, 2007; Burnette, O'Boyle, VanEpps, Pollack, & Finkel, 2013; Cury, DaFonseca, Zahn, & Elliot, 2008). These implicit theories principally might not only concern an individual's cognitive ability but also might emerge in a domain-specific manner and refer to the perceived malleability of certain skills or competencies (Dweck & Molden, 2005). Consequently, they should also play a motivationally crucial role in the students' learning of statistics. With respect to the statistics domain, an entity view of own competencies would diminish or even suspend any control perspective. Unfortunately, previous studies in the field had seldom analyzed the role of implicit theories. If at all, they had referred to the students' general ability beliefs, but not to specific beliefs about statistical competencies (Zonnefeld, 2015) or had only considered the students' beliefs to master statistical demands through strategy use and effortful behavior (Schutz, Drogosz, White, & Distefano, 1998) – thus, reflecting an incremental view of learning approach. However, these learning control beliefs could be demonstrated to significantly predict course grades. Accordingly, against the background of research findings the WAESTA scale scores should be reasonably assumed to correlate positively and substantially with the students' entity view of less or not malleable statistical competence.

Furthermore, as another motivational criterion variable, the students' task values were considered. Conceptually, from an expectancy-value perspective on achievement motivation task values concern the students' perceived importance or adequacy of a certain activity to fulfill their personal needs and to attain their personal goals (Eccles & Wigfield, 2002). In particular,

these task values had been demonstrated to regulate the students' motivational orientations, learning strategies, and academic choices (Wigfield, Hoa, & Klauda, 2009). Task values evidently emerge in a task- or at least domain-specific manner (Gaspard, Häfner, Parrisius, Trautwein, & Nagengast, 2017; Selkirk, Bouchev, & Eccles, 2011). Thus, they should characteristically affect the students' learning and performance in the statistics domain as well. Previous research in the field had primarily analyzed the students' perceived utility or worth of statistical knowledge and competencies as an attitudinal construct (Nolan, Beran, & Hecker, 2012) focusing on the usefulness of statistics in personal and professional contexts (Cruise et al., 1985; Dauphinee, Schau, & Stevens, 1997). The students' ratings of the worth of statistics appeared to positively correlate with their statistical achievement as well as with their learning strategies to a slight extent only (Emmioğlu & Capa-Aydin, 2012). In comparison, the relations of utility perceptions with domain-specific measures of academic self-beliefs were stronger. Students with low competence beliefs valued statistics as less important (Baloğlu, 2002; Chiesi & Primi, 2009; Dauphinee et al., 1997; Vanhoof, Kuppens, Sotos, Verschaffel, & Onghena, 2011). Similarly, students' statistics anxiety was also moderately correlated with their utility ratings – indicating those students suffering from a heightened level of statistics anxiety tendentially perceived statistics to a lesser extent as useful (Baloğlu, 2002; Chew & Dillon, 2014; Nasser, 2004; Papanastasiou & Zembylas, 2008; Papousek et al., 2012). Accordingly, the WAESTA scale scores should be assumed to correlate inversely and substantially with the students' perceived value of statistical competence.

As a relevant background variable to explain the students' statistical self-beliefs and competencies their prior mathematical learning, in particular their latest school grade had been well proven. From the perspective of self-concept development (Marsh & O'Mara, 2008), previous failure experience in mathematics will evidently lead to form low competence beliefs in the statistics domain and, eventually, contribute to strengthening the emergence of domain-specific anxiety responses. In various studies students with poor school grades in mathematics reported a heightened level of statistics anxiety (Beurze, Donders, Zielhuis, de Vegt, & Verbeek, 2013; Birenbaum & Eylath, 1994; Chiesi & Primi, 2010; Lalonde & Gardner, 1993).

Accordingly, the WAESTA scale scores should be assumed to correlate negatively, but low in magnitude with the students' mathematics grade they had last earned at school.

Method

Participants and procedure

In the present study the data of both a construction sample and a validation sample were analyzed. The construction sample consisted of 113 graduate students ($n = 94$ females, $n = 19$ males) from a German university Master's course in educational sciences ($n = 80$) and special education ($n = 33$). They all were enrolled in a compulsory course on empirical research methods. Therefore, the participation rate was sufficiently high at 82 per cent. Seventy-four of the students had already acquired elementary statistical knowledge during their first degree, whereas 39 were required to attend a course in basic descriptive and inferential statistics. Both the subgroup with and without statistical knowledge did not significantly differ with respect to gender (chi-square test, $p > .05$) and age (Mann-Whitney U-test, $p > .05$). Also, there was no significant difference of gender ratio within each subgroup (binomial test, $p > .05$).

The validation sample was thought to scrutinize the findings from the construction sample one year later. It consisted of 87 graduate students from the same Master's courses: educational sciences ($n = 59$) and special education ($n = 28$). The sample was predominantly female ($n = 74$). As with the construction sample all the students were enrolled on a compulsory course on empirical research methods. The participation rate was rather high at 89 per cent. Fifty of the students had acquired statistical knowledge during their first degree whereas 37 had to attend an introductory statistics course. Once again there were no subgroup differences in gender, gender ratio, or age (Stappert, 2017).

In both samples all relevant data concerning the self-belief and background variables under consideration were gathered on the course's first term. For that purpose, a questionnaire including all items to measure the students' self-concept, statistics anxiety, implicit theories, task values, and relevant background information was administered. To prevent a priming

effect of the self-concept items, they were presented at the end of the questionnaire.

Both samples had missing data (5.7% and 7.3%). As they did not produce any systematic pattern in the construction (MCAR test $p = .182$) and in the validation sample (MCAR test $p = .178$), they were treated as “missing completely at random” (Little, 1988). The missing values were estimated by means of the two-step iterative expectation-maximization algorithm (Graham, 2012).

Measures

Students' academic self-concepts in mathematics and language (German) were assessed using nine six-point rating items for each subject. These items referred to the students' most recent learning experiences at school and addressed their competence beliefs with regard to meeting subject-specific demands. The wording of the items was strictly parallel. In the majority, the items originate from well proven instruments (Faber, 2012a; Möller, Streblov, Pohlmann, & Köller, 2006; Rost, Sparfeldt, & Schilling, 2007). For the purpose of this study they were adapted and phrased retrospectively. Sample item: “I tried hard in mathematics/German, but I did not perform very well.” Principal component analysis (with varimax rotation) revealed a two-factor solution allowing for a clear distinction between the subject-specific self-concept facets. Hence, it was possible to build two scales for measuring the subject-specific academic self-concepts. Their reliability was most appropriate for both the mathematics and the language self-concept scale (Table 2). High scale scores indicated the students' competence beliefs being positive. According to the multifaceted feature of the construct, the self-concept variables appeared to lowly correlate in the construction sample ($r = .14, p > .05$) and in the validation sample ($r = -.05, p > .05$).

To assess students' implicit theory of statistical competencies, a short scale with five four-point rating items was administered in the construction sample. As current instruments in the field only allowed for measuring the students' implicit intelligence theory (İlhan & Cetin, 2013; Kookan, Welsh, McCoach, Johnston-Wilder, & Lee, 2016), a new scale was created. Following the recommendations of Hong, Chiu, Dweck, Lin and Wan

(1999), all items tapped an entity view of personal statistical competence. Unfortunately, due to their insufficient item-test correlation ($r_{it} < .22$) two items had been deleted. With an average item intercorrelation of Fisher’s $z' = .44$ and an average item-test correlation of Fisher’s $z' = .52$ the final scale’s reliability appeared to be just acceptable. Sample item: “To work with statistics, you need a talent that I simply do not have.” High scale scores indicated the students to perceive their statistical ability being fixed, hence as less malleable in nature. In the validation sample, a slightly revised scale with four four-point Likert items was used (Stappert, 2017). In view of sample size and item number its reliability appeared to be sufficient (Table 2).

Table 2

Descriptive statistics and reliabilities of the scales for measuring validation variables

	Items	<i>AM</i>	<i>SD</i>	z_S	z_K	α
Latest School Grade in Mathematics						
Construction Sample	1	3.06	0.91	0.76	-0.15	
Validation Sample	1	3.11	1.36	-0.50	-2.28*	
Academic Self-Concept in Mathematics						
Construction Sample	9	29.81	9.81	1.37	-2.04	.93
Validation Sample	9	31.01	11.43	-0.33	-1.84	.94
Academic Self-Concept in Language (German)						
Construction Sample	9	45.52	7.32	-3.61***	-0.01	.92
Validation Sample	9	41.45	8.53	-2.12*	-0.16	.90
Implicit Entity Theory of Statistical Competencies						
Construction Sample	3	5.43	2.02	3.59***	0.13	.70
Validation Sample	4	7.87	2.61	2.15*	-0.32	.81
Negative Instrumental Value of Statistics						
Construction Sample	5	11.06	2.89	0.41	-1.18	.72
Validation Sample	8	16.18	3.95	0.55	-0.10	.75

Table 2 (continued)

Descriptive statistics and reliabilities of the scales for measuring validation variables

Significance: * $p \leq .05$, *** $p \leq .001$

AM = arithmetic mean, SD = standard deviation, z_S = z-standardized skewness, z_K = z-standardized kurtosis, α = internal consistency (Cronbach's coefficient alpha)

The utility value students attributed to statistical competence was measured by means of a short scale. In the case of the construction sample it consisted of five four-point rating items dealing with the perceived utility of statistics for the students' current studies and intended career. Sample item: "Statistics will not play an important role in my future professional life". With an average item intercorrelation of Fisher's $z' = .40$ and an average item-test correlation of Fisher's $z' = .49$ the scale's reliability was just acceptable (Table 2). High scale scores indicated the students to consider statistics being less important. In the validation sample, an extended version of the scale was used. It consisted of eight four-point Likert items. Its reliability was once more just acceptable (Table 2). Here again, high sum scores indicated students to perceive statistics as being less useful for their current studies and later professional development (Eichhorn, 2018).

Finally, as relevant background variable the students' most recent school grade in mathematics was inquired in both samples.

Results

Scale formation

For determining the final version of the WAESTA scale in the construction sample, first of all, descriptive item statistics were calculated. The avoidance item 04 as well as the worry item 11 showed a significant negative skew indicating most students to agree with the statements – in detail they would preferably give a presentation without any statistical content and during a presentation they would strongly hope not being asked statistical questions. Furthermore, as the analysis revealed a significant negative kurtosis score for the items 03, 06, 14, and 15, their distribution appeared to be platykurtic.

Accordingly, the students' relevant item responses denoted a heightened variance or difference among them (Table 3).

In the construction sample, the Kaiser-Meyer-Olkin measure of sampling adequacy and Bartlett's test of sphericity demonstrated the inter-item correlations being appropriately strong (KMO = .878, BTS $p < .001$). Therefore, a principal component analysis (PCA) was conducted in order to clarify the latent scale structure. However, it could not statistically separate the three anxiety components. Neither a varimax nor an oblique rotation procedure could yield any loading pattern to separate the worry, avoidance, and emotionality items in a conceptually proper way. Rather all analyses led to a unidimensional structure (Table 3). This solution revealed sufficiently high factor loadings and could explain 43.59 per cent of extracted variance. Though for further clarification three provisional subscales representing the students' worry, emotionality, and avoidance cognitions were formed and the relations among their sum scores examined. In line with the PCA result, the subscales were strongly correlated – in particular, the worry with the emotionality scale $r = .80$ and with the avoidance scale $r = .72$, the emotionality with the avoidance scale $r = .69$ (all $p < .001$).

Consequently, all WAESTA items could be used to build the scale's final version. For its total score, neither the z-standardized scores of skewness and kurtosis nor the Shapiro Wilk W-test ($W = .988$, $df = 113$, $p = .443$) could evince any significant deviation from the normal distribution assumption. High total scores indicated the students' to report stronger worry, avoidance, and emotionality cognitions. The scale's reliability was estimated in various ways and turned out to be adequate: Its internal consistency (Cronbach's coefficient alpha) amounted to $\alpha = .92$, its split-half reliability (odd-even method using Spearman-Brown correction) to $r_{12} = .89$, and its standard error (based on coefficient alpha) was $s_e = 2.67$.

Table 3

Descriptive statistics, factor loadings and corrected item-test correlations of WAESTA items (WR = worry, AV = avoidance, EM = emotionality): Results from the construction sample

Item	AM	SD	z_S	z_K	a	r_{it}
01 WR	2.37	0.82	1.49	-0.67	.486	.436

Table 3 (continued)

Descriptive statistics, factor loadings and corrected item-test correlations of WAESTA items (WR = worry, AV = avoidance, EM = emotionality): Results from the construction sample

02 EM	2.71	0.94	-1.30	-1.69	.688	.645
03 AV	2.74	1.08	-1.33	-2.61**	.581	.515
04 AV	2.88	0.97	-2.12*	1.65	.719	.664
05 WR	2.51	0.91	-0.50	-1.68	.725	.663
06 AV	2.50	0.97	-0.21	-2.08*	.758	.697
07 EM	2.17	0.96	1.71	-1.78	.697	.620
08 WR	2.24	0.84	1.49	-0.83	.718	.670
09 EM	2.20	0.87	1.59	-1.04	.668	.583
10 WR	2.69	0.93	-1.63	-1.43	.647	.694
11 WR	2.91	0.97	-2.09*	-1.75	.740	.676
12 WR	2.26	0.80	1.52	-0.45	.734	.618
13 EM	2.96	0.93	-1.51	-0.45	.683	.493
14 WR	2.69	1.07	-0.67	-2.18*	.551	.474
15 EM	2.87	0.84	-1.30	-2.77**	.532	.637
16 WR	2.75	0.84	-1.45	-0.83	.689	.442
17 AV	2.07	0.87	1.55	-1.52	.488	.442
Total Sum Score (Items 01-17)						
WAESTA	42.66	10.20	-0.49	-1.34		

Significance: * $p \leq .05$, ** $p \leq .01$

AM = arithmetic mean, SD = standard deviation, z_S = z-standardized skewness, z_K = z-standardized kurtosis, a = factor loading, r_{it} = corrected item-test correlation

In spite of the small number of participants in the validation sample, a principal component analysis (PCA) of the WAESTA items was conducted. As their communalities ranged from $h = .412$ to $h = .660$ and the Kaiser-Meyer-Olkins measure revealed an appropriately high score ($KMO = .901$),

this procedure appeared to be most reasonable (de Winter, Dodou, & Wieringa, 2009; MacCallum, Widaman, Zhang, & Hong, 1999). The results revealed one common factor with considerably high loadings (ranging from $a = .532$ to $a = .804$). Accordingly, the unidimensional scale feature could be fully replicated and explained 50.66 per cent of extracted variance. For its total sum score, z-standardized skewness ($z_s = -0.330$) and kurtosis values ($z_k = -0.190$) did not indicate any significant deviation from the normal distribution assumption. Here again, the scale’s reliability was estimated in various ways and turned out to be adequate: its internal consistency (Cronbach’s coefficient alpha) amounted to $\alpha = .94$, its split-half reliability (odd-even method using Spearman-Brown correction) to $r_{12} = .91$, and its standard error (based on coefficient alpha) was $s_e = 2.45$.

Validation results

As an initial approach to analyze the external validity of the WAESTA scale, its zero-order correlations with the criterion variables under consideration were first analyzed. As the results could demonstrate (Table 4), the WAESTA scores were closely and significantly associated with the students’ mathematics self-concept but not with their language self-concept. This particular finding might be considered to provide preliminary evidence that the WAESTA scale measures rather a domain-specific than a general facet of the students’ test anxiety experience.

Table 4

Relations of WAESTA sum scores with background and self-belief variables (zero-order correlations): Results from the construction and the validation sample

Most Recent School Grade Mathematics	Academic Self-Concept Mathematics	Academic Self-Concept Language	Implicit Entity Theory Competence	Negative Instrumental Value
Construction Sample				
-.31***	-.43***	.09	.62***	.49***
Validation Sample				
-.22*	-.38***	-.08	.80***	.32***

Significance: * $p \leq .05$, *** $p \leq .001$

Furthermore, the scale scores were most strongly correlated with the students' entity views of own statistical competence. A heightened level of statistics anxiety came along with a deep understanding of own statistical competencies being less or even not malleable in nature. With the negative instrumental value of statistics, the WAESTA sum score correlated moderately positive. Students reporting a higher level of statistics anxiety tendentially perceived statistical competencies as less important. Finally, the relation between the WAESTA score and the most recent school grade in mathematics appeared to be positive and significant, though low in magnitude. Hence, students with a heightened level of statistics anxiety had been less successful in the mastery of mathematical demands at school.

To get more differentiated validation results, a series of regression analyses with the WAESTA scale as dependent variable were computed for both samples. As this procedure allowed for controlling the covariations among all predictor variables with respect to their empirical overlap and multicollinearity, it should help to unravel the complexity of construct relations. In particular, a sequence of regression models including an advancing number of predictor variables was consecutively tested (Table 5). In both samples the standardized residuals of WAESTA sum scores did not violate the normal distribution assumption. In each case, the Shapiro Wilk W-test could demonstrate the standardized residuals being normally distributed (construction sample: $W = 986$, $df = 113$, $p = .308$, validation sample: $W = 979$, $df = 87$, $p = .182$).

The results for regression model A clearly demonstrated the mathematics self-concept to explain the most part of anxiety variance. However, adding the entity beliefs to the regression equation in model B and C, the predictive power of the students' mathematics self-concept was reduced to a minimal and insignificant extent. Instead, the students' entity beliefs largely contributed to the WAESTA sum score. As the mathematics self-concept and the entity belief variable in both samples were substantially correlated ($r = -.51$ in the construction sample and $r = -.45$ in the validation sample), but the entity belief variable in both samples was more strongly related to the anxiety variable ($r = .63$ in the construction sample and $r = .80$ in the validation sample) – the massive decline in the self-concepts' beta weight must be seen as a result of multicollinearity. This predictive pattern occurred

in both samples. Moreover, only in the construction sample the students' negative value of statistics substantially and independently explained additional variance in the WAESTA sum score. The difference between samples might be due to the fact, that the methods for assessing the value variable were not comparably formatted. Apart from this, all regression analyses demonstrated the students' statistics anxiety to be essentially and most closely predicted by their control beliefs – as reflecting their perceived malleability of individual competencies in the statistics domain.

Table 5.

Multiple regression of WAESTA sum scores on background and self-belief variables (standardized beta weights and squared multiple regression coefficients): Results from the construction and the validation sample

Model	Most Recent School Grade Mathematics	Academic Self-Concept Mathematics	Implicit Entity Theory Competence	Negative Instrumental Value	R ²
Construction Sample					
A	.027	-.452***			.154
Validation Sample					
A	.184	-.533***			.182
Construction Sample					
B	-.070	-.090	.545***		.396
Validation Sample					
B	.031	-.089	.761***		.620
Construction Sample					
C	-.093	-.045	.452***	.286***	.465
Validation Sample					
C	.024	-.083	.789***	-.062	.614

R² = adjusted multiple regression coefficient squared

Significance: * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$

Finally, for further clarification of the WAESTA scores, their mean differences between both the educational science and the special education students were analyzed. As the comparison groups were small and unequal

in size the nonparametric Mann-Whitney U-test for independent samples was used. In the construction sample a significantly higher level of statistics anxiety in the special education subgroup could be found ($Z = -2.314$, $p = .021$, effect size $r = 0.22$). In the validation sample, the level of statistics anxiety did not significantly differ between educational science and special education students ($Z = -0.374$, $p > .05$). However, with the small size of the validation sample in mind, this finding should be considered cautiously.

Discussion

The present study should examine the internal and external validity as well as the psychometric properties of the newly developed WAESTA scale for measuring educational science students' worry, avoidance, and emotionality cognitions in the domain of statistics learning. Conceptually, this measurement approach should integrate both the strengths of a more situation- and a more reaction-focused research line in the field. As a substantive result, this scale could be demonstrated to represent the construct in a unidimensional manner. The final scale version included all items as initially administered in both samples. Its internal consistency was most sufficient. Furthermore, its relations with self-belief and attainment variables most widely turned out as theoretically predicted. Specifically, total WAESTA score was more strongly correlated with the students' mathematics than with their language self-concept – and, thus, the scale could claim domain-specific validity. These findings correspondingly held for both the construction and the validation sample. For the time being, the WAESTA scale can be considered internally and externally valid as well as having adequate psychometric properties. Nevertheless, some results definitely deserve further attention.

In particular, the scale's underlying structure consistently appeared to be unidimensional. This finding indicates the strong empirical overlap among the worry, avoidance, and emotionality responses – and, thus, the cognitive-motivational interplay of anxiety components. Similarly, close relations had been already found elsewhere (Deffenbacher, 1980; Hodapp & Benson, 1997; Cassady & Johnson, 2002; Chin, Williams, Taylor, & Harvey, 2017;

Hong & Karstensson, 2002; Sarason, 1984), especially with respect to domain- or task-specific facets of test anxiety (Faber, 1995, 2012b). By no means, this result does challenge the need for a separate assessment of worry, avoidance, and emotionality cognitions. Rather this approach should ensure to obtain a more differentiated measuring of statistics anxiety and, thereby, should contribute to reducing the interpretation ambiguity of item responses. In that regard, it should certainly increase the scale's cognitive-motivational representativity and content validity. Likewise, the students' avoidance cognitions were found to be most closely related to their worry, but only slightly less closely related to their emotionality cognitions. Therefore, according to relevant findings (Galassi et al., 1981; Hagtvet & Benson, 1997), avoidance cognitions must be seen as an important feature within the students' anxiety experience and, thus, should have contributed to completing and refining the measuring of statistics anxiety (Putwain, 2008).

With respect to the validation results, both the correlation and regression analyses suggest, at first glance, that students' implicit entity beliefs are sufficient to explain their statistics anxiety. The entity beliefs appear to obviously play a crucial role in the prediction of statistics anxiety – as could be expected from the view of social-cognitive theories (Dweck & Leggett, 1988; Schunk & Zimmerman, 2006). However, in the students' cognitive-motivational processing, they will operate in a more complex manner. According to relevant theoretical conceptions and empirical findings (Blackwell et al., 2007; Chiesi & Primi, 2010; Emmioğlu, 2011; Onwuegbuzie, 2003; Sesé, Jiménez, Montano, & Palmer, 2015), it should be assumed that implicit beliefs actually mediate the effects of students' self-concept and their learning background on the dependent anxiety variable. The massive decline in the self-concept variable's beta weights, when adding the entity belief variable to the regression equation, apparently supports this assumption (Table 5). Indeed, within the validation framework this indirect effect cannot be adequately substantiated with correlational or regression analysis, but only with multivariate modeling method (Kline, 2011). Future research should make every effort to apply such modelling techniques in order to clarify the role of entity beliefs in the statistics domain.

Beyond the purpose of scale validation, the empirical findings concerning the students' entity beliefs might even extend the previous research in two respects: at the level of construct specificity, the measuring of entity beliefs

did not refer to the perceived malleability of general cognitive abilities but enquired the perceived malleability of statistical competencies. As this entity belief variable was only very weakly correlated with language self-concept (construction sample $r = .07$; validation sample $r = -.12$) it should be considered domain-specific. Hence, for the domain of statistical learning, this particular finding appears to be in line with the recommendations of the implicit theories approach (Dweck & Molden, 2005). Accordingly, at the level of construct relations, the results allow for refining the nomological scope of the statistics anxiety framework – at least, as it refers to the type and role of self-belief variables (Bandalos et al., 1995; González et al., 2016; Onwuegbuzie & Wilson, 2003; Zeidner, 1991).

The present study undeniably suffers from some conceptual and empirical limitations. First of all, composition and size of both student samples do not allow for generalizing the empirical findings as should be required. Instead, the findings reported here might claim a sort of local validity – all the more, as their data basis referred to a certain university setting. Further analyses should necessarily remedy this problem and examine the WAESTA scale with other student samples from other educational science contexts.

Moreover, the validation framework is still lacking in several respects and should be further completed (Benson, 1998). The present study assessed students' mathematics self-concept retrospectively. As a proportion of both samples did not have any prior experience of statistics using a measure of statistical self-concept would have been misguided. However, provided that further research could include participants being most comparable in their statistical background, their self-concept in the statistics domain should be absolutely used to elaborate scale validation (González et al., 2016). Likewise, concurrent measures of the students' self-efficacy to master certain statistical tasks could help to further differentiate the scale's criterion validity (Finney & Schraw, 2003; Perepiczka, Chandler, & Becerra, 2011). Not least, an appropriate validation of the WAESTA scale will require analysis of its relations with other instruments for measuring statistics anxiety – for instance, by comparing it with the German adaptation of the STARS questionnaire (Papousek et al., 2012).

Another considerable lack of the present study concerns the missing of a relevant performance measure. As only a certain part of students in both samples had yet to pass an exam in introductory statistics, sufficiently robust data were not available. Further validation studies should analyze the relation between the WAESTA scores and suitable measures of students' actual statistics performance. Especially, this relation should be most instructive – in as much as relevant studies commonly reported low to moderate correlations (Bandalos et al., 1995; Finney & Schraw, 2003; Macher, Paechter, Papousek, & Ruggeri, 2012; Sesé et al., 2015; Tremblay, Gardner, & Heipel, 2000; Vigil-Colet et al., 2008; Zeidner, 1991). However, these results do not really indicate a general flaw in the measures' criterion validity. Rather, they reflect the motivational consequences of statistics anxiety within a strongly restricted setting (Pekrun, 1988). As the successful passing of statistical requirements in the Master's degree is mandatory, the students' increasingly experienced worry, avoidance tendencies, and feelings of apprehension could dispose them to strengthen their learning effort in order to avoid an impending failure outcome (Macher, Papousek, Ruggeri, & Paechter, 2015; Martin & Marsh, 2003). Accordingly, for the WAESTA scale, also a moderate relation with the students' statistical performance should be assumed.

Finally, as both samples in this study were small and predominantly female, gender was not included in the validation analyses. Relevant findings in the field could consistently demonstrate the females to report a higher level of statistics anxiety (Benson, 1989; Hong & Karstensson, 2002; Macher et al., 2012; Onwuegbuzie & Wilson, 2003). Interestingly, despite the apparently heightened anxiety level of female students, some studies could not substantiate any significant disadvantage in their exam performance (Bradley & Wygant, 1998; Macher, Paechter, Papousek, Ruggeri, Freudenthaler, & Arendasy, 2013). This finding needs further clarification with respect to the underlying motivational and behavioral processes. Hence, female students might have overrated their individually existing anxiety level (Zeidner, 1998) – possibly due to a self-derogatory gender stereotyping effect (Bieg, Goetz, Wolter, & Hall, 2015; Pomerantz, Altermatt, & Saxon, 2002). As well, pursuing a more adaptive coping strategy to avoid feared failure, they might have ramped up their learning approach (Martin & Marsh, 2003). Given a larger sample size with a more

adequately balanced gender ratio, this issue should also be examined with respect to the WAESTA scale.

In summary, the present findings yield important information concerning the internal and external validity of the newly developed WAESTA scale. However, they must be seen as preliminary in nature. Therefore, they should represent just a very first step in method development.

References

- Ahmed, W., Minnaert, A., Kuyper, H., & van der Werf, G. (2012). Reciprocal relationships between math self-concept and math anxiety. *Learning and Individual Differences, 22*, 385-389. doi: [10.1016/j.lindif.2011.12.004](https://doi.org/10.1016/j.lindif.2011.12.004)
- Baloğlu, M. (2002). Psychometric properties of the Statistics Anxiety Rating Scale. *Psychological Reports, 90*, 315-325. doi: [10.2466/pr0.2002.90.1.315](https://doi.org/10.2466/pr0.2002.90.1.315)
- Bandalos, D. L., Yates, K., & Thorndike-Christ, T. (1995). Effects of math self-concept, perceived self-efficacy, and attributions for failure and success on test anxiety. *Journal of Educational Psychology, 87*, 611-623. doi: [10.1037/0022-0663.87.4.611](https://doi.org/10.1037/0022-0663.87.4.611)
- Benson, J. (1989). Structural components of statistical test anxiety in adults: An exploratory model. *Journal of Experimental Education, 57*, 247-261. doi: [10.1080/00220973.1989.10806509](https://doi.org/10.1080/00220973.1989.10806509)
- Benson, J. (1998). Developing a strong program of construct validation: A test anxiety example. *Educational Measurement: Issues and Practice, 17*, 10-17, 22. doi: [10.1111/j.1745-3992.1998.tb00616.x](https://doi.org/10.1111/j.1745-3992.1998.tb00616.x)
- Benson, J., Bandalos, D., & Hutchinson, S. (1994). Modeling test anxiety among men and women. *Anxiety, Stress, and Coping, 7*, 131-148. doi: [10.1080/10615809408249340](https://doi.org/10.1080/10615809408249340)
- Beurze, S. M., Donders, A. R. T., Zielhuis, G. A., de Vegt, F., & Verbeek, A. L. M. (2013). Statistics anxiety: A barrier for education in research methodology for medical students? *Medical Science Educator, 23*, 377-384. doi: [10.1007/BF03341649](https://doi.org/10.1007/BF03341649)
- Bieg, M., Goetz, T., Wolter, I., & Hall, N. (2015). Gender stereotype endorsement differentially predicts girls' and boys' trait-state

- discrepancy in math anxiety. *Frontiers in Psychology*, 6, 1404. doi:10.3389/fpsyg.2015.01404
- Birenbaum, M., & Eylath, S. (1994). Who is afraid of statistics? Correlates of statistics anxiety among students of educational sciences. *Educational Research*, 36, 93-98. doi: 10.1080/0013188940360110
- Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, 78, 246-263. doi:10.1111/j.1467-8624.2007.00995.x
- Blankenstein, K.R., Flett, G. L., & Watson, M. S. (1992). Coping and academic problem-solving ability in test anxiety. *Journal of Clinical Psychology*, 48, 37-46.
- Bradley, D. R., & Wygant, C. R. (1998). Male and female differences in anxiety about statistics are not reflected in performance. *Psychological Reports*, 82, 245-246. doi: 10.2466/pr0.1998.82.1.245
- Burnette, J. L., O'Boyle, E. H., VanEpps, E. M., Pollack, J. M., & Finkel, E. J. (2013). Mind-sets matter: A meta-analytic review of implicit theories and self-regulation. *Psychological Bulletin*, 139, 655-701. doi: 10.1037/a0029531
- Cassady, J. C., & Johnson, R. E. (2002). Cognitive test anxiety and academic performance. *Contemporary Educational Psychology*, 27, 270-295. doi: 10.1006/ceps.2001.1094
- Chew, P. K. H., & Dillon, D. (2014). Reliability and validity of the Statistical Anxiety Scale among students in Singapore and Australia. *Journal of Tropical Psychology*, 4(7), 1-7. doi: 10.1017/jtp.2014.7
- Chiesi, F., & Primi, C. (2009). Assessing statistics attitudes among college students: Psychometric properties of the Italian version of the Survey of Attitudes toward Statistics (SATS). *Learning and Individual Differences*, 19, 309-313. doi: 10.1016/j.lindif.2008.10.008
- Chiesi, F., & Primi, C. (2010). Cognitive and non-cognitive factors related to students' statistics achievement. *Statistics Education Research Journal*, 9(1), 6-26.
- Chin, E. C. H., Williams, M. W., Taylor, J. E., & Harvey, S. T. (2017). The influence of negative affect on test anxiety and academic performance: An examination of the tripartite model of emotions. *Learning and Individual Differences*, 54, 1-8. doi: 10.1016/j.lindif.2017.01.002

- Cruise, R. J., Cash, R. W., & Bolton, D. L. (1985). Development and validation of an instrument to measure statistical anxiety. Paper presented at the annual meeting of the Statistical Education Section. *Proceedings of the American Statistical Association, Chicago 1985* (pp. 92-97). Washington: American Statistical Association.
- Cury, F., Da Fonseca, D., Zahn, I., & Elliot, A. (2008). Implicit theories and IQ test performance: A sequential mediational analysis. *Journal of Experimental Social Psychology, 44*, 783-791. doi: [10.1016/j.jesp.2007.07.003](https://doi.org/10.1016/j.jesp.2007.07.003)
- Dauphinee, T. L., Schau, C., & Stevens, J. J. (1997). Survey of Attitudes Toward Statistics: Factor structure and factorial invariance for women and men. *Structural Equation Modeling, 4*, 129-141. doi: [10.1080/10705519709540066](https://doi.org/10.1080/10705519709540066)
- Deffenbacher, J. L. (1980). Worry and emotionality in test anxiety. In I. G. Sarason (Ed.), *Test anxiety: Theory, research and applications* (pp. 111-128). Hillsdale: Erlbaum.
- de Winter, J. C. F., Dodou, D., & Wieringa, P. A. (2009). Exploratory factor analysis with small sample sizes. *Multivariate Behavioral Research, 44*, 147-181. doi: [10.1080/00273170902794206](https://doi.org/10.1080/00273170902794206)
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256-273. doi: [10.1037/0033-295X.95.2.256](https://doi.org/10.1037/0033-295X.95.2.256)
- Dweck, C. S., & Molden, D. C. (2005). Self-theories. Their impact on competence motivation and acquisition. In A. J. Elliot & C. S. Dweck (Eds.), *Handbook of competence and motivation* (pp. 122-140). New York: Guilford Press.
- Earp, M. S. (2007). *Development and validation of the Statistics Anxiety Measure*. University of Denver, College of Education: Dissertation.
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*, 109-132. doi: [10.1146/annurev.psych.53.100901.135153](https://doi.org/10.1146/annurev.psych.53.100901.135153)
- Edmundson, E. W., Koch, W. R., & Silverman, S. (1993). A facet analysis approach to content and construct validity. *Educational and Psychological Measurement, 53*, 351-368. doi: [10.1177/0013164493053002005](https://doi.org/10.1177/0013164493053002005)

- Eichhorn, J. (2018). *Die Erfassung von studentischen Valenzkognitionen zur Nützlichkeit statistischer Kompetenzen. Entwicklung, psychometrische Prüfung und vorläufige Validierung eines entsprechenden Befragungsinstruments* [Measuring educational science majors' utility value of statistical competence. Scale development, psychometric properties, and preliminary validation results]. Leibniz University Hannover, Faculty of Humanities (Institute of Psychology): unpublished Master's thesis.
- 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. doi: 10.1037/0022-3514.76.4.628
- Emmioğlu, E. (2011). *A structural equation model examining the relationships among mathematics achievement, attitudes toward statistics, and statistics outcomes*. Middle East Technical University Ankara, Graduate School of Social Sciences: Doctoral thesis.
- Emmioğlu, E., & Capa-Aydin, Y. (2012). Attitudes and achievement in statistics: A meta-analysis study. *Statistics Education Research Journal*, 11(2), 95-102.
- Faber, G. (1995). Die Diagnose von Leistungsangst vor schulischen Rechtschreibsituationen: Neue Ergebnisse zu den psychometrischen Eigenschaften und zur Validität einer entsprechenden Kurzsкала [Measuring elementary school children's spelling-specific anxiety: Further psychometric and validation results]. *Praxis der Kinderpsychologie und Kinderpsychiatrie*, 44, 110-119.
- Faber, G. (2012a). *Elementary school children's spelling-specific self-beliefs. Longitudinal analyses of their relations to academic achievement, school attitudes, and self-esteem*. New York: Nova Science Publishers.
- Faber, G. (2012b). Measuring self-perceptions of oral narrative competencies and anxiety in the EFL context. *Electronic Journal of Research in Educational Psychology*, 10, 1343-1382.
- Finney, S. J., & Schraw, G. (2003). Self-efficacy beliefs in college statistics courses. *Contemporary Educational Psychology*, 28, 161-186. doi: 10.1016/S0361-476X(02) 00015-2
- Galassi, J. P., Frierson, H. T., & Sharer, R. (1981). Behavior of high, moderate, and low test anxious students during an actual test situation. *Journal of Consulting and Clinical Psychology*, 49, 51-62.

- Gaspard, H., Häfner, I., Parrisius, C., Trautwein, U., & Nagengast, B. (2017). Assessing task values in five subjects during secondary school: Measurement structure and mean level differences across grade level, gender, and academic subject. *Contemporary Educational Psychology*, 48, 67-84. doi: [10.1016/j.cedpsych.2016.09.003](https://doi.org/10.1016/j.cedpsych.2016.09.003)
- Goetz, T., Frenzel, A. C., Pekrun, R., Hall, N. C., & Lüdtke, O. (2007). Between- and within-domain relations of students' academic emotions. *Journal of Educational Psychology*, 99, 715-733. doi: [10.1037/0022-0663.99.4.715](https://doi.org/10.1037/0022-0663.99.4.715)
- Goetz, T., Pekrun, R., Hall, N., & Haag, L. (2006). Academic emotions from a social-cognitive perspective: Antecedents and domain-specificity of students' affect in the context of Latin instruction. *British Journal of Educational Psychology*, 76, 289-308. doi: [10.1348/000709905X42860](https://doi.org/10.1348/000709905X42860)
- Gogol, K., Brunner, M., Preckel, F., Goetz, T., & Martin, R. (2016). Developmental dynamics of general and school-subject-specific components of academic self-concept, academic interest, and academic anxiety. *Frontiers in Psychology*, 7, 356. doi: [10.3389/fpsyg.2016.00356](https://doi.org/10.3389/fpsyg.2016.00356)
- González, A., Rodríguez, Y., Failde, J. M., & Carrera, M. V. (2016). Anxiety in the statistics class: Structural relations with self-concept, intrinsic value, and engagement in two samples of undergraduates. *Learning and Individual Differences*, 45, 214-221. doi: [10.1016/j.lindif.2015.12.019](https://doi.org/10.1016/j.lindif.2015.12.019)
- Graham, J. W. (2012). *Missing data. Analysis and design*. New York: Springer.
- Green, J., Martin, A. J., & Marsh, H. W. (2007). Motivation and engagement in English, mathematics and science high school subjects: Towards an understanding of multidimensional domain specificity. *Learning and Individual Differences*, 17, 269-279. doi: [10.1016/j.lindif.2006.12.003](https://doi.org/10.1016/j.lindif.2006.12.003)
- Griffith, J. D., Mathna, B., Sappington, M., Turner, R., Evans, J., Gu, L., Adams, L. T., & Morin, S. (2014). The development and validation of the Statistics Comprehensive Anxiety Response Evaluation. *International Journal of Advances in Psychology*, 3(2), 21-29. doi: [10.14355/ijap.2014.0302.01](https://doi.org/10.14355/ijap.2014.0302.01)

- Guttman, R., & Greenbaum, C. W. (1998). Facet theory: Its development and current status. *European Psychologist*, 3, 13-36. doi: [10.1027//1016-9040.3.1.13](https://doi.org/10.1027//1016-9040.3.1.13)
- Hagtvet, K., & Benson, J. (1997). The motive to avoid failure and test anxiety responses: Empirical support for integration of two research traditions. *Anxiety, Stress, and Coping*, 10, 35-57. doi: [10.1080/10615809708249294](https://doi.org/10.1080/10615809708249294)
- Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment*, 7, 238-247. doi: [10.1037/1040-3590.7.3.238](https://doi.org/10.1037/1040-3590.7.3.238)
- Hodapp, V., & Benson, J. (1997). The multidimensionality of test anxiety: A test of different models. *Anxiety, Stress, and Coping*, 10, 219-244. doi: [10.1080/10615809708249302](https://doi.org/10.1080/10615809708249302)
- Hong, E., & Karstenson, L. (2002). Antecedents of state test anxiety. *Contemporary Educational Psychology*, 27, 348-367. doi: [10.1006/ceps.2001.1095](https://doi.org/10.1006/ceps.2001.1095)
- Hong, Y.-Y., Chiu, C.-Y., Dweck, C. S., Lin, D. M.-S., & Wan, W. (1999). Implicit theories, attributions, and coping: A meaning system approach. *Journal of Personality and Social Psychology*, 77, 588-599. doi: [10.1037/0022-3514.77.3.588](https://doi.org/10.1037/0022-3514.77.3.588)
- Hox, J. J. (1997). From theoretical concept to survey question. In L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz, & D. Trewin (Eds.), *Survey measurement and process quality* (pp. 47-69). Hoboken: Wiley. doi: [10.1002/9781118490013.ch2](https://doi.org/10.1002/9781118490013.ch2)
- İlhan, M., & Cetin, B. (2013). Mathematics Oriented Implicit Theory of Intelligence Scale: Validity and reliability study. *GESJ: Education Science and Psychology*, 25, 116-134.
- Kieffer, K. M., & Reese, R. J. (2009). Measurement of test and study worry and emotionality in college students. A psychometric evaluation of the Test and Study Attitudes Inventory. *Educational and Psychological Measurement*, 69, 303-321. doi: [10.1177/0013164408323231](https://doi.org/10.1177/0013164408323231)
- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd. ed.). New York: Guilford Press.
- Kooker, J., Welsh, M. E., McCoach, D. B., Johnston-Wilder, S., & Lee, C. (2016). Development and validation of the Mathematical Resilience

Scale. *Measurement and Evaluation in Counseling and Development*, 49, 217-242. doi: [10.1177/07481756155596782](https://doi.org/10.1177/07481756155596782)

- Lalonde, R. N., & Gardner, R. C. (1993). Statistics as a second language? A model for predicting performance in psychology students. *Canadian Journal of Behavioural Science*, 25, 108-125. doi: [10.1037/h0078792](https://doi.org/10.1037/h0078792)
- Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83, 1198-1202. doi: [10.1080/01621459.1988.10478722](https://doi.org/10.1080/01621459.1988.10478722)
- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4, 84-99. doi: [10.1037/1082-989X.4.1.84](https://doi.org/10.1037/1082-989X.4.1.84)
- Macher, D., Paechter, M., Papousek, I., & Ruggeri, K. (2012). Statistics anxiety, trait anxiety, learning behavior, and academic performance. *European Journal of Psychology of Education*, 27, 483-498.
- Macher, D., Paechter, M., Papousek, I., Ruggeri, K., Freudenthaler, H. H., & Arendasy, M. (2013). Statistics anxiety, state anxiety during an examination, and academic achievement. *British Journal of Educational Psychology*, 83, 535-549. doi: [10.1111/j.2044-8279.2012.02081.x](https://doi.org/10.1111/j.2044-8279.2012.02081.x)
- Macher, D., Papousek, I., Ruggeri, K., & Paechter, M. (2015). Statistics anxiety and performance: blessings in disguise. *Frontiers in Psychology*, 6, 1116. doi: [10.3389/fpsyg.2015.01116](https://doi.org/10.3389/fpsyg.2015.01116)
- Mandler, G., & Sarason, S. B. (1952). A study of anxiety and learning. *Journal of Abnormal and Social Psychology*, 47, 166-173. doi: [10.1037/h0062855](https://doi.org/10.1037/h0062855)
- Marsh, H. W., & O'Mara, A. J. (2008). Self-concept is as multidisciplinary as it is multidimensional. A review of theory, measurement, and practice in self-concept research. In H. W. Marsh, R. G. Craven & D. M. McInerney (Eds.), *Self-processes, learning, and enabling human potential. Dynamic new approaches* (pp. 87-115). Charlotte: Information Age Publishing.
- Martin, A. J., & Marsh, H. W. (2003). Fear of failure: Friend or foe? *Australian Psychologist*, 38, 31-38. doi: [10.1080/00050060310001706997](https://doi.org/10.1080/00050060310001706997)
- Möller, J., Streblow, L., Pohlmann, B., & Köller, O. (2006). An extension to

the internal/ external frame of reference model to two verbal and numerical domains. *European Journal of Psychology of Education*, 21, 467-487.

- Nasser, F. M. (2004). Structural model of the effects of cognitive and affective factors on the achievement of Arabic-speaking pre-service teachers in introductory statistics. *Journal of Statistics Education*, 12(1), 1-18.
- Nolan, M. M., Beran, T., & Hecker, K. G. (2012). Surveys assessing students' attitudes toward statistics: A systematic review of validity and reliability. *Statistics Education Research Journal*, 12(2), 103-123.
- Onwuegbuzie, A. J. (2003). Modeling statistics achievement among graduate students. *Educational and Psychological Measurement*, 63, 1020-1038. doi: [10.1177/0013164402250989](https://doi.org/10.1177/0013164402250989)
- Onwuegbuzie, A. J. (2004). Academic procrastination and statistics anxiety. *Assessment and Evaluation in Higher Education*, 29, 3-19. doi: [10.1080/0260293042000160384](https://doi.org/10.1080/0260293042000160384)
- Onwuegbuzie, A. J., & Wilson, V.A. (2003). Statistics anxiety: Nature, etiology, antecedents, effects and treatments – a comprehensive review of the literature. *Teaching in Higher Education*, 8, 195-209. doi: [10.1080/1356251032000052447](https://doi.org/10.1080/1356251032000052447)
- Papanastasiou, E. C., & Zembylas, M. (2008). Anxiety in undergraduate research method courses: its nature and implications. *International Journal of Research and Method in Education*, 31, 155-167. doi: [10.1080/17437270802124616](https://doi.org/10.1080/17437270802124616)
- Papousek, I., Ruggeri, K., Macher, D., Paechter, M., Heene, M., Weiss, E.M., Schuler, G., & Freudenthaler, H. (2012). Psychometric evaluation and experimental validation of the Statistics Anxiety Rating Scale. *Journal of Personality Assessment*, 94, 82-91. doi: [10.1080/00223891.2011.627959](https://doi.org/10.1080/00223891.2011.627959)
- Pekrun, R. (1988). Anxiety and motivation in achievement settings: towards a system-theoretical approach. *International Journal of Educational Research*, 12, 307-323. doi: [10.1016/0883-0355\(88\)90008-0](https://doi.org/10.1016/0883-0355(88)90008-0)
- Pekrun, R., Elliot, A. J., & Maier, M. A. (2009). Achievement goals and achievement emotions: Testing a model of their joint relations with academic performance. *Journal of Educational Psychology*, 101, 115-135. doi: [10.1037/a0013383](https://doi.org/10.1037/a0013383)
- Pekrun, R., Goetz, T., Perry, R. P., Kramer, K., Hochstadt, M., & Molfenter,

- S. (2004). Beyond test anxiety: Development and validation of the Test Emotions Questionnaire (TEQ). *Anxiety, Stress, and Coping*, 17, 287-316. doi: 10.1080/10615800412331303847
- Perepiczka, M., Chandler, N., & Becerra, M. (2011). Relationship between graduate students' statistics self-efficacy, statistics anxiety, attitude toward statistics, and social support. *The Professional Counselor*, 1(2), 99-108. doi:10.15241/mpa.1.2.99
- Pomerantz, E. M., Altermatt, E. R., & Saxon, J. L. (2002). Making the grade but feeling distressed. Gender differences in academic performance and internal distress. *Journal of Educational Psychology*, 94, 396-404. doi: 10.1037/0022-0663.94.2.396
- Putwain, D. W. (2008). Deconstructing test anxiety. *Emotional and Behavioural Difficulties*, 13, 141-155. doi: 10.1080/13632750802027713
- Putwain, D. W., & Symes, W. (2012). Achievement goals as mediators of the relationship between competence beliefs and test anxiety. *British Journal of Educational Psychology*, 82, 207-224. doi: 10.1111/j.2044-8279.2011.02021.x
- Rost, D. H., & Schermer, F. J. (1989). The various facets of test anxiety: A subcomponent model of test anxiety measurement. In R. Schwarzer, H. M. van der Ploeg & C. D. Spielberger (Eds.), *Advances in test anxiety research. Volume 6* (pp. 36-52). Lisse: Swets and Zeitlinger.
- Rost, D. H., Sparfeldt, J. R., & Schilling, S. R. (2007). *DISK-Gitter mit SKSLF-8. Differentielles Selbstkonzept-Gitter mit Skala zur Erfassung des Selbstkonzepts schulischer Leistungen und Fähigkeiten* [The self-concept grid: A questionnaire for measuring subject-specific facets of academic self-concept]. Göttingen: Hogrefe.
- Sarason, I. G. (1984). Stress, anxiety, and cognitive interference: Reactions to tests. *Journal of Personality and Social Psychology*, 46, 929-938. doi: 10.1037/0022-3514.46.4.929
- Schunk, D. H., & Zimmerman, B. J. (2006). Competence and control beliefs: Distinguishing the means and ends. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 349-367). New York: Erlbaum.
- Schutz, P. A., Drogosz, L. M., White, V. E., & Distefano, C. (1998). Prior knowledge, attitude, and strategy use in an introduction to statistics

- course. *Learning and Individual Differences*, 10, 291-308. doi: [10.1016/S1041-6080\(99\)80124-1](https://doi.org/10.1016/S1041-6080(99)80124-1)
- Schwarzer, R. (1996). Thought control of action: Interfering self-doubts. In I. G. Sarason, G. R. Pierce & B. R. Sarason (Eds.), *Cognitive interference: Theory, methods and findings* (pp. 99-115). Mahwah: Erlbaum.
- Schwarzer, R., & Quast, H.-H. (1985). Multidimensionality of the anxiety experience: Evidence for additional components. In H. M. van der Ploeg, R. Schwarzer & C. D. Spielberger (Eds.), *Advances in test anxiety research. Volume 4* (pp. 3-14). Lisse: Swets and Zeitlinger.
- Selkirk, L. C., Bouchey, H. A., & Eccles, J. S. (2011). Interaction among domain-specific expectancies, values, and gender: Predictors of test anxiety during early adolescence. *Journal of Early Adolescence*, 31, 361-389. doi: [10.1177/0272431610363156](https://doi.org/10.1177/0272431610363156)
- Sesé, A., Jiménez, R., Montano, J.-J., & Palmer, A. (2015). Can attitudes towards statistics and statistics anxiety explain students' performance? *Revista de Psicodidáctica*, 20, 285-304. doi: [10.1387/RevPsicodidact.13080](https://doi.org/10.1387/RevPsicodidact.13080)
- Skaalvik, E. M. (1997). Self-enhancing and self-defeating ego orientation: Relations with task and avoidance orientation, achievement, self-perceptions, and anxiety. *Journal of Educational Psychology*, 89, 71-81. doi: [10.1037/0022-0663.89.1.71](https://doi.org/10.1037/0022-0663.89.1.71)
- Stappert, A. (2017). *Die Erfassung impliziter Fähigkeitstheorien von Studierenden im Umgang mit statistischen Anforderungen. Entwicklung, psychometrische Prüfung und vorläufige Validierung eines entsprechenden Befragungsinstruments* [Measuring educational science majors' implicit theories of statistical competence. Scale development, psychometric properties, and preliminary validation results]. Leibniz University Hannover, Faculty of Humanities (Institute of Educational Psychology): unpublished Master's thesis.
- Tremblay, P. F., Gardner, R. C., & Heipel, G. (2000). A model of the relationships among measures of affect, aptitude, and performance in introductory statistics. *Canadian Journal of Behavioural Science*, 32, 40-48. doi: [10.1037/h0087099](https://doi.org/10.1037/h0087099)
- von der Embse, N., Jester, D., Roy, D., & Post, J. (2018). Test anxiety effects, predictors, and correlates: A 30-year meta-analytic review. *Journal of Affective Disorders*, 227, 483-493. doi: [10.1016/j.jad.2017.11.048](https://doi.org/10.1016/j.jad.2017.11.048)

- Vanhoof, S., Kuppens, S., Sotos, A. E. C., Verschaffel, L., & Onghena, P. (2011). Measuring statistics attitudes: Structure of the Survey of Attitudes Toward Statistics (SATS-36). *Statistics Education Research Journal*, 10(1), 35-51.
- Vigil-Colet, A., Lorenzo-Seva, U., & Condon, L. (2008). Development and validation of the Statistical Anxiety Scale. *Psicothema*, 20, 174-180.
- Wigfield, A., Hoa, L. W., & Klauda, S. L. (2009). The role of achievement values in the regulation of achievement behaviors. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning. Theory, research, and applications* (pp. 169-195). New York: Routledge.
- Williams, A. (2014). An exploration of preference for numerical information in relation to math self-concept and statistics anxiety in a graduate statistics course. *Journal of Statistics Education*, 22, 1-16. doi: [10.1080/10691898.2014.11889693](https://doi.org/10.1080/10691898.2014.11889693)
- Williams, A. S. (2015). Statistics anxiety and worry: The roles of worry beliefs, negative problem orientation, and cognitive avoidance. *Statistics Education Research Journal*, 14(2), 53-75.
- Zeidner, M. (1991). Statistics and mathematics anxiety in social science students: Some interesting parallels. *British Journal of Educational Psychology*, 61, 319-328. doi: [10.1111/j.2044-8279.1991.tb00989.x](https://doi.org/10.1111/j.2044-8279.1991.tb00989.x)
- Zeidner, M. (1998). *Test anxiety. The state of the art*. New York: Plenum.
- Zonnefeld, V. L. (2015). *Mindsets, attitudes, and achievement in undergraduate statistics courses*. University of South Dakota, Graduate Faculty: Dissertation.

Appendix. WAESTA items for measuring education science students' worry (WR), avoidance (AV), and emotionality (EM) cognitions encountering statistical demands (originally worded in German)

01 WR I will hardly be able to meet my degree program's statistics requirements right away.

05 WR It would be difficult for me to discuss statistical content adequately in my papers.

08 WR I have difficulty understanding statistical content in a lecture.

10 WR If I had to comment on statistical data in a course, I would be worried that I would make a fool of myself.

11 WR If I had to give a presentation including statistical findings in a course, I would hope that no one had any follow-up questions.

12 WR I would hardly be able to present a report on statistical research findings adequately.

14 WR Despite careful preparation for a statistics exam, I would worry about not passing it.

16 WR If I took a statistics course, I would be concerned that I would quickly forget everything I had learned.

03 AV If I could, I would rather take two other courses than do one statistics course.

04 AV When presentation topics are being assigned in the course, I would make sure that I receive a topic that doesn't involve statistics.

06 AV When preparing presentations, I would rather omit anything that has to do with statistics.

284 *Faber, Drexler, Stappert & Eichhorn – Education science students' statistics anxiety*

17 AV In scientific texts, I would skip over statistical tables and diagrams if possible.

02 EM I would be very uncomfortable if I had to work on a statistical problem.

07 EM I would be quite nervous if I were asked to explain a chart from a research report.

09 EM I would have trouble extracting the relevant information from a table of statistical values.

13 EM I would feel very tense if I had to apply a statistical formula.

15 EM The thought of having to explain a statistical problem in a course makes me quite nervous.

Günter Faber, Dr., is a lecturer at Leibniz University Hannover, Institute of Psychology.

Heike Drexler, Dr., is a lecturer at Leibniz University Hannover, Institute of Psychology.

Alexander Stappert, M.A., is a research assistant at the University of Vechta, Department of Social and Education Sciences.

Joana Eichhorn, B.A., is a student in the Master's Programme of Educational Sciences, Leibniz University Hannover.

Contact address: Günter Faber, Institute of Psychology, Leibniz University Hannover, Schloßwender Straße 1, D-30159 Hannover.
Email: faber@psychologie.uni-hannover.de