

Vanessa Hübner & Maximilian Pfof

Operationalization of Academic Risk-Taking in University Students

Abstract

Academic Risk-Taking (ART) is a type of student behavioral engagement characterized by uncertainty regarding the outcome. Students for example share their ideas on difficult topics during class. When taking academic risks, students are confronted with the possibility of being perceived as less competent by others. Students and instructors ascribe a beneficial effect to ART, and ART may foster academic achievement by promoting deep learning. Still, research shows that students avoid taking academic risks. This phenomenon is largely unexplained and there is only little research on ART in the context of higher education. This may be due to a lack of short survey instruments. This study addresses this issue by proposing an extensive self-report measure to evaluate students' general (G-ART) as well as seminar-specific ART (S-ART). First results indicate that two dimensions of ART, seminar group and peer context, may be differentiated. The third dimension of instructor context shows low consistency and needs to be revised. Possible applications of the scale include closer exploration of the nature of the construct ART regarding stability over time and different contexts. The instrument enables analyzing possible social and gender disparities as well as the relationship between seminar characteristics, ART, and academic achievement.

Keywords

academic risk-taking, higher education, student engagement, self-report scale

Vanessa Hübner, M. Sc., ORCID: 0000-0003-3568-7618 (corresponding author) •
Prof. Dr. Maximilian Pfof, ORCID: 0000-0002-7066-0456, University of Bamberg,
Department of Educational Research, Markusplatz 3, 96045 Bamberg, Germany
email: vanessa.huebner@uni-bamberg.de
maximilian.pfof@uni-bamberg.de

Operationalisierung akademischer Risikobereitschaft Studierender

Zusammenfassung

Akademische Risikobereitschaft (ART) ist eine Form der behavioralen Partizipation und wird durch die Unsicherheit Studierender bezüglich des Ergebnisses einer Lernsituation charakterisiert. Dies ist beispielsweise der Fall, wenn Studierende während Seminaren ihre Ideen zu schwierigen Themen teilen. Gehen Studierende akademische Risiken ein, sind sie mit der Möglichkeit konfrontiert, von anderen als weniger kompetent wahrgenommen zu werden. Studierende und Dozierende schreiben ART einen lernförderlichen Effekt zu und ART kann durch die Unterstützung tiefen Lernens positiv auf akademischen Erfolg wirken. Dennoch werden akademische Risiken von Studierenden gemieden. Dieses Phänomen ist weitgehend unerklärt, es existiert nur wenig Forschung zu ART im Hochschulbereich. Dies kann auf einen Mangel an simplen Erhebungsinstrumenten zurückzuführen sein. Die vorliegende Studie befasst sich mit diesem Problem durch die Entwicklung eines umfassenden Selbstauskunftsinstruments, das ART sowohl in allgemeinen (G-ART) als auch in seminarspezifischen (S-ART) Kontexten erfasst. Erste Ergebnisse deuten auf die Differenzierung von zwei Dimensionen hin: Seminargruppe und Peer-Kontext. Die dritte Dimension, der Dozentenkontext, weist eine geringe Konsistenz auf und benötigt weitere Überarbeitung. Mögliche Anwendungen der Skala umfassen die nähere Untersuchung der Beschaffenheit des Konstrukts ART hinsichtlich der Stabilität über die Zeit und verschiedene Kontexte hinweg, die Analyse möglicher sozialer und geschlechtsspezifischer Disparitäten sowie die Beziehung zu Seminarcharakteristika und akademischem Erfolg.

Schlagworte

Akademische Risikobereitschaft, Hochschule, Partizipation, Fragebogenskala

1. Introduction

Academic Risk-Taking (ART) is a special type of student engagement (Beghetto et al., 2020; Clifford, 1991) which, since the 1990s, generates a growing corpus of research. ART is characterized by students' uncertainty regarding the correctness of their contribution and thus, the outcome of the learning situation (Clifford, 1991). Therefore, students risk making mistakes, receiving implicit or explicit negative feedback, and subsequently being perceived as less intelligent or competent by their peers or instructors (Beghetto, 2009). Students who take academic risks may for example share their ideas on a topic during class even though they are not sure about the quality of their contribution, or ask their peers for feedback on term papers that still need editing. Unlike illegal or dangerous behavior which is classified as negative risk, ART, being socially accepted and constructive can be classified as a positive risk (Duell &

Steinberg, 2019). A study by Krochmal and Roth (2017) suggests that ART has a beneficial effect on students' academic achievement by fostering deep learning. This assumption is supported by Özbay and Köksal (2021), who found an association between ART and secondary students' science achievement. Additionally, university instructors think that students should take more risks in terms of challenging their own beliefs, communicating and networking with others, and seeking academic exploration (Ravert & Schneller, 2019). Students likewise perceive ART to be worthwhile, however, they hesitate to take academic risks (Teagarden et al., 2018) as this poses for example a threat to their self-concept (Ellis, 2015). Özbay and Köksal (2021) state that students' ART levels are dependent on their willingness to participate, their enjoyment, their resilience, their problem-solving skills, as well as motivational factors such as their goal orientation. Especially mastery goal orientation seems to be beneficial for ART (Dachner et al., 2017), the effect possibly mediated by preferred task difficulty. Abercrombie, Carbonneau, and Hushman (2022) found that students who are mastery goal-oriented tend to prefer difficult tasks and therefore take more academic risks. Furthermore, students take more academic risks, when more autonomy is granted by the instructor and when the instructor poses high performance expectations on the students (Dachner et al., 2017).

To our knowledge, no attempts have yet been made to systematically and conjointly investigate the associations between ART, class characteristics, students' individual characteristics, and academic success. Additionally, there is a lack of research on ART in university students compared to students attending primary or secondary education. Research shows, for example, that school students are generally quite uninterested in taking academic risks (Clifford & Chou, 1991), but those who do show ART tend to achieve higher academic outcomes (Bal-İncebacak et al., 2019). Research on whether the same holds for university students is missing. This may also be due to a lack of instruments measuring ART. In fact, we are aware of three instruments currently available. One method includes the observation of primary students' behavior, either during regular lessons or after administering mathematics, spelling, and vocabulary items of varying difficulty (Clifford, 1988; Clifford et al., 1990), which is a laborious approach, subject-specific and not adequate for university students. Beghetto's (2009) self-report measure is generic and therefore often used, but it is unidimensional and does not consider context and persons who might observe the behavior. However, students' individual judgement about the magnitude of risk and, therefore, their behavior may differ according to different situations (Rohrman, 2005). For example, in the context of their model of identity-based disengagement, Lund Dean and Jolly (2012) state that students' perception of risk and, in turn, their engagement may differ according to the role they take on in front of certain persons, but also within certain situations with varying degree of formality. Therefore, it tends to be important to consider and differentiate environmental circumstances when measuring ART, such as formal in-seminar situations, in which students only have limited control over who is watching, as well as less formal settings like study groups, mostly consisting of members coming to-

gether based on sympathy or joint interests. Finally, some authors (e.g., Akbay & Delibalta, 2020; Çetin et al., 2014) measure ART using Clifford's (1988) School Failure Tolerance scale as a proxy measure for ART. The measure considers affect, cognition, and behavior by evaluating students' emotional reaction after making mistakes, their preference concerning the difficulty of a task, and their tendency to engage in adaptive behavior after having made a mistake or when having difficulties during the learning process. Those aspects are important determinants of ART (Clifford, 1988, 1991). While we acknowledge the usefulness of this instrument to approximate and predict ART, even after the recent revision by Abercrombie, Carbonneau, and Hushman (2022), the instrument has a strong focus on the aspect of dealing with mistakes that have already happened, which is not necessarily a defining feature of ART. Contrary to this retrospective perspective on mistakes, academic risk-taking focuses on the acceptance toward potentially making mistakes in front of others and therefore also taking situational contexts into account. Consequently, Clifford's (1988) School Failure Tolerance scale falls short of reflecting this aspect and just in parts captures the construct of ART. The soundness of interpretations and conclusions based on the results of studies using this instrument to measure ART are thus questionable at best.

By proposing a multidimensional self-report scale that can be used for evaluating either ART in general contexts (G-ART) or ART in a specific context (S-ART), this paper addresses the lack of generic and easy-to-use instruments for academic risk-taking in higher educational contexts. Additionally and by focusing on the prospective component of ART, that is, the acceptance of making potential mistakes that have not yet happened, the article illuminates the structure of the construct in more detail. Concerning the internal structure of ART, we expect to find three distinguishable dimensions within G-ART and S-ART, although we expect all dimensions to correlate positively with each other. Additionally, ART being a special type of engagement, we expect S-ART to correlate positively with seminar-specific engagement. Concerning relations to further criterion variables, we expect to find positive correlations of G-ART and S-ART with mastery goal orientation and grade point average (GPA). However, we expect negative correlations with performance-avoidant goal orientation, which would be in alignment with the theoretical conceptualization of the construct and previous research mentioned above (Beghetto, 2009; Clifford, 1991; Özbay & Köksal, 2021). There is currently no research on the association between ART and teaching quality, but Leach (2016) suggests that student engagement can be enhanced by teaching quality. Therefore, using students' subjective rating of the specific seminar as a proxy for teaching quality, we expect S-ART to correlate positively with seminar rating. Finally, we expect to see a positive association between S-ART and seminar-specific learning success, as ART is meant to promote deep learning (Krochmal & Roth, 2017). We make no assumption on the association of ART with performance-approach goal orientation due to inconclusive prior research findings (Abercrombie, Bang, & Vaughan, 2022; Abercrombie, Carbonneau, & Hushman, 2022) and no clear-cut theoretical derivations.

2. Method

2.1 Development and Description of the Instrument

First, we formulated 22 items based on the theoretical conceptualization of ART, which meant for each item to suggest a certain risk of failure that is realistic for standard academic contexts. We considered the results of interview studies conducted with university instructors and students (Figueira et al., 2018; Ravert & Schneller, 2019; Teagarden et al., 2018) and in which the participants were asked to describe risky academic situations. Drawing from the model of identity-based disengagement (Lund Dean & Jolly, 2012), which assumes that students' engagement levels differ according to the role they take on in front of certain persons and within certain situations, we divide ART into three dimensions which are characterized by who is observing the behavior. The first dimension Seminar Group includes situations in which students' peers as well as their instructors are present (12 items, e.g., "To participate in seminar discussions even on difficult topics."). The second dimension Instructors refers to situations in which only the students' instructor is present (5 items, e.g., "To request feedback from instructors on term papers for which I have received a poor grade."). The third dimension Peers refers to situations in which only students' peers are present (5 items, e.g., "To form a study group with fellow students, even if I feel I know less than the others."). Participants were asked to rate the likelihood of the described behavior on a 5-point Likert scale (*very unlikely, rather unlikely, undecided, rather likely, very likely*). Work by Fredricks et al. (2004) suggests that engagement possesses both, a state and a trait component, so in order to consider context sufficiently, we used the same scales twice, asking the students to rate the items regarding university seminars in general as well as answering them regarding one specific seminar.

In a second step, we conducted cognitive interviews with a sample of five participants. The aim of those interviews was to spot item formulations that were arbitrary, hard to understand, or otherwise misleading. The participants read all items aloud and were asked to verbalize all of their thoughts before marking their answer. The cognitive interviews resulted in minor specifications or changes of wording, but there were no systematic problems occurring repeatedly. The Appendix includes all items of the preliminary instrument as well as the introductory text and answer categories in German language. In addition and for linguistic purposes only, an English translation, which has not been tested, is attached.

2.2 Design and Participants

In order to test the instrument regarding its internal consistency, dimensional structure, and validity, we conducted a pilot study in the second half of the summer semester 2021. We chose German university students studying within the field of social sciences or humanities as our target audience. We collected the data using

an online questionnaire as well as a paper-pencil questionnaire that were identical in content and item order.

Overall, we recruited 159 participants, 116 responding to the online questionnaire and 43 responding to the paper-pencil version. The students' mean age was 24.92 years ($SD = 6.20$ years). Forty-five percent of them were Bachelor students, 23% Master students, 23% students aiming for state examination, and 9% did not specify their studies. Seventy-two percent of participants were females.

2.3 Measures

2.3.1 Academic Risk-Taking

In order to measure students' ART levels, we used the 3-dimensional instrument described in the previous chapter. Each participant estimated their ART behavior twice: The items were first introduced referring to a specific context. In this regard, participants were asked to refer to the seminar that they last attended. Second, the items were introduced referring to a general context.

2.3.2 Goal Orientation

We used the scales developed for the StEG study and slightly adjusted them so they fit the university context (Leibniz-Institut für Bildungsforschung und Bildungsinformation, 2020a, 2020b, 2020c). Contrary to the original version, a 5-point Likert scale instead of a 4-point Likert scale was used. The subscale of mastery goal orientation consists of five items and in our study obtains $\alpha = .66$. The subscale of performance-approach goal orientation consists of five items and shows an internal consistency of $\alpha = .87$. The subscale of performance-avoidant goal orientation consists of four items and internal consistency in our study is $\alpha = .84$.

2.3.3 Auxiliary Variables

We asked participants to specify their age and gender. We also inquired the participants' seminar-specific subjective learning success and their own seminar-specific engagement. Those items were measured on a Likert scale from 1 (*very low*) to 6 (*very high*). Furthermore, the participants were asked to rate the overall seminar quality, using German grades from 1 (*very good*) to 6 (*not sufficient*).

2.4 Analysis Strategy

For data analyses, we used R Version 3.6.3 (R Core Team, 2020) in order to optimize and test the ART scale. Using *psych* Version 2.1.9 (Revelle, 2021) we con-

ducted exploratory factor analysis (EFA) for ordinal data. Since we had assumptions concerning the loading pattern of the items, we used *GPArotation* Version 2014.11-1 (Bernaards & Jennrich, 2005) for target rotation (*targetQ*) towards a pre-defined matrix. Items with loadings $< .3$ on their assigned factor or cross-loadings $> .5$ were deleted. We also considered internal consistency using Cronbach's α and item-total correlations. We used *ltm* Version 1.1-1 (Rizopoulos, 2006) for graded response models (GRM) in order to investigate the items' discriminatory power and difficulty. GRM, which are based on item response theory, are used for ordinal data and assume that the probability for a person to answer in a certain answer category is different for each item, even if the person's score on the latent construct is stable. The discrimination parameter indicates how strongly an item and the measured latent construct are associated. The between category thresholds indicate the point on the latent variable where the likelihood of answering in a given category is equal to 0.5 (for a more detailed overview on GRM see Baker, 2001). We deleted those items that had discriminatory power below 1.35 and thus, only keeping those items with at least high discrimination (Baker, 2001). We explored our instrument's validity using correlation.

3. Results

3.1 Descriptive Statistics

Table 1 shows means and standard deviations of the individual ART items of both contexts as well as item correlations with the total score of their respective subscale and the scales' internal consistencies. With the exception of Item 17, floor or ceiling effects were not present. Item-total correlations of the items with their respective scales lay between .00 and .74 for the general and between $-.06$ and .74 for the specific context. We observe Cronbach's α that is acceptable or good ($> .7$) for the group and peer subscales, but insufficient ($< .6$) for the instructor subscale. This result holds for both contexts.

3.2 Factor Analysis and Reliability

3.2.1 Factor Structure

Table 2 presents the results of EFA with target-rotation that we conducted in order to explore the assumed underlying structure with three factors for both, the general and the specific context of the ART scale. With 28% in the general context and 26% in the specific context, the seminar group dimension explains the highest proportion of variance. The peer factor in the general context explains 14% and 13% in the specific context. The instructor factor explains the smallest proportion of variance with 8% or 6%, respectively. In the general context, Factors 1 and 2 correlate

Table 1: Descriptive Statistics of the Preliminary ART Scale

	General				Specific			
	<i>M</i>	<i>SD</i>	r_{it}	α [95% CI]	<i>M</i>	<i>SD</i>	r_{it}	α [95% CI]
Seminar group	Item 1	3.39	1.09	.64	4.42	1.11	.60	
	Item 2	3.38	1.02	.69	3.48	1.09	.70	
	Item 3	3.21	1.24	.67	3.30	1.21	.60	
	Item 4	2.90	1.21	.69	3.00	1.31	.64	
	Item 5	3.34	1.12	.63	3.41	1.17	.66	
	Item 6	3.48	1.13	.29	3.17	1.25	.25	.86
	Item 7	2.60	1.15	.57	2.63	1.17	.53	[.82, .89]
	Item 8	3.12	1.18	.60	3.43	1.20	.57	
	Item 9	3.20	1.12	.61	3.02	1.12	.48	
	Item 10	3.23	1.12	.41	3.16	1.11	.28	
	Item 11	2.66	1.09	.41	2.62	1.12	.35	
	Item 12	2.90	1.08	.74	3.01	1.22	.70	
Instructors	Item 13	3.18	1.27	.56	3.13	1.30	.50	
	Item 14	3.43	1.30	.49	3.54	1.33	.54	
	Item 15	2.23	1.05	.20	2.16	0.95	.17	.55
	Item 16	3.30	1.12	.42	3.18	1.31	.42	[.46, .65]
	Item 17	4.00	0.91	.00	4.01	0.88	-.06	
Peers	Item 18	3.28	1.29	.62	2.92	1.36	.55	
	Item 19	3.27	1.26	.65	2.88	1.39	.74	
	Item 20	3.10	1.28	.59	2.93	1.40	.48	.74
	Item 21	3.69	1.10	.66	3.37	1.21	.56	[.68, .80]
	Item 22	3.58	0.94	.28	3.51	1.01	.23	

at $\phi = .42$ while Factors 1 and 3 correlate at $\phi = .17$ and Factors 2 and 3 correlate at $\phi = .10$. In the specific context, Factors 1 and 2 correlate at $\phi = .41$ while Factors 1 and 3 correlate at $\phi = .15$ and Factors 2 and 3 correlate at $\phi = .15$.

Six items of the general and the specific context, respectively, have factor loadings $< .3$ on their assigned factor. Concerning the seminar group dimension, very high cross-loadings $> .5$ concern Items 6 and 11 in both contexts and Item 10 in the specific context. Concerning the instructor dimension, except for one item, low factor loadings on all items are found in both contexts. In short, for the instructor dimension a unidimensional loading pattern is not found. Consequently, the dimension was not considered for further analysis. Item 22 was deleted due to low factor loading on the peer factor.

At this stage, the deletion of above discussed items results in a 2-dimensional instrument for both contexts. The group subscale of the general context consists of 10 items, while in the specific context nine items seem appropriate. The peer subscale of both, the general and the specific context, consists of four items.

3.2.2 Item Discrimination

GRM are based on item response theory, which assumes unidimensionality for its models, so we evaluated thresholds and discrimination parameters for each sub-scale individually. The results are shown in Table 3 and Table 4. Observing the items' thresholds between categories, we do not find any peculiarities; five answer categories seem adequate. Item 7 and Item 10 of the general context have discriminatory power < 1.35. Item 7 and Item 9 have discriminatory power below the cutoff in the specific context.

Table 2: Results of EFA With Target Rotation

	General				Specific				
	F1	F2	F3	com	F1	F2	F3	com	
Seminar group	Item 1	.78	-.11	-.04	.54	.70	-.17	.32	.58
	Item 2	.81	.08	-.10	.69	.82	.03	-.07	.67
	Item 3	.79	-.10	.00	.64	.69	.07	-.14	.51
	Item 4	.86	-.23	.04	.64	.77	.05	-.25	.63
	Item 5	.78	.03	-.25	.63	.90	-.08	-.25	.76
	Item 6	.19	.03	-.25	.31	.07	.28	.41	.32
	Item 7	.54	.05	.21	.40	.55	-.03	.31	.44
	Item 8	.62	.22	-.01	.54	.65	.16	-.11	.52
	Item 9	.74	-.03	-.08	.52	.61	-.17	.27	.43
	Item 10	.43	-.03	.16	.22	.15	.23	.37	.28
	Item 11	.26	.02	.76	.73	.09	.35	.40	.38
	Item 12	.76	.14	.04	.70	.79	.06	-.14	.65
Instructors	Item 13	.25	.33	.12	.28	.26	.39	-.05	.30
	Item 14	.27	.27	.06	.22	.30	.34	-.05	.28
	Item 15	.22	.01	.75	.67	.00	.40	.27	.27
	Item 16	.34	.34	.08	.35	.42	.35	-.17	.41
	Item 17	.10	.12	-.04	.03	.18	-.18	.31	.13
Peers	Item 18	.08	.70	.00	.54	.02	.56	.09	.34
	Item 19	-.10	.85	-.08	.67	.08	.76	.06	.65
	Item 20	-.15	.74	.16	.52	-.08	.74	-.19	.50
	Item 21	-.05	.80	.01	.61	.19	.58	.00	.46
	Item 22	.52	.17	-.11	.36	.66	-.08	.28	.52
Prop Var	.28	.14	.08		.26	.13	.06		

Note. Factor loadings > +/- .3 in bold.

Table 3: Between Category Thresholds and Discriminations of the Items in the General Context

	Threshold 1	Threshold 2	Threshold 3	Threshold 4	Discrm		
Seminar group	Item 1	-2.123	-1.020	0.065	1.315	1.986	
	Item 2	-1.918	-0.873	-0.107	1.438	3.011	
	Item 3	-1.460	-0.720	-0.069	1.557	2.021	
	Item 4	-1.362	-0.243	0.347	1.851	2.121	
	Item 5	-1.960	-0.766	-0.114	1.441	2.257	
	Item 7	-1.352	-0.080	1.161	2.885	1.258	
	Item 8	-1.743	-0.642	0.211	1.649	1.797	
	Item 9	-1.954	-0.772	0.350	1.455	1.900	
	Item 10	-3.497	-1.395	0.160	2.433	0.852	
	Item 12	-1.545	-0.331	0.561	1.764	2.806	
	Peers	Item 18	-1.608	-0.592	-0.030	1.254	1.869
		Item 19	-1.328	-0.676	-0.015	1.197	2.721
Item 20		-1.469	-0.520	0.168	1.534	1.784	
Item 21		-2.055	-1.213	-0.382	0.808	2.621	

Note. Discrimination < 1.35 in bold.

Table 4: Between Category Thresholds and Discrimination of the Items in the Specific Context

	Threshold 1	Threshold 2	Threshold 3	Threshold 4	Discrm	
Seminar group	Item 1	-2.251	-1.167	-0.045	1.331	1.697
	Item 2	-1.807	-1.032	-0.134	0.957	3.068
	Item 3	-1.807	-0.952	-0.032	1.235	1.820
	Item 4	-1.227	-0.398	0.202	1.372	2.225
	Item 5	-1.768	-0.720	-0.188	1.006	3.041
	Item 7	-1.484	-0.083	1.036	2.727	1.198
	Item 8	-2.008	-0.991	-0.240	1.139	1.742
	Item 9	-2.690	-0.524	0.517	2.263	1.169
	Item 12	-1.450	-0.304	0.300	1.162	3.015
	Peers	Item 18	-1.163	-0.324	0.307	1.985
Item 19		-0.791	-0.177	0.273	1.217	4.052
Item 20		-1.242	-0.189	0.328	1.547	1.456
Item 21		-1.912	-0.763	-0.008	1.104	1.936

Note. Discrimination < 1.35 in bold.

3.2.3 Testing the Final Instrument

Finally, the resulting scales were again scrutinized with regard to content and scale heterogeneity. Since one aim was to develop an instrument that can be used for general and specific contexts, we matched both scales. Therefore, despite low discrimination, we kept Item 9 within the specific context. However, it seemed appropriate to delete Item 3 due to high redundancy with Item 4. In addition, Item 5 was deleted due to strong content overlap with Item 2. Therefore, the final instrument consists of 10 items depicted in Table 5, which also shows the results of the re-assessed internal consistencies and factor loadings conducting EFA. A coherent loading pattern is observed, and the two factors explain a substantial proportion of variance in both contexts. They correlate at $\phi = .35$ in the general context and at $\phi = .43$ in the specific context.

Table 5: Results of EFA with Target Rotation and Cronbach’s Alpha of the Final Instrument

		General				Specific			
		F1	F2	com	α [95% CI]	F1	F2	com	α [95% CI]
Seminar group	Item 1	.76	-.13	.52		.68	-.08	.43	
	Item 2	.82	.07	.71		.84	.03	.73	
	Item 4	.77	-.17	.54	.86 [.83, .90]	.72	.01	.52	.83 [.79, .87]
	Item 8	.65	.16	.52		.66	.12	.52	
	Item 9	.73	-.03	.51		.53	-.06	.26	
	Item 12	.83	.13	.77	.83 [.79, .87]	.83	.00	.68	.83 [.77, .86]
Peers	Item 18	.14	.66	.51		-.04	.65	.40	
	Item 19	-.03	.84	.69	.82 [.77, .86]	.00	.93	.86	.78 [.72, .84]
	Item 20	-.07	.74	.51		-.09	.70	.44	
	Item 21	.02	.81	.68		.16	.64	.52	
Prop Var		.35	.24			.31	.22		

Note. Cronbach’s α reported in Columns 4 and 8 were generated including the items of the individual subscales, while the parameters in Columns 5 and 9 include all 10 items of the scale.

3.3 Validity

We used aggregated means of the subdimensions as individual indicators for ART for further analyses concerning validity. Sample means and standard deviations of all variables are presented in Table 6. As expected, we find strong and significant associations between the group dimensions and the peer dimensions of the general and specific context, respectively. The correlations between the respective group dimensions and peer dimensions are moderate. Additionally, we observe that the group dimensions of ART correlate significantly with students’ perceived own engagement within the specific seminar. Furthermore, we observe that students’ ART

Table 6: Interrelations of the Dimensions of ART and Auxiliary Variables

	<i>M</i>	<i>SD</i>	1	2	3	4	5	6	7	8	9	10	11
1. G-ART: Group	3.13	0.80	1										
2. G-ART: Peers	3.33	0.99	.27**	1									
3. S-ART: Group	3.26	0.87	.76**	.24**	1								
4. S-ART: Peers	3.03	1.04	.26**	.77**	.31**	1							
5. Mastery GO	4.38	0.46	.28**	.09	.23**	.11	1						
6. Approach GO	2.04	0.90	-.13	.01	-.07	-.02	-.08	1					
7. Avoidant GO	2.76	1.02	-.51**	-.14	-.46**	-.20*	-.12	.49**	1				
8. LS	4.72	1.40	.16**	.22**	.29**	.26**	.23**	-.05	-.14	1			
9. Eng. Self	4.72	1.42	.31**	-.03	.42**	.04	.27**	.04	-.19**	.41**	1		
10. Seminar Rating	2.20	1.00	-.12	-.14*	-.31**	-.20**	-.19**	.03	.22	-.53**	-.34**	1	
11. GPA	2.39	1.04	-.12	-.10	-.07	-.09	-.16	-.20**	-.03	-.16*	-.20**	.14*	1

Note. Correlations of Variables 1–7 with each other are parametric (Pearson's r), while all other correlations are non-parametric (Kendall's tau). Negative correlations with GPA, and seminar rating indicate an association of the respective variable with better grades/better rating, since the German grading system ranges from 1 (very good) to 6 (not sufficient). G-ART = ART in general contexts; S-ART = ART in a specific context; GO = goal orientation; LS = seminar-specific learning success; Eng. Self = perceived own engagement; GPA = grade point average.

* $p < .05$, ** $p < .01$.

in the seminar group is associated with goal orientation. Students tend to have higher levels of ART with high mastery goal orientation. In contrast, performance-avoidant goal orientation seems to inhibit students' ART on the seminar group dimension. We do not observe a significant association between students' ART and performance-approach goal orientation. Moreover, seminar rating seems to play a role, especially regarding students' S-ART. The better students perceive the seminar to be, the higher their ART within this context. In terms of learning outcome, we find that students' subjective learning success within a specific seminar tends to be higher with higher levels of G-ART and S-ART. We find the same tendency with regard to students' GPA, although the relation is not significant.

4. Discussion

This study aims to address the lack of instruments to evaluate university students' ART levels by proposing a self-report scale. Assuming three dimensions, we formulated 22 items based on the results of previous interview studies (Figueira et al., 2018; Ravert & Schneller, 2019; Teagarden et al., 2018). Those were reduced in a stepwise process, taking into consideration aspects of dimensionality, internal consistency, item discrimination, and, lastly, content. This process resulted in an instrument of 10 items on two dimensions: the seminar group dimension and the peer dimension.

4.1 Internal Structure

The seminar group dimension and the peer dimension are distinguishable in both contexts using exploratory factor analysis and are internally consistent with respect to Cronbach's α . Additionally, we observe a significant correlation between the ART seminar group dimension and student engagement in both the general and the specific context. This result indicates that ART is, in fact, a type of student engagement. Contrary to the assumed 3-dimensional structure, however, items that were intended to represent the instructor dimension do not show a coherent unidimensional loading pattern. The items were formulated with a focus on course requirements, such as presentations and term papers, since those are usually opportunities for personal talk with the instructor. However, those requirements are not a prerequisite in all seminars and therefore it is possible that our items produce too much error variability. Additionally, the correlation between engagement and ART is not found for the peer dimension. This result might be explained by the fact that while the item we used for measuring students' general engagement focuses on on-campus courses, S-ART on the peer dimension especially focuses on interaction beyond on-campus courses.

4.2 ART in General and Specific Contexts

Based on the theoretical assumption that students' engagement is dependent on who is present and observing the behavior (Lund Dean & Jolly, 2012) and in order to explore whether ART also has a trait component as suggested by Fredricks et al. (2004), we aimed to measure ART in general and with regard to a specific seminar. We find first evidence suggesting the relevance of examining ART levels within general and specific contexts separately. Though the respective subscales of G-ART and S-ART correlate significantly and highly with each other, students' G-ART cannot fully explain their S-ART levels. The extent to which students display S-ART may not only be a matter of how willing they are to engage in academic risks generally but also of characteristics of the seminar itself, as indicated by the significant correlation coefficient between seminar rating and students' S-ART.

4.3 External and Content Validity

In terms of external validity, students with performance-avoidant goal orientation display lower levels of ART on the group dimensions, which is in alignment with the theoretical assumption that appearing less competent in front of others is a main characteristic of ART (Beghetto, 2009; Clifford, 1991). We find the same relation for the peer dimension, which is not significant in the general context. In line with Dachner et al. (2017), we find that students with mastery goal orientation tend to display higher levels of ART on the group dimensions. Again, the same tendency is found on the peer dimension, which is not significant. The lack of significant correlations on the peer dimension could be caused by the fact that the quality of peer relationships clearly comes into play as an additional source of variance and that goal orientation therefore plays a subordinate role in this context. We did not make any assumptions about the correlation between ART and performance-approach goal orientation and find no association for the group dimensions or for the peer dimension.

We also find a non-significant tendency of students with higher levels of ART achieving higher GPAs. The fact that students' GPAs depend on a variety of individual factors (e.g., intelligence, metacognition) and contextual factors (e.g., type of examination), which may act as confounding variables, can explain the weak nature of the correlation. However, seminar-specific learning success, which we assume to be less confounded than GPA, does correlate significantly with students' ART and thus serves as first indicator that ART may have an influence on academic achievement.

In terms of content validity, our instrument only considers actions with a possibility of failure as these are the core element of ART. Actions and affect following failure as well as individual preferences for task difficulty conceptually play a minor role in ART. Our developed ART scale offers the opportunity to extensively evaluate self-reported, prospective ART, taking into consideration circumstantial and person-specific deviations regarding students' individual perception of risk. The the-

oretical premise of uncertainty about one's contribution as well as the possibility of making mistakes (Beghetto, 2009; Clifford, 1991) is considered in all individual items. Furthermore, the assumption that students perceive risk and therefore behave differently depending on the people who are present and the degree of formality of the situation (Lund Dean & Jolly, 2012; Rohrmann, 2005) is reflected within the two dimensions of the seminar group and peers. Finally, the assumption that ART has both, a state and a trait component (Fredricks et al., 2004) is captured by the differentiation between S-ART and G-ART.

4.4 Limitations

In addition to the aspects already mentioned, our study has some methodological limitations. Since we acquired our sample during the corona pandemic, at a time when German universities had been in partial lockdown for over a year, a replication of the results found under normal teaching conditions is desirable. These circumstances are also partly responsible for the fact that we tested the instrument on a rather small opportunity sample. Within the acquisition process, an open request was made to instructors of the university via an email distribution list, asking them to forward the survey link to their students. Therefore, we cannot make any statements regarding the response rate, as we do not have the necessary information on the number of students reached and to determine the pool of potential participants. Additionally, we limited our sample to students studying within the fields of social sciences and humanities and transferability of the results to students of other study fields such as natural sciences is not verified. Lastly, at the present stage of our research project, a reproduction of the current results with regard to the internal structure of the items was not feasible. Confirming the suggested factorial structure using confirmatory factor analysis within an independent sample is a clear desideratum for future research.

4.5 Implications and Future Research

Since previous research on school students indicates an association between students' ART and their academic success (Bal-İncebacak et al., 2019) and comparable associations were found in the present study, it is of interest to consider ART in research on learning processes also in the context of higher education. For the use of our instrument, we recommend treating and analyzing the peer and seminar group scales individually instead of using an aggregated parameter across all 10 items as a person-specific estimator for ART.

Concerning the research potential of our scale, we see a wide range of important research questions that might be addressed such as whether ART is a stable characteristic of students over time and contexts. Another question concerns possible social and gender disparities. Research on secondary school students shows that

females tend to take less academic risks (Byrnes et al., 1999; Karademir & Akgul, 2019) and that engagement levels are lower in students from lower socioeconomic backgrounds (Tomaszewski et al., 2020), thus enhancing educational inequalities. Whether those disparities persist into higher education may be explored using the presented scales. However, whether ART can be fostered by instructors' teaching behavior or other seminar characteristics and whether this has an influence on students' academic outcomes are questions that are especially relevant.

Since students' ART levels are a factor when it comes to their reported seminar-specific learning success, encouraging students to take more academic risks is a long-term goal that should be pursued. Previous research suggests that psychological safety in the classroom is an important parameter concerning students' ability to learn from their mistakes (Lee, 2020) and that fear of negative feedback inhibits school students' willingness to take academic risks (Üztemur, 2020). Since showing ART implies the ability to handle potential mistakes and negative feedback, we see potential in focusing on psychological safety within the seminar as well as improving students' and instructors' mistake-handling practices and feedback literacy. Sensitizing instructors and students for the learning opportunities that come with making mistakes as well as providing input on how to give and take feedback may improve seminar atmosphere as well as students' confidence and therefore may result in higher ART levels.

Acknowledgement

We would like to thank Dr. Timo Gnambs (Leibniz Institute for Educational Trajectories) for his constructive feedback on an earlier version of this article.

Data availability statement

To ensure the reproducibility of our results for our readers, a minimal dataset, a codebook, and the analysis script were published on the online repository Psych-Archives:

<https://doi.org/10.23668/psycharchives.7080>

<https://doi.org/10.23668/psycharchives.7079>

References

- Abercrombie, S., Bang, H., & Vaughan, A. (2022). Motivational and disciplinary differences in academic risk taking in higher education. *Educational Psychology, 42*(7), 895–912. <https://doi.org/10.1080/01443410.2022.2076810>
- Abercrombie, S., Carbonneau, K. J., & Hushman, C. J. (2022). (Re)Examining academic risk taking: Conceptual structure, antecedents, and relationship to productive

- failure. *Contemporary Educational Psychology*, 68, Article 102029. <https://doi.org/10.1016/j.cedpsych.2021.102029>
- Akbay, S.E., & Delibalta, A. (2020). Academic risk taking behavior in university students: Academic procrastination, academic locus of control, and academic perfectionism. *Eurasian Journal of Educational Research*, 89, 159–178. <https://doi.org/10.14689/ejer.2020.89.8>
- Baker, F.B. (2001). *The basics of item response theory* (2nd ed.). ERIC Clearinghouse on Assessment and Evaluation.
- Bal-İncebacak, B., Yaman, S., & Sarışan-Tunç, A. (2019). The relation between intellectual risk-taking regarding science classes and test anxiety inventory of secondary school. *South African Journal of Education*, 39(1), Article 1670. <https://doi.org/10.15700/saje.v39n1a1670>
- Beghetto, R.A. (2009). Correlates of intellectual risk taking in elementary school science. *Journal of Research in Science Teaching*, 46(2), 210–223. <https://doi.org/10.1002/tea.20270>
- Beghetto, R.A., Karwowski, M., & Reiter-Palmon, R. (2020). Intellectual risk taking: A moderating link between creative confidence and creative behavior? *Psychology of Aesthetics, Creativity, and the Arts*, 15(4), 637–644. <https://doi.org/10.1037/aca0000323>
- Bernaards, C.A., & Jennrich, R.I. (2005). Gradient projection algorithms and software for arbitrary rotation criteria in factor analysis. *Educational and Psychological Measurement*, 65(5), 676–696. <https://doi.org/10.1177/0013164404272507>
- Byrnes, J.P., Miller, D.C., & Schafer, W.D. (1999). Gender differences in risk taking: A meta-analysis. *Psychological Bulletin*, 125(3), 367–383. <https://doi.org/10.1037/0033-2909.125.3.367>
- Çetin, B., İlhan, M., & Yılmaz, F. (2014). An investigation of the relationship between the fear of receiving negative criticism and of taking academic risk through canonical correlation analysis. *Educational Sciences: Theory & Practice*, 14(1), 146–158. <https://doi.org/10.12738/estp.2014.1.1616>
- Clifford, M.M. (1988). Failure tolerance and academic risk-taking in ten- to twelve-year-old students. *British Journal of Educational Psychology*, 58(1), 15–27. <https://doi.org/10.1111/j.2044-8279.1988.tb00875.x>
- Clifford, M.M. (1991). Risk taking: Theoretical, empirical, and educational considerations. *Educational Psychologist*, 26(3–4), 263–297. <https://doi.org/10.1080/00461520.1991.9653135>
- Clifford, M.M., & Chou, F.C. (1991). Effects of payoff and task context on academic risk taking. *Journal of Educational Psychology*, 83(4), 499–507. <https://doi.org/10.1037/0022-0663.83.4.499>
- Clifford, M.M., Chou, F.C., Mao, K.-N., Yun Lan, W., & Kuo, S.-Y. (1990). Academic risk taking, development, and external constraint. *The Journal of Experimental Education*, 59(1), 45–64. <https://doi.org/10.1080/00220973.1990.10806550>
- Dachner, A.M., Miguel, R.F., & Patena, R.A. (2017). Risky business: Understanding student intellectual risk taking in management education. *Journal of Management Education*, 41(3), 415–443. <https://doi.org/10.1177/1052562917695775>
- Duell, N., & Steinberg, L. (2019). Positive risk taking in adolescence. *Child Development Perspectives*, 13(1), 48–52. <https://doi.org/10.1111/cdep.12310>
- Ellis, D.E. (2015). What discourages students from engaging with innovative instructional methods: Creating a barrier framework. *Innovative Higher Education*, 40(2), 111–125. <https://doi.org/10.1007/s10755-014-9304-5>
- Figueira, C., Theodorakopoulos, N., & Caselli, G. (2018). Unveiling faculty conceptions of academic risk taking: A phenomenographic study. *Studies in Higher Education*, 43(8), 1307–1320. <https://doi.org/10.1080/03075079.2016.1250074>
- Fredricks, J.A., Blumenfeld, P.C., & Paris, A.H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59–109. <https://doi.org/10.3102/00346543074001059>

- Karademir, C. A., & Akgul, A. (2019). Students' social studies-oriented academic risk-taking behaviours and autonomous learning skills. *Cypriot Journal of Educational Sciences*, *14*(1), 56–68. <https://doi.org/10.18844/cjes.v14i1.4038>
- Krochmal, A. R., & Roth, T. C. (2017). From comfort to confidence: Modeling science as a process of risk-taking in the classroom. In R. Keltly & B. A. Bunten (Eds.), *Risk-taking in higher education: The importance of negotiating intellectual challenge in the college classroom* (S. 103–118). Rowman & Littlefield.
- Leach, L. (2016). Enhancing student engagement in one institution. *Journal of Further and Higher Education*, *40*(1), 23–47. <https://doi.org/10.1080/0309877X.2013.869565>
- Lee, W.S. (2020). An experimental investigation into the application of a learning-from-mistakes approach among freshmen students. *SAGE Open*, *10*(2). <https://doi.org/10.1177/2158244020931938>
- Leibniz-Institut für Bildungsforschung und Bildungsinformation. (2020a). *Skala: Annäherungsleistungsziele* [Scale: Performance-Approach Goal Orientation]. Retrieved from https://www.fdz-bildung.de/skala.php?erhebung_id=216&skala_id=6692
- Leibniz-Institut für Bildungsforschung und Bildungsinformation. (2020b). *Skala: Lernzielorientierung* [Scale: Mastery Goal Orientation]. Retrieved from https://www.fdz-bildung.de/skala.php?erhebung_id=216&skala_id=6693
- Leibniz-Institut für Bildungsforschung und Bildungsinformation. (2020c). *Skala: Vermeidungsleistungsziele* [Scale: Performance-Avoidance Goal Orientation]. Retrieved from https://www.fdz-bildung.de/skala.php?erhebung_id=216&skala_id=6694
- Lund Dean, K., & Jolly, J.P. (2012). Student identity, disengagement, and learning. *Academy of Management Learning & Education*, *11*(2), 228–243. <https://doi.org/10.5465/amle.2009.0081>
- Özbay, H. E., & Köksal, M.S. (2021). Middle school students' scientific epistemological beliefs, achievements in science and intellectual risk-taking. *Science & Education*, *30*(5), 1233–1252. <https://doi.org/10.1007/s11191-021-00217-y>
- R Core Team. (2020). *R. A language and environment for statistical computing* (Version 3.6.3) [Computer software]. R Foundation for Statistical Computing. <https://www.r-project.org/>
- Ravert, R. D., & Schneller, J. (2019). Risks worth taking: Advisors' views on risks that college students should take more often. *Educational Review*, *71*(6), 792–799. <https://doi.org/10.1080/00131911.2018.1515725>
- Revelle, W. (2021). *psych: Procedures for psychological, psychometric, and personality research* (Version 2.1.9) [Computer software]. <https://CRAN.R-project.org/package=psych>
- Rizopoulos, D. (2006). *ltm: An R package for latent variable modeling and item response theory analysis* [Computer software]. <https://doi.org/10.18637/jss.v017.i05>
- Rohrmann, B. (2005). *Risk attitude scales: Concepts, questionnaires, utilizations: Project report*. Retrieved from <http://www.rohrmannresearch.net/pdfs/rohrmann-ra-report.pdf>
- Teagarden, A., Commer, C., Cooke, A., & Mando, J. (2018). Intellectual risk in the writing classroom: Navigating tensions in educational values and classroom practice. *Composition Studies*, *46*(2), 116–136. <https://www.jstor.org/stable/26642478>
- Tomaszewski, W., Xiang, N., & Western, M. (2020). Student engagement as a mediator of the effects of socio-economic status on academic performance among secondary school students in Australia. *British Educational Research Journal*, *46*(3), 610–630. <https://doi.org/10.1002/berj.3599>
- Üztemur, S. (2020). What if people judge me unfairly: The mediating role of fear of negative evaluation on the relationship between perceived autonomy support and academic risk-taking behaviour in social studies courses. *Journal of International Social Studies*, *10*(1), 62–91.

Appendix

Items of the Academic Risk-Taking Scale – German Language With English Translations in Brackets

Instruktion [instruction]:

Im Nachfolgenden geht es um unterschiedliche Formen der Seminarbeteiligung. Schätzen Sie ein, wie wahrscheinlich folgendes Verhalten in Seminaren generell/in diesem spezifischen Seminar für Sie ist.

[The following statements focus on different forms of seminar participation. Assess how likely you are to engage in the following behavior during seminars in general/in this specific seminar.]

Dimension 1: Seminar Gruppe [Dimension 1: Seminar Group]

1. (*) Mich auch zu schwierigen Inhalten an Semindiskussionen beteiligen. [To participate in seminar discussions even on difficult topics.]
2. (*) Ideen mit dem Plenum teilen, auch wenn sie noch unvollständig sind. [To share ideas with the full seminar group, even if they are still incomplete.]
3. Die Ergebnisse von Gruppenarbeiten unaufgefordert im Plenum vorstellen, auch wenn wir uns bei unseren Ergebnissen nicht sicher sind. [Without being asked, to present the results of small group work in front of the full seminar group, even if we are not sure about our results.]
4. (*) Meine Ergebnisse einer Einzelarbeit unaufgefordert im Plenum vorstellen, auch wenn ich mir bei meinen Ergebnissen nicht sicher bin. [Without being asked, to present my results of individual work to the full seminar group, even if I am not sure about my results.]
5. Im Seminar auf Fragen der Lehrenden antworten, auch wenn ich mir bei meiner Antwort nicht sicher bin. [To answer the instructors' questions in the seminar, even if I am not sure about my answer.]
6. Neue Methoden und Präsentationsmedien in Referaten ausprobieren. [To try out new methods and presentation media during presentations.]
7. Aussagen der Dozierenden im Plenum hinterfragen. [To question statements made by instructors in front of the full seminar group.]
8. (*) Im Plenum zu unklaren Seminarinhalten Fragen stellen, auch wenn ich das Gefühl habe, dass meine Mitstudierenden die Inhalte verstanden haben. [To ask questions about unclear seminar content in front of the full seminar group, even if I feel that my fellow students understand the content.]

9. (*) In Seminardiskussionen eine von der Mehrheit abweichende eigene Meinung vertreten. [During seminar discussions, to express an opinion that differs from that of the majority.]
10. Mich für ein Referat entscheiden, auch wenn die zur Verfügung gestellte Literatur schwierig erscheint. [To choose a presentation topic, even if the literature provided seems difficult.]
11. Referate so gestalten, wie ich es für richtig halte, auch wenn das bedeutet, mich über Vorschläge der Dozierenden hinwegzusetzen. [To prepare presentations as I see fit, even if that means disregarding the instructors' suggestions.]
12. (*) Im Seminar aktiv mitreden, auch wenn ich das Gefühl habe, dass die Inhalte über meinem Kompetenzlevel liegen. [To engage in seminar discussions actively, even if I feel that the content is above my level of competence.]

Dimension 2: Dozierende [Dimension 2: Instructors]

13. Nach Referaten, die meiner Meinung nach hätten besser sein können, Feedback von Dozierenden einfordern. [To ask for feedback from instructors after presentations which, in my opinion, could have been better.]
14. Zu schriftlichen Arbeiten, auf die ich eine schlechte Note erhalten habe, Feedback von Dozierenden einfordern. [To request feedback from instructors on written work for which I have received a poor grade.]
15. Schriftliche Arbeiten so gestalten, wie ich es für richtig halte, auch wenn das bedeutet, mich über Vorschläge der Dozierenden hinwegzusetzen. [To design written work as I see fit, even if this means disregarding instructors' suggestions.]
16. Dozierenden nach der Veranstaltung zu unklaren Seminarinhalten Fragen stellen, auch wenn ich das Gefühl habe, dass meine Mitstudierenden die Inhalte verstanden haben. [To ask instructors questions about unclear seminar content after the course, even if I feel that my fellow students understood the content.]
17. Aussagen der Dozierenden für mich persönlich hinterfragen. [To question statements made by instructors for myself personally.]

Dimension 3: Peers [Dimension 3: Peers]

18. (*) Mitstudierenden auch schriftliche Arbeiten zum Gegenlesen geben, über deren Qualität ich mir unsicher bin. [To ask fellow students to proofread my written work, even though I am unsure about its quality.]
19. (*) Nach Veranstaltungen auch schwierige Seminarinhalte mit Mitstudierenden besprechen. [To discuss difficult seminar content with fellow students after courses.]
20. (*) Mit Mitstudierenden eine Lerngruppe bilden, auch wenn ich das Gefühl habe, weniger zu wissen als die anderen. [To form a study group with fellow students, even if I feel I know less than the others.]

21. (*) Mitstudierenden zu Seminarinhalten, die mir unklar geblieben sind, Fragen stellen. [To ask fellow students questions about seminar content that remained unclear to me.]
22. Gegenüber Mitstudierenden eine abweichende eigene Meinung vertreten. [To represent a dissenting opinion in front of fellow students.]

Anmerkungen. (*) Item wurde in das finale Instrument aufgenommen. Die englische Übersetzung wurde nicht überprüft und dient nur dem sprachlichen Verständnis.

[*Note.* (*) Item selected for the final instrument. The English translation has not been tested and only serves the purpose of linguistic understanding.]

Antwortkategorien [response options]:

Die Items werden auf einer fünfstufigen Likert-Skala von 1 (*sehr unwahrscheinlich*), 2 (*eher unwahrscheinlich*), 3 (*teils, teils*), 4 (*eher wahrscheinlich*) bis 5 (*sehr wahrscheinlich*) beantwortet.

[Items are answered on a 5-point Likert scale from 1 (*very unlikely*), 2 (*rather unlikely*), 3 (*undecided*), 4 (*rather likely*) to 5 (*very likely*).]