Personality and Individual Differences 50 (2011) 259–263

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

How would you like to be evaluated? The correlates of students' preferences for assessment methods

Adrian Furnham [a,*], Mark Batey [b], Neil Martin [c]

[a] Psychology Department, University College London, UK
[b] Department of Organisational Psychology, Manchester Business School, UK
[c] Department of Psychology, University of Middlesex, UK

ARTICLE INFO

Article history:
Received 28 April 2010
Received in revised form 23 September 2010
Accepted 29 September 2010
Available online 28 October 2010

Keywords:
Assessment
Personality
Intelligence

ABSTRACT

This study is concerned with individual correlates of students' preferences for different examination methods. The aim was to examine the incremental validity of learning approaches, intelligence and personality (Big Five traits) in predicting the preferences for six different assessment methods. Four hundred and ten British students from three British Universities completed a general intelligence test, a personality test, a measure on their Approach to Learning and a University assessment method preference test that specified six different methods. Students favoured multiple choice and continuous assessment most and Oral and Group work least. Deep learning style and Openness to Experience were the most consistent predictors of assessment preferences. The three sets of measures accounted for between 9% and 29% of the variance in preferences for different methods. Specific learning approaches and personality traits, but not intelligence is clearly linked to preferences for assessment.

1. Introduction

Over the years various studies have examined scholar and student preferences for such things as the personality of their lecturers or teachers (Furnham & Chamorro-Premuzic, 2005; Swami et al., 2007); their teaching method (Chamorro-Premuzic, Furnham, & Lewis, 2007) and their assessment preferences (van de Watering, Gijbels, Dochy, & van der Rijt, 2008). These have been linked to studies looking at individual differences correlates of academic performance (Mellanby & Zimdars, in press; Van Bragt, Bakx, Bergen, & Croon, in press). Studies have been done on teachers (Birgin & Baki, 2009) and school children (Peterson & Irving, 2008). This study concerns students' preferences for their method of assessment in relation to general intelligence, personality and Approaches to Learning as assessed by the Study Process Questionnaire (SPQ; Biggs, 1987; Fox, McManus, & Winder, 2001).

Student assessment method has been an active field of educational research and application (Meyer et al., 2010) in many countries (Ismail, Hussain, & Jamaluddin, 2010; Tait, 2010). It has seen changes in ideas about the assessment *of* learning to the assessment *for* learning. Further, new forms of assessment have been proposed like peer or self-assessment (Scouller & Prosser, 1994). There are also numerous studies on issues such as preferences for particular teaching methods (Apperson, Laws, & Scepansky, 2008; Askell-Williams & Lawson, 2001; Cowman, 1995).

There is now an extensive and growing literature on students' perceptions about assessment and evaluation in higher education (Birenbaum, 1997; Kniveton, 1996; Marlin, 1987). Much of this literature has concentrated on one particular assessment technique namely multiple choice questions (MCQ's) (Alker, Carlson, & Hermann, 1969) and one individual difference factor namely Approaches to Learning. Thus, Scouller and Prosser (1994) showed deep and achievement oriented students had different perceptions of, and preparation strategies for, MCQs. Scouller (1998) later showed students tended to employ a surface learning approach when examined by MCQ and to perceive them as assessing knowledge based lower levels of intellectual processing. Indeed Deep Learning approaches were associated with poorer performance in MCQs but the opposite in assignment essays. Birenbaum (2007) looked at students' test anxiety and learning strategies influence on their preference for type of instruction/teaching and assessment. He found evidence of a learning regulation dimension from self-regulation to instructor dependency and concluded the need for a dialogue between instructor and students to structure expectations to fit the goals of higher education.

Others have looked at the possibility of using peer assessment. Segers and Dochy (2001) compared students' own, their peers', and their tutor's assessments with examination scores. Interestingly, the highest correlations were between peer ratings and the examination score.

Struyven, Dochy, and Janssens (2005) reviewed 35 salient papers on the topic. They concluded that a students' Approach to Learning is a clear and logical correlate of assessment preferences

* Corresponding author.
E-mail address: a.furnham@ucl.ac.uk (A. Furnham).

0191-8869/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.paid.2010.09.040

and that assessment methods have an impact on the students' learning approach and vice versa. They noted that studies have focused on essay vs MCQ's examinations and that students low in test taking anxiety and good learning skills prefer essays over MCQs, but that high anxious, poor learning skill students prefer MCQs, which elicit surface Approaches to Learning.

There appear to have been three salient and recent studies in the area of student preference for assessment method. Furnham and Chamorro-Premuzic (2005) in the first paper in the area conducted a two study examination of ability (IQ) and personality (Big Five) correlates of preference for assessment method. They found in their first study ($N = 103$) modest, but interpretable, correlations between ability test scores, personality trait measures and assessment preferences. The second study ($N = 93$) replicated some of the results. They found that the students' academic performance had no relationship to their preference for assessment methods. Chamorro-Premuzic, Furnham, Dissou, and Heaven (2005) in a study of Australian University students ($N = 125$) looked at self-assessed intelligence and Big Five personality trait predictors of seven assessment-method preferences. As before, the idea of being assessed by viva (Oral examination) seemed the most closely linked to personality, in particular Extraversion. Self-assessed intelligence seemed unrelated to preference for assessment methods.

More recently Furnham, Christopher, Garwood, and Martin (2008) reported a study on over 400 British and American students' preferences for assessment method. They examined Approaches to Learning using the SPQ and Big Five personality correlates of preferences. They found Surface Learning was associated with a preference for Multiple-Choice tests and Group work exercises, but a disinclination to be assessed by traditional exam Essays or Dissertations. Deep learners on the other hand favoured assessment by essay, oral exam and dissertation. They also found Extraverts preferred Multiple Choice, Oral examinations and Group-work while Conscientious people showed a preference for course work, but not Group work. Openness-to-Experience was related to a preference for Oral examination and Essays, but not Multiple-Choice and Group work. However, of the twelve regressions conducted, only three showed the individual difference predictor variables to account for >10% of the variance.

This study extends this research programme by doing three things. First, it includes a measure of intelligence to see if it provides incremental validity over preference measures in assessment preferences. Second, this study has a more heterogeneous and larger student population to see if more variance can be accounted for. Third, by examining the incremental validity of the three sets of variables considered, namely Approaches to Learning, intelligence and personality traits.

There are numerous studies showing that the three predictor variables alone or together try to predict academic outcomes (Furnham, 2010, 2011) but far less studies on preferences.

This study is predominantly concerned with how much variance the three individual difference variables can account for in each of the preference methods, particularly MMQ and whether intelligence provides incremental validity over personality traits and Approaches to Learning.

2. Method

2.1. Participants

In all, 410 mainly social science students took part in this study of which 108 were male. Their average age was 20.45 years (SD = 5.06). They were drawn from four British Universities. All students were fluent English speakers and provided their informed consent to participate in the study.

2.2. Materials

Choice of Assessment Method (Furnham & Chamorro-Premuzic, 2005): Preference for assessment methods were assessed by this six item questionnaire, where participants indicated how much (1 = Not at all, to 7 = Very much) they would like particular assessment methods to be used to evaluate their final, overall, university degree. The items covered: Multiple choice exams; Essay-type exams (e.g., answer two questions in 2 h); Final year dissertation (e.g., supervised thesis, 4-month research project); Oral examination (viva voce); Continuous assessment (e.g., participation in class, attendance, course work, etc.) and Group work (team assignment, one mark for everybody in the group). Versions of this measure have been used in various other studies (Swami et al., 2007).

Approaches to Learning were assessed by the Study Process Questionnaire (SPQ; Biggs, 1987; Fox et al., 2001). This inventory is used to assess three major Approaches to Learning- Surface, Deep, and Achieving. The SPQ comprises 42 items. Participants' responses were recorded on a five-point Likert-type scale. Psychometric information regarding the reliability and validity of SPQ has been reported elsewhere (Biggs, 1987, 1993; Fox et al., 2001; Zhang, 2004). In the present data set, all Cronbach's alphas for the SPQ factors were higher than .67, indicating acceptable internal consistency.

Personality was assessed using the NEO-FFI (Costa & McCrae, 1992). This inventory is a short 60 item version of the NEO-PI-R and assesses the personality dimensions of Neuroticism (low emotional stability), Extraversion, and Openness to Experience, Agreeableness, and Conscientiousness. Scores range from 12 to 60 for each trait. Hundreds of studies have used this inventory as it demonstrates good internal and external validity (Costa & McCrae, 1992). The inventory has been shown to be a satisfactory tool for assessing relationships between personality and a number of academic variables such as intellectual satisfaction, self-esteem, and teaching effectiveness. Items ask about typical behaviors or reactions and are answered on a five-point Likert-type scale, ranging from "strongly disagree" to "strongly agree".

General Intelligence was assessed by The Wonderlic Personnel Test (Wonderlic, 1992). This 50-item test is administered in 12 min. Scores can range from 0 to 50 with a population mean of 21. Items include word and number comparisons, disarranged sentences, serial analysis of geometric figures and story problems that require mathematical and logical solutions. The test has impressive norms and correlates very highly ($r = .92$) with WAIS-R.

2.3. Procedure

Students completed the measures in a large, quiet, invigilated lecture theatre. Students completed the measure of general intelligence under timed conditions, before completing the remaining self-report measures in their own time.

3. Results

3.1. Descriptive statistics

The means and standard deviations of the variables included in this study are presented in Table 1. In order to check for gender or age differences in the preferred assessment methods an ANOVA was conducted. There were no sex differences or significant age differences on the six assessment preference choices.

Table 1
Mean and SD Preferences for the six assessment methods and variables utilized in the Study ($n = 410$).

		Mean	SD
1	Multiple Choice	5.10	1.55
2	Essay type	4.32	1.70
3	Final year dissertation	3.98	1.68
4	Oral exam	3.55	1.79
5	Continuous assessment	4.80	1.59
6	Group work	3.57	1.68
7	Surface learning	45.07	7.34
8	Deep learning	45.70	7.40
9	Achieving learning	45.99	7.62
10	Neuroticism	34.89	8.61
11	Extraversion	42.26	6.19
12	Openness	40.90	6.28
13	Agreeableness	29.45	5.96
14	Conscientiousness	43.08	6.88
15	Intelligence (WPT)	28.77	6.01

Note: WPT, Wonderlic Personnel Test.

Table 1 shows the mean scores. By far the strongest preference was for multiple choice questions while the students seemed strongly disinclined to have Oral or Group work assessment. The means for the other tests were well within the normal range for this population as given by test manuals.

Table 2 shows correlations and regression results for each of the six preferences. Those with lower Deep, but higher Surface approaches and lower Openness scores tended to favour MCQ for assessment. The opposite pattern was true for assessment by essays: Higher Deep, lower Surface and higher Openness scores predicted preference for essays. Those who favoured dissertations as an assessment method tended to more Open, less Neurotic and favour Deep, but not Surface, Approaches to Learning. Stable, Open, Extraverts with high Deep and low Surface score showed a preference for oral exams. Conscientious Extraverts and those with a Deep and Achieving approach favoured continuous assessment. Finally, the correlations showed that less intelligent, Open and Conscientious students with a low Achieving approach favoured Group work.

Table 2 also shows the results for the six regressions. These results are more important because they show how much of the variance is accounted for by all the independent variables together. Furthermore, it has been shown that there is considerable overlap between the Big Five and Approaches to Learning (Chamorro-Premuzic & Furnham, 2009), which leads to problems of multicollinearity. The predictor variables were Approaches to Learning, intelligence, and personality put in that order. Regressions were done in three steps to examine the incremental validity of the three blocks of variables.

The results showed: *Multiple choice*: Bright, less Open candidates prefer this method and those factors accounted for over a fifth of the variance; *Essays*: Open to Experience, high on a Deep approach students preferred essays and sixteen percent of the variance was accounted for; *Dissertation*: Open to Experience students with a Deep, but not Surface approach preferred dissertations and these accounted for nearly thirty percent of the variance, *Oral/Viva*: Stable, low Conscientious students with a Deep learning style preferred oral exams. Together these factors accounted for just under quarter of the variance; *Continuous Assessment*: the regressions were not significant; *Group work*: Deep learning students and those low in Openness preferred this method.

4. Discussion

This study was part replication and part extension of previous studies. Most students appeared not to like Oral exams or Group work and favoured multiple choice and continuous assessment methods. It confirmed Furnham and Chamorro-Premuzic (2005) findings that intelligence was a modest predictor of assessment choice and confirmed Furnham et al. (2008) results that Approach to Learning and personality were predictable correlates of those preferences.

This study looked at the incremental validity of Approaches to Learning, intelligence and personality. Further, in contrast to previous studies that seem to account for <10% in accounting for these preferences, this study showed that these factors accounted for over twenty percent of the variance for three of the six preferences. Most importantly, it demonstrated that while Approaches to Learning could account for around a tenth of the variance, in two preference judgments, personality variables could double the variance thereby explaining as much as a quarter of the variance. Further, it showed that intelligence adds very little (almost nothing) as a predictor of assessment choice when considered along with Approaches to Learning and personality. This cannot be attributed to a restriction of range as Table 1 shows.

It is particularly interesting that bright, but less imaginative, Surface learners preferred multiple choice exams which can differ

Table 2
Correlational and regressional results with assessment choice as criterion variable ($n = 410$).

Predictor	Multiple choice			Essay			Dissertation			Oral			Continuous assessment			Group work		
	r	Beta	*t*	*r*	Beta	*t*	*r*	Beta	*t*	*r*	Beta	*t*	*r*	Beta	*t*	*r*	Beta	*t*
Surface	.22**	.14	1.40	−.14**	.05	.00	−.34**	−.36	3.82***	−.17**	−.07	.72	.00	−.10	0.94	.03	.17	1.59
Deep	−.13**	−.10	0.97	.27**	.26	2.53*	.24**	.12	1.29	.22**	.27	2.65**	.14**	.16	1.45	.00	.24	2.29*
Achieving	.07	−.05	0.45	.06	−.03	0.26	−.02	.08	0.70	.00	.11	0.99	.15*	.22	1.77	−.11*	−.15	1.24
IQ (WPT)	.10	.21	2.57*	.00	−.11	1.32	.04	−.06	0.75	−.02	−.03	0.41	.02	.02	0.27	−.15*	.07	0.83
Neuroticism	.05	−.03	0.32	−.07	−.10	1.01	−.15**	−.00	0.09	−.16**	−.24	2.58**	.05	.05	0.48	−.09	−.12	1.27
Extraversion	.00	.14	1.49	.03	−.06	0.62	.00	−.07	0.80	.19**	.15	1.70	.14*	.07	0.71	.12*	.10	1.11
Openness	−.27**	−.31	3.09**	.25**	.20	1.97*	.31**	.21	2.19*	.14**	.01	0.13	−.04	−.16	1.52	−.12*	−.24	2.27*
Agreeable	.05	.03	0.36	.06	.06	0.66	−.03	−.03	0.37	.03	.11	1.34	−.09	−.04	0.49	−.03	.01	1.13
Conscient.	.02	.01	0.11	.09	.00	0.57	.04	.03	0.04	−.01	−.27	2.46**	.11**	−.12	1.02	−.17*	.19	1.66
Step	*F*	R^2		*F*	R^2		*F*	R^2		*F*	R^2		*F*	R^2		*F*	R^2	
1	4.41**	.09		6.12**	.12		15.20***	.25		5.58***	.12		2.79*	.06		2.43	.05	
2	4.56**	.12*		4.91**	.13		11.40***	.25		4.61**	.12		2.08	.06		2.26*	.06	
3	3.84**	.21**		2.71**	.16		5.80***	.29		4.23***	.23		1.46	.09		2.47*	.15	

Note: WPT, Wonderlic Personnel Test; Openness, Openness to Experience; Agreeable, Agreeableness; Conscient., Conscientiousness.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

considerably in terms of whether they primarily involve memory for facts or subtle understanding of complex issues. These exams are used extensively in some countries (i.e. the USA) and rarely in others (i.e. the UK) and are favoured by lecturers for their ease of administration and reliability of marking. Usually there are low correlations between cognitive ability and Openness so it is surprising that these variables work in opposite directions in this regression.

Most of the participants in this study would have been used to being primarily examined by the two or three hour written exam where they are required to write three to five essays answering a choice of between six and ten questions. Imaginative, Deep learners prefer this traditional method, though a preference for this method was not at all correlated with intelligence. Presumably they feel able to exhibit and benefit from their wide reading and attempts to really understand problems. It is also common, at least in a final year to do a dissertation. Open students liked this, but those with Surface learning approaches did not, no doubt because they are nearly always in-depth studies with critical analysis of issues.

At post-graduate level, universities begin to introduce oral exams. Indeed most doctorates are examined by viva such that a good written PhD can yet be failed by very poor performance in a viva (and vice versa). This regression accounted for most of the variance and five of the correlations were significant, though preferences were quite unrelated to intelligence. The regression showed that Stable, Deep learners low on Conscientiousness preferred this method. Students often report how anxiety-provoking oral exams can be and this no doubt accounts for the relationship of Neuroticism with a dislike of oral exams.

Neither of the analyses on continuous assessment or Group work showed many effects. This may reflect the rarity of the utilization of these assessment methods.

It is no surprise that Stable Extraverts favour oral exams over Neurotic Introverts or that Conscientious students like continuous assessment. These methods speak to their strengths. Indeed, students have been known to choose particular courses as a function of how they are assessed.

Teachers worry about the reliability, fairness and accuracy of assessment methods as well as the amount of work involved. Indeed the method they choose may be a function of their abilities and preferences. Students likewise, get very concerned about how fair assessment is, and whether they are dispositionally better suited to some methods more than others. Further, some subjects and teaching styles lend themselves to some assessment methods better than others.

There is considerable debate about the stability of personality traits and their mutability. It is, however, assumed that Approaches to Learning are relatively easily modifiable and to some extent a function of the teaching philosophy of educational institutions. Thus, to attempt to change preferences for, as well as attitudes to, particular assessment methods (i.e. MCQ) it seems logical to target Approaches to Learning which accounts for a reasonable amount of the variance.

Like all studies this had limitations. Although the sample was of reasonable size and drawn from a variety of institutions, most participants were social scientists who may have themselves a rather narrow range of examination preferences. Thus, mathematics or engineering students may make rather different choices. Further, some courses lend themselves to different assessment techniques more readily than others. There is also considerable variety within an assessment method: thus the number of options in an MCQ could influence preferences as well as whether marks are deducted for incorrect answers.

Future research in the area would do well to examine such things as whether students' scores differ as a function of the variables examined in this study, if they did *different* types of tests in the *same* course. Most degrees involve different types of assessment and it may well be that intra-individual variations in scores across a course where different methods are used are related systematically to individual differences.

References

Alker, H., Carlson, J., & Hermann, M. (1969). Multiple-choice questions and student characteristics. *Journal of Educational Psychology, 60*, 231–243.

Apperson, J., Laws, E., & Scepansky, J. (2008). An assessment of student preferences for powerpoint presentation structure in undergraduate courses. *Computers and Education, 50*, 148–153.

Askell-Williams, H., & Lawson, M. (2001). Mapping students' perceptions of interesting class lessons. *Social Psychology of Education, 5*, 127–147.

Biggs, J. (1987). *The Study Process Questionnaire manual*. Victoria: Australian Council for Educational Research.

Biggs, J. (1993). What do inventories of students' learning processes really measure? *British Journal of Educational Psychology, 3*, 3–19.

Birenbaum, M. (1997). Assessment preferences and their relationship to learning strategies and orientations. *Higher Education, 33*, 71–84.

Birenbaum, M. (2007). Assessment and instruction preferences and their relationship with test anxiety and learning strategies. *Higher Education, 53*, 749–768.

Birgin, O., & Baki, A. (2009). An investigation of primary school teachers' proficiency perceptions about measurement and assessment methods. *Procedia Social and Behavioral Sciences, 1*, 681–685.

Chamorro-Premuzic, T., & Furnham, A. (2009). Mainly openness: The relationship between the Big Five personality traits and learning approaches. *Learning and Individual Differences, 19*, 524–529.

Chamorro-Premuzic, T., Furnham, A., Dissou, G., & Heaven, P. (2005). Personality and preference for academic assessment. *Learning and Individual Differences, 15*, 247–256.

Chamorro-Premuzic, T., Furnham, A., & Lewis, M. (2007). Personality and Approaches to Learning predict preference for different teaching methods. *Learning and Individual Differences, 17*, 241–250.

Costa, P., & McCrae, R. (1992). *Revised NEO personality inventory (NEO-PI-R) and NEO five factor inventory (NEO-FFI): Professional manual*. Odessa: FL: Psychological Assessment Resources.

Cowman, S. (1995). The teaching/learning preferences of student nurses in the Republic of Ireland. *International Journal of Nursing Studies, 32*, 126–136.

Fox, R., McManus, I. C., & Winder, B. (2001). The shortened Study Process Questionnaire. *British Journal of Educational Psychology, 71*, 511–530.

Furnham, A. (2010). Learning Styles and Approaches to Learning. In T. Urdan, J. Roger, S. Graham, & M. Zeidner (Eds.), *APA educational psychology handbook*. Washington, DC: APA.

Furnham, A. (2011). Personality and Approaches to Learning. In S. Von Stumm, T. Chamorro-Premuzic, & A. Furnham (Eds.), *Handbook of individual differences*. Oxford: Wiley-Blackwell.

Furnham, A., & Chamorro-Premuzic, T. (2005). Individual differences and beliefs concerning preferences for university assessment methods. *Journal of Applied Social Psychology, 35*, 1986–1994.

Furnham, A., Christopher, A., Garwood, J., & Martin, N. (2008). Ability, demography, learning style and personality trait correlates of student preferences for assessment method. *Educational Psychology, 28*, 15–27.

Ismail, A., Hussain, R., & Jamaluddin, S. (2010). Assessment of students' learning styles preferences in the faculty of science, Tishreen University, Syria. *Procedia, Social and Behavioral Sciences, 2*, 4087–4091.

Kniveton, B. (1996). Student perceptions of assessment methods. *Assessment and Evaluations in Higher Education, 21*, 229–238.

Marlin, J. (1987). Student perception of end-of-course-evaluations. *Journal of Higher Education, 58*, 704–716.

Mellanby, J., & Zimdars, A. (in press). Trait anxiety and final degree performance at the University of Oxford. *Higher Education.*

Meyer, L., Davidson, S., McKenzie, L., Rees, M., Anderson, H., Flectcher, Johnson., et al. (2010). An investigation of tertiary assessment policy and practice. *Higher Education Quarterly, 64*, 331–350.

Peterson, E., & Irving, S. (2008). Secondary school students' conceptions of assessment and feedback. *Learning and Instruction, 18*, 238–250.

Scouller, K. (1998). The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher Education, 35*, 453–472.

Scouller, K., & Prosser, M. (1994). Student's experiences in studying for multiple choice question examinations. *Studies in Higher Education, 19*, 267–279.

Segers, M., & Dochy, F. (2001). New assessment forms in problem-based learning. *Studies in Higher Education, 26*, 327–343.

Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education:0020A review. *Assessment and Evaluation in Higher Education, 30*, 325–341.

Swami, V., Furnham, A., Maakup, I., Ahmad, S., Hudani, N., Voo, P., et al. (2007). A 'cross cultural investigation of students' preferences for lecturers' personalities in Britain, Malaysia and the United States. *Learning and Individual Differences, 17*, 307–315.

Tait, C. (2010). Chinese students' perceptions of the effects of western university examination formats on their learning. *Higher Education Quarterly, 64*, 261–275.
Van Bragt, C., Bakx, A., Bergen, T., & Croon, M. (in press). Looking for students' personal characteristics predicting study outcome. *Higher Education.*
van de Watering Gijbels, D., Dochy, F., & van der Rijt, J. (2008). Students' assessment and their relationships to study results. *Higher Education, 56*, 645–658.
Wonderlic, E. (1992). *Wonderlic Personnel Test*. Libertyville, Illinois: Wonderlic.
Zhang, L.-F. (2004). Learning approaches and career personality types. *Personality and Individual Differences, 37*, 65–81.