The Journal of Navigation (2021), **74**:2 293–310
doi:10.1017/S037346332000051X

RESEARCH ARTICLE

Investigating the correlation between students' perception of authenticity in assessment and their academic achievement in the associated assessment tasks

Samrat Ghosh,[1]* Benjamin Brooks,[1] Dev Ranmuthugala,[1] and Marcus Bowles[2]

[1]National Centre for Ports & Shipping, Australian Maritime College, University of Tasmania, Launceston, Tasmania, Australia.
[2]School of Engineering and Built Environment, Deakin University, Burwood, Victoria, Australia.
*Corresponding author. E-mail: sghosh@utas.edu.au

Received: 14 April 2019; **Accepted:** 12 September 2020; **First published online:** 1 December 2020

Keywords: authenticity, assessment, student, perceptions, seafarer, factor analysis

Abstract

The objective of this research was to investigate the factors of assessment that students undergoing authentic assessment perceived to be significant regarding their academic achievement. This project advanced past research by the authors which found that the academic achievement of seafarer students was significantly higher in a formatively implemented authentic assessment compared with a summative traditional assessment. The academic achievement (assessment scores) was based on the students' performance in analysing information presented in a real-world context (authentic assessment) as opposed to the analysis of information presented devoid of a real-world context (traditional assessment). Using the data obtained from students undergoing the authentic assessment, this project correlated their perceptions of authenticity for factors of assessment to their scores in the associated task. Stage 1 focused on deriving the factors conceptually from the definition of the authentic assessment by the authors, based on which a perception survey questionnaire was designed. Stage 2 extracted new factors through a factor analysis conducted using the software SPSS. Both stages of investigation found that the factor of transparency of criteria was a significant predictor of the students' academic achievement.

1. Introduction

Empirical evidence (Emad and Roth, 2007; AMC, 2011; Maringa, 2015) in past research has suggested that seafarer students tended to disengage with learning and assessment when traditional assessments (oral examinations, written assignments and multiple-choice questions) required them to construct responses purely based on the analysis of information presented, devoid of a real-world context, making them rely solely on their ability to regurgitate memorised information. Disengaged students opted for surface-learning approaches (Maltby and Mackie, 2009) relying on rote learning instead of assimilating and analysing information critically towards preparation for such assessment tasks. For example, one of the ways a seafarer is certified as competent to work onboard commercial ships is through an assessment based on memorised answers in an oral examination. However, the ability to memorise is a lower-level cognition, and memory lapses may lead to unintentional skill and knowledge-based errors (Wiggins, 1989) leading to poor academic achievement. Although one may argue that traditional assessments like oral examinations can also be authentic in particular contexts, Mueller (2006) suggests that they are on the lower end of the continuum of authenticity when they focus on the attributes of recall and regurgitation. Traditional assessment methods adopted in seafarer education are promoted by the Standards of Training, Certification and Watchkeeping (STCW) code that was introduced by

the International Maritime Organization (IMO) in 1978 (revised through major amendments in 1995 and 2010) to provide global, minimum standards of competence assessment. Although traditional assessment methods may be effective in assessing knowledge-based components of a task, they are somewhat decontextualised in nature and make it difficult to provide students with a real-world context for skills and knowledge application (Boud and Falchikov, 2006).

Prior to the STCW code, countries established their own standards. However, STCW did not prove to be as effective as expected due to criticisms from stakeholders that complained of vague and unclear standards left to individual interpretations by maritime nations (Maringa, 2015). As a result, the STCW code was revised with amendments in 1995 (referred to as STCW'95) to address previous concerns and improve upon the training mandate to make it outcome based, requiring seafarer students to demonstrate their competence by performing tasks that resembled workplace duties (Emad and Roth, 2007). STCW'95 did not fully eliminate the vagueness in assessment standards, however, as it specified methods to demonstrate competence but did not provide specific methodologies, leaving these to the discretion of the assessor (Robson, 2007). The code specifies methods (simulator, specialist training etc.) to demonstrate competence but does not provide clear and detailed guidelines on how to use these methods to collect evidence of competence. For example, how sophisticated and advanced should the simulators be to reflect STCW standards? The STCW code only provides recommended performance standards for non-mandatory types of simulators. Even after the latest revision of the STCW code in 2010, its vagueness continues to leave too much room for interpretation by maritime education and training (MET) institutes, which use varying combinations of assessments (Bhardwaj, 2009) for students to demonstrate the performance standards in the STCW code.

Performance standards should ideally communicate performance expectations from workplace duties, encompassing not only the technical skills but also the underpinning skills and knowledge. For example, planned distribution of cargo and recording information (as specified by the STCW code) are not the only skills required for carrying dangerous goods. Assessment should also identify essential underpinning skills, such as problem identification and solving if there are any unexpected occurrences with the carriage. The MET institutes complying strictly with the STCW code will assess seafarers in accordance with inadequate performance standards, producing graduate seafarers lacking workplace skills. This is a major concern for seafarer employers.

In education, authentic assessments appear as a model that integrates knowledge and skills acquired in the classroom with employment, replicating the tasks and performance standards typically faced by professionals in the world of work (Wiggins, 1989), making it suitable for implementation in the vocational-based seafarer education and training. Due to a global absence of evidence regarding the impact of authentic assessment in seafarer education, the authors investigated seafarer students' academic achievement (measured through their assessment scores) in authentic assessment as compared with traditional assessments (Ghosh et al., 2020). However, past researchers (Law and Eckes, 1995; Bailey, 1998, p. 205; Dikli, 2003, p. 16; Abeywickrama, 2012) described traditional assessments as a 'one-shot' or single-occasion tests that are implemented at the end of the learning period in a summative manner. Since authentic assessments are characterised by providing students with more than one opportunity (Wiggins, 1989; Gulikers, 2006), the authors also distinguished the two assessments on the basis of their implementation. The traditional assessment was implemented in a summative format while the authentic assessment was implemented formatively.

Hence, in Ghosh et al. (2020), the authors of this paper investigated the difference in seafarer students' academic achievement (traditional versus authentic) for the unit 'Managerial and Leadership Skills' delivered in the third year of the Bachelor of Nautical Science programme. Students completing this unit acquire the knowledge and skills required by a senior seafarer officer to organise and manage efficient operation onboard a merchant ship. Hence, all the students enrolled in this unit had not only completed two years of the Bachelor programme but also had seagoing experience. The students who enrolled in the unit in Semester 1 were classified as the 'control group' that underwent a traditional assessment. The traditional assessment comprised of two case study scenarios presented and described only on paper in the absence of a real-world context. The students provided written responses on paper

https://doi.org/10.1017/S037346332000051X Published online by Cambridge University Press

to essay-type questions based on their analysis of the described scenarios relying solely on their ability to recall how the scenarios would have played out in the real world onboard ships.

In comparison, another cohort of students who enrolled in the same unit in Semester 2 were assessed authentically through the same case studies described on paper. Although the authentically assessed students also provided written responses on paper to the same essay-type questions, the authentic assessment differed from the traditional assessment by providing a real-world authentic context to the assessment task through a simulation and practical demonstration of the same case study scenarios as employed in the traditional assessment, enacted by staff at the Australian Maritime College (AMC). For example, one case study that described ship staff abandoning the ship using a life raft during a fire was demonstrated at the AMC training pool. The pool was equipped with facilities to launch a real life raft in simulated waves, strong winds, darkness, rain, and smoke. The simulation also included ringing of the emergency alarms and staff playing the role of panicking seafarers jumping into the pool to replicate a possible emergency. In comparison with the authentic assessment, the students who were assessed traditionally relied only on their imagination and experience to visualise the described scenarios.

Although one may argue that the descriptive case studies in themselves (without the simulation) may have provided the real-world contexts, the simulations engaged the sensory perceptions of the students, requiring them to demonstrate the ability to analyse, assimilate, and integrate presented information and construct responses towards it. This was similar to the workplace, where professional seafarers analyse available information and take required action, and thus it distinguished the traditional from the authentic assessment.

In addition to the authentic design, the assessments also differed in the nature of their implementation. The authentic assessments were formative in nature and held on two different days (three weeks apart). The second authentic task was implemented once the students had received individual feedback through the assessment rubric on their performance in the first authentic task. In comparison with the authentic assessment, the traditional assessment was summative in nature and both case studies were implemented at the assessment. However, the duration of the authentic assessment (combined) was the same as that of the traditional assessment. The assessment details and rubric were provided to both the student groups at the beginning of the semester. To avoid the introduction of additional variables, the unit, learning content, lecture delivery methods, lecturer, assessment rubric, total duration of the assessment, and assessment questions were kept constant. The number of completed semesters and academic workloads were the same for both groups. Both the assessments were supervised by external invigilators appointed by the AMC.

The first case study used for assessment required students to respond to a ‘man overboard’ situation at sea and to describe a rescue operation; and the second case study required students to respond to a situation that involved abandoning the ship using a life raft. All students were required to complete the courses ‘Proficiency in Survival Craft’ and ‘Personal Survival Techniques’ to acquire the technical skills required to respond to the emergencies described in the cases studies. Additional to the courses, the students were also trained to acquire the underlying competencies required to respond to the case studies in the unit of ‘Managerial and Leadership Skills’. The syllabus of lectures included training the students to develop their abilities of task and workload management (assigning required personnel towards task implementation, adhering to time and resource constraints, prioritisation of tasks etc.), resource management (considering crew experiences in deciding course of action, recognising barriers to communication etc.), and decision making (conducting situation analysis and risk assessment, obtaining and maintaining situational awareness, and selecting the right course of action).

Findings of the research in Ghosh et al. (2020) confirmed that the seafarer students’ academic achievement was significantly higher (student scores improved by 17·3%) in the formative authentic assessment when compared with the summative traditional assessment. Although in past research higher academic achievement was attributed to the ‘authentic’ design of the assessment and the formative nature of its implementation, further research was required to investigate the factors of assessment that the students may have perceived to be significant and that influenced their perception of authenticity in assessment leading to higher academic achievement. These factors will provide guidance to assessors

https://doi.org/10.1017/S037346332000051X Published online by Cambridge University Press

in the designed authentic assessment with the aim of improving scores and the resulting academic achievement. Hence, using the same but independent sample of authentically assessed students, the research presented in this paper investigated student perceptions of authenticity in assessment to reveal the factors of assessment that correlated significantly with their academic achievement.

As a result, the following research question (RQ) was developed:

RQ: What is the correlation between seafarer students' perception of authenticity in assessment and their academic achievement in the associated assessment tasks?

The developed RQ enabled the development of the following research variables:

(1) independent variable: Perceptions of authenticity in assessment; and
(2) dependent variable: Students' academic achievement.

This research identified seafarer students' 'perception of authenticity in assessment' as the independent variable. The term 'authenticity' in this regard referred to the characteristics (e.g. setting assessment tasks in real-world contexts) of the authentic assessment that students may perceive as significant towards the outcomes of: higher student engagement; ability to transfer skills to different contexts; contextual and multiple evidence of competence; and valid (relevant to workplace) and reliable (multiple and consistent) student performance (Ghosh et al., 2017). The defining characteristics of authentic assessment that led to the aforementioned outcomes were derived from the definition of authentic assessment based on the works of the most commonly cited authors in the area. Based on an extensive literature review (Ghosh et al., 2016, 2017), the authors of this paper used the characteristics of authentic assessment provided by the most commonly cited authors and defined the assessment as one encompassing: *tasks* resulting in outcomes in a real-world context that require an integration of competence to solve forward-looking questions and ill-structured problems; *processes* that require performance criteria to be provided beforehand and evidence of competence to be collected by the student; and *outcomes* that result in valid and reliable student performance, contextual and multiple evidence of competence, higher student engagement, and transfer of skills to different contexts.

The characteristics derived from the definition are summarised in Table 1. Subsequently, the key words (bold in Table 1) in the defining characteristics of authentic assessment were used to develop conceptually the factors of assessment (task, context, criteria etc.). The development of the factors is also shown in Table 1.

Based on the conceptually developed factors (Table 1), this project adapted a questionnaire – mostly from Gulikers (2006) – which was used to obtain student responses regarding their perceptions of authenticity in assessment. In Stage 1, the perceptions of authenticity for the conceptually developed factors were correlated to the dependent variable of students' academic achievement (defined by their composite numeric scores obtained in the authentic assessment tasks). Stage 2 extracted new factors of assessment through a factor analysis. Using the student responses from the perception survey, an additional correlational analysis was conducted between students' perception of authenticity for the new factors of assessment and their scores in the authentic assessment. Both stages of investigation revealed significant findings towards the design of authentic assessments for higher academic achievement of students.

2. Research methodology

2.1. Questionnaire design

This paper used a questionnaire to measure seafarer students' perceptions of authenticity in assessment. To develop the questionnaire, past research in the area of authentic assessment was scanned to investigate if existing published questionnaires and/or items could be used for the purpose. Additionally, an internet search was conducted for the same purpose. The final survey document developed for this research used all the questions from Gulikers (2006) to form Questions 5–27. Since that questionnaire was developed for social work students, the word 'social worker' was replaced with the word 'seafarer' in

https://doi.org/10.1017/S037346332000051X Published online by Cambridge University Press

Table 1. *Defining independent variable to provide conceptually developed factors of assessment for measuring seafarer students' perception of authenticity.*

Independent variable	Defining characteristics	Conceptually developed factors of assessment derived from keywords in the defining characteristics
Perception of 'authenticity' in assessment	Assessment outcomes: Higher student engagement; Ability to transfer skills to different contexts; Contextual and multiple evidence of competence; Valid and reliable student performance.	
Authentic assessment outcomes:		
Higher student engagement	Setting assessment **tasks** in real-world **contexts**; Assessment tasks should be **relevant to the workplace**; Assessment's emphasis on active **construction of knowledge**; Performance **criteria** should reflect workplace needs and be provided beforehand to show **transparency**; **Multiple opportunities** for students to improve learning based on feedback on learning achieved.	Task; Context; Relevance to the workplace; Construction of knowledge; Criteria; Transparency of criteria; Multiple opportunity based on feedback
Ability to transfer skills to different contexts	Setting assessment **tasks** in real-world **context**; Students using **feedback** to identify and fill gaps in competence through **multiple opportunities**.	Task; Context; Multiple opportunity based on feedback
Contextual and multiple evidence of competence	Students provided with **multiple opportunities** to improve learning based on previous **feedback**.	Multiple opportunity based on feedback
Valid and reliable student performance	Assessment tasks should be **relevant to workplace**; **Multiple opportunities** to improve learning based on previous feedback.	Relevance to the workplace. Multiple opportunity based on feedback

https://doi.org/10.1017/S037346332000051X Published online by Cambridge University Press

the questionnaire developed for this project. One question was adopted from the National Survey of Student Engagement (NSSE) to form Question 28a–28e. Two questions were devised by the authors of this paper to form Questions 29 and 30a–30b. The first four questions enquired about the student's demographic details. Questions 5–27 and 29–30 were scored on a five-point Likert scale ranging from 1 (totally disagree) to 5 (totally agree). Only Question 28 was scored on a four-point Likert scale ranging from 1 (very little) to 4 (very much). The Likert scale was reverse coded for negatively worded questions (i.e., Questions 10, 11, 18, 23, 26, 28a). Question 30a required a response on the nominal scale of 'Yes' or 'No'.

2.2. Validity and reliability of the questionnaire

Since the questionnaire constructed for this research was mainly drawn (barring three questions) from Gulikers (2006), it initially derived its validity and reliability from the values published by that author. According to Gulikers (2006), all scales of the survey had a reasonable internal consistency, shown in Cronbach's alpha ranging from 0·63 to 0·83. The Cronbach's alpha for the survey used in this research had a value ranging from 0·69 to 0·75. The adaptation of Gulikers's questionnaire for the purposes of this research study was validated through an expert validation process. The questionnaire was reviewed through a pilot survey by 12 fellow academics and researchers within the AMC, where the research was conducted. The pilot survey respondents suggested retaining most of the original questions but defining the terms 'context', 'criteria', 'oriented', 'under-graduate', post-graduate', and 'output' used in the survey, for the students. The respondents also suggested excluding the demographic question enquiring about the age of the students and including a question related to their educational qualifications.

2.3. Data collection

The survey was administered on completion of the authentic assessments for the treatment group. A general announcement was made in class and an email was sent inviting students to participate in the survey. A minimal risk ethics application approval was obtained for this research project.

2.4. Sampling considerations and response rate

The sampling technique used in this research was based on convenience sampling, which relies on opportunity and participant accessibility and is used when the study population is large, and the research is unable to test every individual (Robson, 2011; Clark, 2014). A key consideration while sampling was to ensure that the treatment group was comprised of randomly assigned students in which each participant had an equal chance of being chosen based only on the sequence of enrolment in the individual semesters. The groups were not sorted based on any other pre-determined characteristics, such as qualifications, academic ability, age or work experience that may have affected the outcomes of this research. This ensured that the relationship between the two variables remained the same in all segments of the sample, which is essential for correlational research (Graziano and Raulin, 2000). Moreover, in correlational research the coefficient of determination (r^2) that allows us to estimate how useful the relationship between the dependent and independent variables might be in a prediction (and is a measure of effect size), should be considered significant only if the minimum sample size is 30 (Graziano and Raulin, 2000; Blondy, 2007; Suresh and Chandrashekara, 2012). This research thus exceeded the recommended minimum sample size. Although 102 students were asked to respond to the survey, just 98 students participated in the study. Out of the 98 respondents, 93 surveys were usable for analysis, as five surveys were discarded due to incomplete/absent responses.

The assessments required students to respond to case study scenarios based on situations that they might encounter on board ships. It was realised that students with work experience may have encountered similar situations and hence were better equipped to answer the questions than students without enough sea experience. Hence, students enrolled in the selected unit (and the respondents) were expected to have

https://doi.org/10.1017/S037346332000051X Published online by Cambridge University Press

completed the minimum work experience of one and half (operational level) to three years (management level) on ships.

2.5. Data analysis

The correlation analysis was conducted in two stages using the statistical software package SPSS 23.

2.5.1. Stage 1: Correlation analysis between students' perception of authenticity in assessment (for factors derived conceptually) and their scores

The questionnaire statements were categorised under the conceptually developed factors of assessment (task, context, criteria, etc.) as set out in Table 1. Questions categorised under a common factor were subjected to an inter-reliability analysis (Cronbach's alpha) to ensure that they were significantly correlated to each other. This is detailed in Table 2.

For the purposes of this paper, a Cronbach's alpha value of greater than 0·70 (Tavakol and Dennick, 2011) was considered statistically significant for reporting. Table 2 showed that an inter-reliability analysis of the categorised survey questions revealed an acceptable Cronbach's alpha value (0·70 or greater) for only two factors of assessment, i.e., relevance to workplace and transparency of criteria. Since an acceptable value of Cronbach's alpha was found for a low number, i.e., only two factors, a correlation analysis between seafarer students' perception of authenticity in authentic assessment for all the developed factors and their scores in the associated assessment task was conducted. The correlation between the variables (perception of authenticity and scores) was considered significant if the correlation coefficient (R) value was higher than 0·25 (Clark, 2014).

The findings of the correlation analysis conducted in stage 1 are discussed in the 'Results' section below.

2.5.2. Stage 2: Correlation analysis between students' perception of authenticity in assessment (for factors extracted through factor analysis) and their scores

Since the majority of the conceptually developed factors of assessment (except transparency of criteria and relevance to workplace) had a low value (less than 0·70) of Cronbach's alpha, a factor analysis to develop new factors of assessment statistically was conducted. Next, a factor analysis to remove multicollinearity and extract factors that are relatively independent of one another was conducted. The survey questions loaded cleanly (without overlap) under the seven new factors. The construction of knowledge questions (28b–28d) clustered in Factor 2, the 'context' questions (question 12–14) item in Factor 4, the 'transparency of criteria' questions (24–26) in Factor 5, and the 'multiple opportunity' questions (29–30b) in Factor 7. Hence, these factors retained the original titles. The questions that were reverse coded clustered in Factor 6, which was therefore titled irrelevant to the profession.

Conversely, the questions related to the conceptually developed factors of relevance to the workplace, task and criteria did not cluster in the expected way; and loaded unevenly (split loading) in Factors 1 and 3. Although a limitation of factor analysis is that factor names may not accurately reflect the variables within the factor, especially in the case of split loadings (Yong and Pearce, 2013), this research used the factor naming technique suggested by Neill (2008). Neill advocated for using the majority of the loading items for naming each factor. The items in Factors 1 and 3 were reviewed to provide meaningful names for the extracted factors based on the top loadings for each factor. Additionally, each factor was subjected to an inter-reliability analysis (Cronbach's alpha) to verify if the values were greater than 0·70. Table 3 details the survey question numbers with their factor loadings, together with the factor titles, and the Cronbach's value of inter-reliability analysis.

Based on the inter-reliability values of Cronbach's alpha, Table 3 revealed that the factor analysis extracted five factors with an acceptable value of more than 0·70. Factors 6 and 7 were rejected due to low Cronbach's alpha values of less than 0·70. The selected factors (1–5) cumulatively explained 60% of the variance in the data, which was considered significant (Williams et al., 2010) for further correlation and regression analysis. Thus, stage 2 investigated the correlation between seafarer students'

https://doi.org/10.1017/S037346332000051X Published online by Cambridge University Press

Table 2. *Survey questions categorised under conceptually developed factors of assessment, and their inter-reliability values.*

Question number	Question statement	Factors of assessment	Cronbach's alpha
5	This assessment was oriented to my future profession as a seafarer.	Relevance to the workplace	0·840
6	This assessment was clearly directed to my professional requirements.		
7	This assessment prepared me for my future profession.		
15	This way of assessing is an effective way of assessing professional skills.		
16	This way of assessing fits well with the seafarer's profession.		
17	The output that I had to produce in this assessment is part of the seafarer's job.		
18	The output that was evaluated in this assessment is different from what is being evaluated in practice.		
19	The result that I had to produce in this assessment is something that a real seafarer also has to produce in practice.		
27	In this assessment, both knowledge and professional skills were important.		
8	The task of the assessment resembled the task of a real seafarer.	Task	0·478
9	The task of this assessment was an important part of the seafarer profession.		
10	The task of this assessment differed from the tasks of a real seafarer.		
11	The context in which I had to perform the assessment was fake.	Context	0·650
12	The context in which I had to perform the assessment looked like a seafarer's workplace.		
13	The context in which I had to perform the assessment looked just like the real world.		
14	The context in which I had to perform the assessment was realistic.		
20	The criteria resembled the criteria that I have to meet in practice.	Criteria	0·547
21	The criteria that I had to meet in this assessment resembled the criteria used in practice.		
22	In this assessment, I was evaluated on criteria important for the seafarer's profession.		
23	In this assessment, I was evaluated on things that I never have to use in real professional practice.		
24	The criteria that I had to meet in this assessment were clear enough.	Transparency of criteria	0·763
25	Before I started the assessment, it was clear to me what was expected of me.		

Continued.

https://doi.org/10.1017/S037346332000051X Published online by Cambridge University Press

Table 2. *Continued.*

26	It was hard to find out what was expected of me in this assessment.		
28	The following requirements of the assessment helped me to improve my score:	Construction of knowledge	0·540
28a	Memorising course material.		
28b	Applying facts, theories, or methods to practical problems or new situations.		
28c	Analysing an idea, experience or line of reasoning in depth by examining its parts.		
28d	Evaluating a point of view, decision or information source.		
28e	Forming a new idea or understanding from various pieces of information.		
29	The feedback provided in this assessment helped me to identify the strengths and weaknesses in my learning.	Multiple opportunity based on feedback	0·697
30a	This assessment provided more than one (1) opportunity to improve my score.		
30b	If 'Yes', the feedback provided on my first performance helped me to improve my assessment score in the next performance.		

Table 3. *Factors extracted using factor analysis: categorised survey questions, titles, and inter-reliability values.*

Factor	Survey questions	Factor title	Cronbach's alpha
Factor 1	5, 6, 7, 8, 9, 17, 22	Relevance to the profession	0·865
Factor 2	28b, 28c, 28d, 28e	Construction of knowledge	0·806
Factor 3	15, 16, 19, 20, 21, 27	Assessing competence to job-relevant criteria	0·868
Factor 4	12, 13, 14	Context	0·732
Factor 5	24, 25, 26	Transparency of criteria	0·763
Factor 6	10, 11, 23	Irrelevant to the profession	0·616
Factor 7	29, 30b	Multiple opportunity	0·697

perceptions of authenticity for the new factors (1–5) of assessment extracted through factor analysis and their scores in the associated assessment task. The correlation between the variables (perception of authenticity and scores) was considered significant if the correlation coefficient (R) value was higher than 0·25 (Clark, 2014). The findings of the correlation analysis conducted in stage 2 are discussed in the 'Results' section below.

3. Results

The results of the RQ are summarised for each stage of investigation in Table 4.

Reporting of the results below is organised by each stage of data analysis.

https://doi.org/10.1017/S037346332000051X Published online by Cambridge University Press

Table 4. *Summary of results.*

RQ	Stage	Results summary
RQ1	Stage 1	Significant correlation found between seafarer students' perceptions of authenticity for the conceptually developed factor of transparency in criteria and their scores in the authentic assessment. Transparency of criteria was also found to be a significant predictor of student scores in the authentic assessment.
RQ1	Stage 2	Significant correlation found between seafarer students' perceptions of authenticity for Factors 2 (construction of knowledge) and 5 (transparency in criteria) extracted through a factor analysis and their scores in the authentic assessment. Factor 5 was also found to be significant predictor of student scores in the authentic assessment.

3.1. Stage 1

The *R*-values for the correlation between the students' perceptions of authenticity (for the conceptually developed factors) in authentic assessment and their scores in the associated assessment task are detailed in Table 5. The *R*-values in Table 5 show significant correlation (*R*-value higher than 0·25 is outlined in bold) between students' perceptions of authenticity for the factor transparency of criteria and their scores in the authentic assessment. Using the significantly correlated factor (transparency of criteria) and the scores in the authentic assessment, a linear regression analysis based on the recommended (Sarkar et al., 2011) confidence level of 95% (or *P*-value 0·05 or less) was conducted. Although confidence levels can be represented as 90%, 95%, 99% or any percentage (between 0 and 100%), the authors of this paper chose the most commonly used confidence level of 95% (Tan and Tan, 2011). The findings of the regression analysis are detailed in Figure 1.

The bold *P*-value (less than 0·05) of the factor transparency of criteria, as shown in Figure 1, revealed the factor to be a significant predictor of student scores in authentic assessment. However, this finding was based on a relatively lower value (8·8%) of the adjusted *R*-square.

3.2. Stage 2

The *R*-values for the correlation between the students' perception of authenticity for the factors of assessment (extracted through factor analysis) and their scores in authentic assessment are detailed in Table 6. The *R*-values in Table 6 show significant correlation (*R*-value higher than 0·25 is outlined in bold) between students' perceptions of authenticity for Factors 2 and 5 and their scores in the authentic assessment. Using the significantly correlated factors (2 and 5) and the scores in the authentic assessment, a multiple regression analysis based on the recommended (Sarkar et al., 2011) confidence level of 95% (or *P*-value 0·05 or less) was conducted. The findings of the regression analysis are detailed in Figure 2.

The bold *P*-value (less than 0·05) of the factor transparency of criteria, as shown in Figure 2, revealed the factor to be a significant predictor of student scores in authentic assessment. However, this finding was based on a relatively lower value (10·4%) of the adjusted *R*-square.

4. Discussion

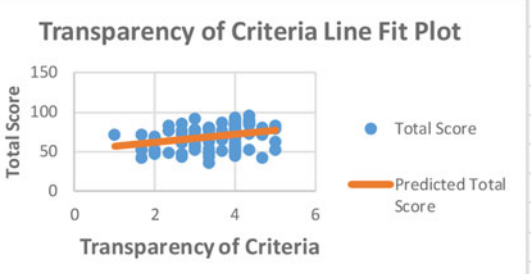
4.1. Transparency of assessment criteria as a significant predictor of academic achievement

Transparency of assessment criteria is essential for learning (Reddy, 2007; Biggs and Tang, 2011) and providing the criteria at the beginning of the learning period (thus making the assessment transparent)

https://doi.org/10.1017/S037346332000051X Published online by Cambridge University Press

Table 5. *R-values of student perceptions of authenticity (conceptually developed factors) in authentic assessment and their scores in the associated assessment task.*

	Relevance	Task	Criteria	Trans Cri	Context	Constn Kn	Mul Opp	Total Score
Relevance	1							
Task	0·7151807	1						
Criteria	0·7434481	0·557436508	1					
Trans Cri	0·5277081	0·357781976	0·57022824	1				
Context	0·5614482	0·595435346	0·430059512	0·252956709	1			
Constn Kn	0·6740382	0·481201314	0·45849741	0·497534833	0·417284643	1		
Mul Opp	0·4850502	0·302732244	0·367019161	0·393821471	0·256092255	0·448560931	1	
Total Score	0·171476	0·194456237	0·136400319	**0·31225956**	0·109658853	0·188522718	−0·022574955	1



SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.3122596
R Square	0.097506
Adjusted R Square	0.0875885
Standard Error	13.967688
Observations	93

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	1918.128195	1918.128195	9.831698956	0.002310037
Residual	91	17753.76428	195.0963108		
Total	92	19671.89247			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	51.299776	6.090641979	8.422720581	5.07168E-13	39.20146347	63.39808764	39.20146347	63.39808764
Trans Cri	5.2648653	1.679086156	3.135554011	**0.002310037**	1.929566951	8.600163712	1.929566951	8.600163712

Figure 1. *Regression analysis of seafarer students' perceptions of authenticity in transparency of criteria and their scores in authentic assessment.*

Table 6. *R-values of students' perceptions of authenticity (factors extracted through factor analysis) in authentic assessment and their scores in the associated assessment task.*

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Total Score
Factor 1	1					
Factor 2	0·51256627	1				
Factor 3	0·778340265	0·608966905	1			
Factor 4	0·464486019	0·349940441	0·470277922	1		
Factor 5	0·418538139	0·503530125	0·566614622	0·190884736	1	
Total Score	0·160202988	**0·296878226**	0·169033928	0·030550485	**0·31225956**	1

is an essential and key requirement for authentic assessment (Wiggins, 1989). Findings of this research project (Figures 1 and 2) confirmed that the factor transparency of criteria is essential for learning, since students had significantly higher achievement when they found the assessment criteria to be transparent. Knowing the assessment criteria (detailing standards of performance) beforehand provided a roadmap of the subject to be learned, while allowing the students to construct their understanding of the topic. This project provided the assessment criteria at the beginning of the learning period through the use of assessment rubrics. Although the rubric was provided to the control as well as the treatment group, the real-world scenarios demonstrated in the authentic assessment allowed the authentically assessed students to analyse the scenarios and construct responses towards the achievement of the standards described in the rubric. For example, when the students were asked to 'recognise all the barriers to effective communication' in the assessment rubric, the authentically assessed students were able to experience the wind and rain that hampered communication in the emergency scenario. In comparison with the authentically assessed students, the traditionally assessed students were unable to recognise the same barriers to communication from the descriptive case studies which lacked the demonstrated scenarios. The authentically assessed students thus used the rubrics to reflect on their learning and carry out self-assessments of their thinking and practices towards achievement of the required standards.

This finding corroborated the findings of the past research (Gulikers, 2006; Jonsson, 2008) where transparency of assessment criteria enhanced student achievement in authentic assessment. Gulikers (2006) also found the transparency of assessment criteria to be the strongest influence on social work students' learning and their development of skills out of other factors such as task and context. Jonsson (2008) focused only on the correlation between transparency of assessment criteria and student scores and revealed that increasing transparency of criteria improved students' performances. The significance of transparency of criteria on student achievement was also found by Hattie (2009) and Hattie and Timperley (2007). To ascertain the major influence on student achievement, Hattie (2009) and Hattie and Timperley (2007) synthesised more than 800 meta-analyses in education and found that making the criteria more explicit leads to improvement of skills since students become more aware of what constitutes a successful performance. Clarity in expectations engages students in the task, which further increases the chance of enhancing their achievement. Hattie (2009) and Hattie and Timperley (2007) also argued that without transparency in assessment criteria, providing students with the feedback on their performance is devoid of context. Feedback directed to the transparent criteria enables students to reduce the gap between their current level of competence and the expected level. Well-directed feedback can then be used by students to adjust their learning strategies towards higher achievement (Hattie and Timperley, 2007; Hattie, 2009).

4.2. Impact of feedback provided to students on their academic achievement

The formative assessment employed in this research project provided students with an opportunity to receive individual feedback on their performance in the first authentic case study (AA_1) before

https://doi.org/10.1017/S037346332000051X Published online by Cambridge University Press

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.351611748
R Square	0.123630821
Adjusted R Square	0.104155951
Standard Error	13.84029554
Observations	93

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	2	2432.052221	1216.026	6.348223	0.002635842
Residual	90	17239.84025	191.5538		
Total	92	19671.89247			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*	*Lower 95.0%*	*Upper 95.0%*
Intercept	45.50989075	6.994083563	6.506913	4.23E-09	31.61492299	59.4048585	31.61492299	59.4048585
Factor 2	3.868048422	2.361499213	1.637963	0.104921	-0.823481943	8.559578786	-0.823481943	8.559578786
Factor 5	3.676608053	1.925711831	1.90922	**0.05942**	-0.149154615	7.502370721	-0.149154615	7.502370721

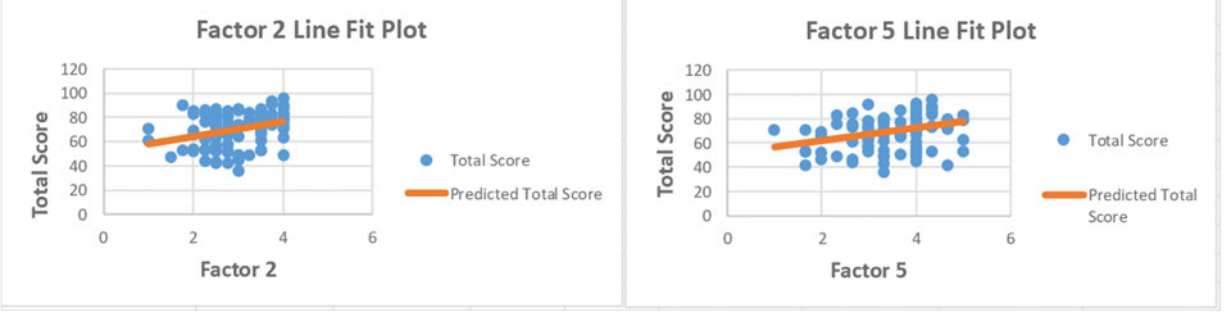


Figure 2. *Regression analysis of seafarer students' perceptions regarding authenticity in Factors 2 and 5 and their scores in authentic assessment.*

attempting the second case study (AA_2). According to Zhang and Zheng (2018), feedback on a student's current ability to perform an assessment task and providing suggestions to improve and to attain their expected levels, encourages the student to take necessary actions to close the gap in their ability. This was confirmed empirically in Ghosh et al. (2020). For example, higher academic achievement in AA_2 (student scores improved by 12%) as compared with AA_1 indicated that using the feedback obtained, seafarer students recognised the gaps in their knowledge, re-evaluated their learning approaches and implemented new strategies to improve their scores. In comparison with the formative assessment, the feedback obtained by the students in the summative traditional assessment task proved to be too late for the control group students to make any adjustments to their learning process to improve their scores.

The positive impact of providing feedback to students on their academic achievement was reaffirmed empirically in this paper, thus advancing past research by the authors (Ghosh et al., 2020). The findings of that research evidenced that higher academic achievement in authentic assessment was not only due to the authentic element of the assessment but could also be attributed to the formative nature of its implementation. The correlation study in this paper confirmed that the group of seafarer students that had significantly higher academic achievement in the authentic assessment perceived the transparency of criteria factor to be the most significant predictor of their achievement (Figures 1 and 2). This also indicated that the seafarer students who underwent formative authentic assessment were able to improve their performance in the second assessment task resulting in an improved academic achievement using the feedback provided to them on their performance in the first task. The individual feedback provided (through the assessment rubrics) enabled the students to conduct a self-assessment of their existing knowledge and skills using the assessment criteria provided at the beginning of the learning period. The students then adopted learning strategies towards obtaining higher academic achievement in the second authentic assessment task.

4.3. Significance of construction of knowledge in authentic assessment

In stage 2, Factor 2 (construction of knowledge) also significantly correlated to the student scores in authentic assessment. A further regression analysis, assuming a 95% confidence level, did not find the factor to be a significant predictor of scores however. If this paper had assumed a 90% confidence level, Factor 2 also would have been considered a significant predictor of students' academic achievement. The choice of whether to use a 90 or 95% confidence interval is somewhat arbitrary (Tan and Tan, 2011), and the 95% confidence level for this research was chosen due to its common use. However, this should not diminish the value of the factor of construction of knowledge and, hence, should be included in designing authentic assessment for students.

4.4. Low value of adjusted R-square

The findings of the regression analysis presented in this paper are based on a relatively low value (8·8% in stage 1 and 10·4% in stage 2) of adjusted *R*-square. The adjusted *R*-square value focuses on explaining the observed variation in the dependent variable due to the independent variable (Lukacs et al., 2010). This implied that the significant factor (transparency of criteria) in this study, although important, did not explain the majority of the variance in the student scores. This was also evidenced by the fact that Factor 1 accounted for the majority of the variance (38·5%) and, did not correlate significantly to the scores. Hence, it was a possibility that the correlation and regression model adopted in this paper may not have included important factors of assessment before measuring the independent variable of perception of authenticity in assessment. For example, factors of assessment such as collaborative assessment (Gulikers, 2006; Ashford-Rowe et al., 2014), student ownership of task design (Gulikers, 2006), completion of task and collation of evidence of competence by students over a sustained period (Morrissey, 2014), and presentation of student work to an audience (Herrington, 1997) were rejected at a theoretical level for the following reasons:

https://doi.org/10.1017/S037346332000051X Published online by Cambridge University Press

- Collaborative assessment was rejected since the research by Gulikers (2006) revealed that students and teachers rated this factor (described as 'social context') as the least important dimension of authentic assessment. Moreover, demonstrating individual competence in the units of learning is essential for seafarer certification (IMO, 2011).
- Factors such as collaborative assessment, student ownership of task design, and completion of task over a sustained period of time were also rejected to avoid plagiarism in student work. This research required seafarer students to complete the assessment task under the supervision of externally employed invigilators. The factors were also rejected since inclusion of these factors in the assessment design would have created uncontrolled additional variables (e.g. variation in student groups, variation in task design and variation in time taken to complete task) other than the authentic design which would have affected student performance.
- The factor requiring presentation of student work to an audience was rejected since it was incongruous to the nature of the assessment task developed for this paper.

The relatively low value of adjusted *R*-square may have resulted also from the use of the quantitative survey to measure student perceptions. This is because the use of Likert scales in the quantitative survey may have limited the students from outlining, describing and adequately conveying the other factors of authentic assessment that they perceive to be significant towards obtaining a higher academic achievement. Instead, the students were compelled to choose the significant factors amongst the choices provided through the survey, which may have led to an inadvertent omission of factors. This was also evidenced by the perception study by Gulikers (2006) in which the quantitative data did not reveal an overall differing perception of authenticity in task, but the qualitative investigation revealed otherwise.

Goodwin and Leech (2006) recommended examining the variability in the data (dependent and independent variable) if the resulting correlation was lower than expected. Lack of variability (indicated through low values of standard deviation) lowers the correlation value between variables (Goodwin and Leech, 2006). To examine the variability, this research calculated the standard deviation values for the student survey responses for perception in authenticity (independent variable) and the composite student scores (dependent variable). The standard deviation for student scores was 14·6 (mean score 69·8/100; minimum 36/100; maximum 96/100). The standard deviation thus indicated a relatively low value, which may have contributed to the lower correlation between the variables. Similar to the dependent variable, the standard deviation values of the student responses to the perception survey had relatively low values, which may also have contributed to the lower correlation.

Lack of variation in student scores indicated evenness in student performance. This may imply that the evenness in performance may have been due to the transparency in assessment criteria that provided all students with the same guidelines to obtain higher academic achievement. This argument is based on past researchers (Black and Wiliam, 1998a, 1998b; Sadler and Good, 2006; Jonsson, 2008), who claimed that transparency in criteria is not only an effective means to improve performance but also a provider of equality in academic achievement. The researchers argued that in studies characterised by formative assessments and transparent criteria, the difference in student achievement between high- and low-performing students is typically reduced.

5. Conclusion

Past research (Ghosh et al., 2020) by the authors found that seafarer students' academic achievement was significantly higher in formatively implemented authentic assessment in which students constructed responses based on the assimilation, integration, and analysis of information presented in real-world settings. This was opposed to a summative traditional assessment that focused on students constructing responses based on memorisation and regurgitation of information. Building on past research findings from Ghosh et al. (2020), the authors investigated factors of authentic assessment (task, context, etc.) that correlated significantly to higher academic achievement (measured using scores obtained in the assessment tasks). Findings derived through factor analysis confirmed that the factor transparency of

https://doi.org/10.1017/S037346332000051X Published online by Cambridge University Press

criteria correlated significantly with student scores. This finding confirmed that providing students with assessment criteria at the beginning of the learning period provides them with clear indications on standards of performance expected in the assessment tasks. Using the feedback provided on their performances in formatively implemented authentic assessment tasks, the students conduct a self-assessment of their learning. Once the gaps in their knowledge and skills are recognised, the students focus on aspects of learning that will improve their performance and overall scores, making them autonomous learners and eventually, skilled professionals.

One may argue that a key limitation of this paper was that the findings of this project were based on a relatively low value of adjusted *R*-square. The adjusted *R*-square value focuses on explaining the observed variation in the dependent variable due to the independent variable. However, the focus of this paper was not to explain variation, but to find an association through correlation between the independent variable (perception of authenticity) and dependent variable (scores). In this context, the adjusted *R*-square value was irrelevant; and the low *R*- square value with statistically significant parameters was more valuable than a high *R*-square value accompanied with statistically insignificant parameters. The researcher acknowledge that a limitation of this project resulted due to the use of the quantitative methodology adopted to enhance generalisability of findings. The quantitative methodology used a survey based on a Likert scale that limited the response of the seafarer students to a perception survey. Hence, certain variables (collaborative assessment and student ownership of task design) may have been rejected at a theoretical level and intentionally omitted from the data analysis model used in this project. Therefore, future research will investigate seafarer students' perceptions through the use of qualitative methodologies, such as interviews, focus groups etc. Although certain variables were excluded, this research uncovered significant factors of assessment which if included, in the design of the assessment, will guide authentically assessed students towards higher academic achievement.

References

Abeywickrama, P. (2012). Rethinking traditional assessment concepts in classroom-based assessment. *The CATESOL Journal*, **23**(1), 205–214.

AMC. (2011). *Report on the Review of the National Centre for Ports & Shipping: Deck Officer Courses*. Launceston, Tasmania: Australian Maritime College.

Ashford-Rowe, K., Herrington, J. and Brown, C. (2014). Establishing the critical elements that determine authentic assessment. *Assessment & Evaluation in Higher Education*, **39**(2), 205–222. doi:10.1080/02602938.2013.819566

Bailey, K. M. (1998). *Learning About Language Assessment: Dilemmas, Decisions, and Directions*. New York: Heinle & Heinle.

Bhardwaj, S. (2009). Quality Maritime Education and Training. In: Loginovsky, V. (ed.) *MET Trends in the XXI Century: Shipping Industry and Training Institutions in the Global Environment – Area of Mutual Interests and Cooperation*, 19–21 September. Saint Petersburg, Russia: Admiral Makarov State Maritime Academy, 29–32.

Biggs, J. and Tang, C. (2011). *Teaching for Quality Learning at University: What the Student Does*. Kuala Lumpur, England: McGraw-Hill.

Black, P. and Wiliam, D. (1998a). Assessment and classroom learning. *Assessment in Education. Principles, Policy & Practice*, **5**(1), 7–74. doi: 10.1080/0969595980050102.

Black, P. and Wiliam, D. (1998b). Inside the black box: raising standards through classroom assessment. *Phi Delta Kappan*, **80**(2), 139–144.

Blondy, L. C. (2007). *A Correlational Study Between the Critical Thinking Skills of Nursing Faculty and Their Perceived Barriers to Teaching Critical Thinking Skills to Nursing Students*. Minneapolis, MN: Capella University.

Boud, D. and Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education*, **31**(4), 399–413. doi:10.1080/02602930600679050

Clark, R. A. (2014). *Correlation Study: The Effect of Student-Teacher Rapport on High School Student Performance Rate*. Lynchburg, VA: Liberty University.

Dikli, S. (2003). Assessment at a distance: traditional vs. alternative assessments. *The Turkish Online Journal of Educational Technology*, **2**(3), 13–19.

Emad, G. and Roth, W. M. (2007). Evaluating the competencies of seafarers: challenges in current practice. In: Pelton, T., Reis, G. and Moore, K. (eds.). *Proceedings of the University of Victoria Faculty of Education Research Conference - Connections'07*, Victoria, BC: University of Victoria, 71–76.

Ghosh, S., Bowles, M., Ranmuthugala, D. and Brooks, B. (2016). Authentic assessment in seafarer education: Using literature review to investigate its' validity and reliability through rubrics. *WMU Journal of Maritime Affairs*, **15**(2), 317–330. doi:10.1007/s13437-015-0094-0

Ghosh, S., Bowles, M., Ranmuthugala, D. and Brooks, B. (2017). Improving the validity and reliability of authentic assessment in seafarer education: A conceptual and practical framework to enhance resulting assessment outcomes. *WMU Journal of Maritime Affairs*, **16**(3), 455–472. doi:10.1007/s13437-017-0129-9

Ghosh, S., Brooks, B., Ranmuthugala, D. and Bowles, M. (2020). Authentic versus traditional assessment: An empirical study investigating the difference in seafarer students' academic achievement. *Journal of Navigation*. Published Online. doi:10.1017/S0373463319000894

Goodwin, L. D. and Leech, N. L. (2006). Understanding correlation: factors that affect the size of r. *The Journal of Experimental Education*, **74**(3), 251–266. doi:10.3200/JEXE.74.3.249-266.

Graziano, A. M. and Raulin, M. L. (2000). *Research Methods: A Process of Inquiry*. New York, USA: A Pearson Education Company.

Gulikers, J. T. M. (2006). *Authenticity Is in the Eye of the Beholder: Beliefs and Perceptions of Authentic Assessment and the Influence on Student Learning*. Heerlen: Open Universiteit Nederland.

Hattie, J. A. C. (2009). *Visible Learning: A Synthesis of Over 800 Meta-Analyses Relating to Achievement*. London: Routledge.

Hattie, J. A. C. and Timperley, H. (2007). The power of feedback. *Review of Educational Research*, **77**(1), 81–112. doi: 10.3102/003465430298487

Herrington, J. A. (1997). *Authentic Learning in Interactive Multimedia Environments*. Perth, Western Australia: Edith Cowan University.

IMO. (2011). *International Convention on Standards of Training, Certification, and Watchkeeping for Seafarers*. London: International Maritime Organization.

Jonsson, A. (2008). *Educative Assessment for/of Teacher Competency. A Study of Assessment and Learning in the 'Interactive Examination' for Student Teachers*. Malmo, Sweden: Malmo University.

Law, B. and Eckes, M. (1995). *Assessment and ESL: On the Yellow Big Road to the Withered of Oz*. Winnipeg, Peguis: Manitoba.

Lukacs, P. M., Burnham, K. P. and Anderson, D. R. (2010). Model selection bias and Freedman's paradox. *Annals of the Institute of Statistical Mathematics*, **62**, 117–125. doi: 10.1007/s10463-009-0234-4

Maltby, A. and Mackie, S. (2009). Virtual learning environments – help or hindrance for the 'disengaged' student? *ALT-J*, **17**(1), 49–62. doi: 10.1080/09687760802657577

Maringa, M. T. H. (2015). *Assessment of Quality of Training and Education of Seafarers in South Africa and Ghana*. Malmo, Sweden: World Maritime University.

Morrissey, P. E. (2014). *Investigating how an Authentic Task can Promote Student Engagement when Learning About Australian History*. Wollongong, NSW, Australia: University of Wollongong.

Mueller, J. (2006). Authentic assessment toolbox. Retrieved from http://jonathan.mueller.faculty.noctrl.edu/toolbox/whatisit.htm#looklike

Neill, J. (2008). *Writing up a Factor Analysis.* Centre for Applied Psychology. Canberra, Australia: University of Canberra.

Reddy, M. Y. (2007). Rubrics and the enhancement of student learning. *Educate*, **7**(1), 3–17.

Robson, C. (2011). *Real World Research*. Chichester, UK: John Wiley & Sons.

Robson, C. S. (2007). Toward an International Rubric: A Compilation of STCW Competency Assessment Methodologies. In: Zhukov, D. (ed.), *8th IAMU Annual General Assembly*. Odesa, Ukraine: Odesa National Maritime Academy, 247–258.

Sadler, P. M. and Good, E. (2006). The impact of self- and peer-grading on student learning. *Educational Assessment*, **11**(1), 1–31. doi: 10.1207/s15326977ea1101-1

Sarkar, E., Keskin, S. and Unver, H. (2011). Using of factor analysis scores in multiple linear regression model for prediction of kernel weight in Ankara walnuts. *The Journal of Animal & Plant Sciences*, **21**(2), 182–185.

Suresh, K. P. and Chandrashekara, S. (2012). Sample size estimation and power analysis for clinical research studies. *Journal of Human Reproductive Sciences*, **8**(3), 186–192. doi: 10.4103/0974-1208.97779

Tan, S. H. and Tan, S. B. (2011). The correct interpretation of confidence intervals. *Proceedings of Singapore Healthcare*, **19**(3), 276–278.

Tavakol, M. and Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, **2**, 53–55. doi: 10.5116/ijme.4dfb.8dfd

Wiggins, G. (1989). A true test: toward more authentic and equitable assessment. *The Phi Delta Kappan*, **70**(9), 703–713. doi: 10.1177/003172171109200721

Williams, B., Brown, T. and Onsman, A. (2010). Exploratory factor analysis: A five-step guide for novices. *Journal of Emergency Primary Health Care*, **8**(3), 1–13.

Yong, A. G. and Pearce, S. (2013). A beginner's guide to factor analysis: focusing on exploratory factor analysis. *Tutorials in Quantitative Methods for Psychology*, **9**(2), 79–94. doi: 10.20982/tqmp.09.2.p079

Zhang, L. and Zheng, Y. (2018). Feedback as an assessment for learning tool: How useful can it be?, **43**(7), 1120–1132. doi: 10.1080/02602938.2018.1434481

Cite this article: Ghosh S, Brooks B, Ranmuthugala D, Bowles M (2021). Investigating the correlation between students' perception of authenticity in assessment and their academic achievement in the associated assessment tasks. *The Journal of Navigation* **74**: 2, 293–310. https://doi.org/10.1017/S037346332000051X