Learning and Individual Differences 50 (2016) 1–10

Contents lists available at ScienceDirect

Learning and Individual Differences

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

Does trait test anxiety compromise the measurement fairness of high-stakes scholastic achievement tests?

Markus Sommer, Martin E. Arendasy *

University of Graz, Austria

ARTICLE INFO

Article history:
Received 16 July 2015
Received in revised form 25 May 2016
Accepted 30 June 2016
Available online xxxx

Keywords:
Admission testing
Trait test anxiety
Interference hypothesis
Deficit hypothesis
Measurement bias

ABSTRACT

Based on meta-analytic evidence of a moderate negative correlation between trait test anxiety and test performance some researchers hypothesized that trait test anxiety may induce measurement bias. Two competing models have been advanced to account for the test anxiety–test performance relationship: the deficit hypothesis and the interference hypothesis. The interference hypothesis predicts that trait test anxiety induces measurement bias in items of intermediate difficulty, while the deficit hypothesis claims that test anxiety has no causal effect on test performance. In the present study we tested these competing predictions in a high-stakes setting by means of structural equation modeling. Test-takers ($N = 1768$) solved a knowledge test as part of an admission test and filled a questionnaire measuring trait worry, trait task-irrelevant thinking and trait emotionality. In line with the deficit hypothesis the present results indicated that neither of the three trait test anxiety components induced measurement bias.

1. Introduction

Admission tests have important consequences for the later life trajectories of test-takers and there is evidence that test-takers differ considerably in the way they approach and respond to this challenge (cf. Strack & Esteves, 2015). While some test-takers seem to be relatively unbothered others tend to react to the challenge of admission tests with an increase in test anxiety. The existence of individual differences in test-takers' appraisal of admission tests raises the question whether test anxiety has a causal influence on test performance and therefore threatens the measurement fairness and construct validity of admission tests.

Based on the meta-analytic evidence of a moderate negative (meta-analytic mean $r = -0.18$ to $r = -0.24$) correlation between test anxiety and performance in scholastic achievement measures (cf. Hembree, 1988; Richardson, Abraham, & Bond, 2012; Seipp, 1991) some researchers (e.g. Hembree, 1988) hypothesized that test anxiety may prevent more test anxious test-takers from performing at their true level of ability. This implies that scholastic achievement test scores reflect individual differences in construct-irrelevant factors (here: trait test anxiety) in addition to individual differences in the latent trait(s) of interest. If this would be the case, the admission test would be biased because more test anxious test-takers with the same level of ability than their less anxious counterparts would have lower expected test scores (cf. Drasgow, 1987; Millsap, 2011; Mislevy et al., 2013). Despite the theoretical and practical relevance of this hypothesis only few studies (Halpin, da-Silva, & De Boeck, 2014; Reeve & Bonaccio, 2008; Sommer & Arendasy, 2014) directly tested this prediction. Unfortunately only one of these studies (Halpin et al., 2014) used a scholastic achievement test. Furthermore, all three studies were conducted in a low-stakes assessment situation. These two characteristics may limit the generalizability of previous research findings. Therefore the present study aims to examine, whether individual differences in trait test anxiety induce measurement bias in a medical school scholastic achievement admission test.

1.1. Definition of test anxiety

Trait test anxiety can be defined as a disposition to experience situation-specific anxiety in evaluative situations (Putwain, 2008; Zeidner, 1998). Factor analytic research (e.g. Benson & Bandalos, 1992; Hodapp & Benson, 1997; Keith, Hodapp, Schermelleh-Engel, & Moosbrugger, 2003; Lowe, 2015; Lowe, Ang, & Loke, 2011; Mowbray, Jacobs, & Boyle, 2015; Sarason, 1984; Wacker, Jaunzeme, & Jaksztat, 2008) indicated that trait test anxiety consists of *cognitive components* and *affective components*. The affective component comprises physiological reactions (*bodily symptoms*) and the feeling of being nervous and tense (*emotionality*) while the cognitive component can be further subdivided into *worry* and *task-irrelevant thinking*. The cognitive component *worry* refers to concerns about the outcome and consequences of the assessment and has been shown to be linked to individual differences in task-orientation and test preparation (cf. Stöber, 2004). *Task-irrelevant thinking*,

* Corresponding author at: Psychological Methods and Assessment Group, Department of Psychology, University of Graz, Universitätsplatz 2, 8010 Graz, Austria.
E-mail address: martin.arendasy@uni-graz.at (M.E. Arendasy).

http://dx.doi.org/10.1016/j.lindif.2016.06.030
1041-6080/

on the other hand, denotes interfering thoughts unrelated to the content and the outcome of the assessment and has been shown to be related to avoidance coping (cf. Stöber, 2004; Schutz, Di Stefano, Benson, & Davis, 2004).

1.2. Relation between test anxiety components and scholastic achievement tests

The cognitive and affective components of trait test anxiety have also been shown to be differentially related to scholastic achievement and performance in standardized admission tests (e.g. Cassady, 2004a; Cassady & Johnson, 2002; Chapell et al., 2005; Hembree, 1988; Powers, 1986; Putwain & Daly, 2013; Seipp, 1991). In general, the cognitive components of trait test anxiety were more strongly correlated with scholastic achievement and admission test performance than the affective components. Furthermore, controlling for individual differences in the cognitive components of trait test anxiety decreased the correlation between the affective components and admission test performance while controlling for individual differences in the affective component left the correlation between the cognitive components of trait test anxiety and test performance essentially unchanged (cf. Cassady, 2004a; Hembree, 1988; Powers, 1986). Thus, the test anxiety–test performance relationship seems to be mainly driven by the cognitive components.

1.3. Moderator variables of the test anxiety–test performance relationship

Research indicated that the size of the correlation coefficient between trait test anxiety and performance in scholastic achievement tests is moderated by (1) characteristics of the scholastic achievement test, (2) characteristics of the test-takers, and (3) situational characteristics of the assessment setting.

1.3.1. Effect of test characteristics

The difficulty of the scholastic achievement test has been shown to affect the size of the correlation coefficient between trait test anxiety and scholastic achievement test performance (cf. Hembree, 1988; Hong, 1999). More precisely, test anxiety was more closely linked to performance on scholastic achievement tests of intermediate to higher difficulty (meta-analytic mean $r = -0.45$) than for tests of lower difficulty (meta-analytic mean $r = -0.07$). A possible explanation for this finding may be that more difficult scholastic achievement test items are more prone to be affected by trait test anxiety because these items require more attentional control during the planned retrieval of the relevant knowledge from long-term memory (Unsworth, Brewer, & Spiller, 2013). Alternatively, a more difficult test may simply be more anxiety provoking than less difficult tests.

1.3.2. Effect of characteristics of the test-takers

Chapell et al. (2005) reported a lower correlation between trait worry and scholastic achievement in graduate student samples than in samples of undergraduate students. One possible explanation for this finding could be that graduate students constitute a more restricted sample that is presumably more cognitively able than undergraduate students. This interpretation would be in line with studies indicating that test anxiety is more strongly correlated with test performance in samples of less able test-takers than in samples of more able test-takers (cf. Goetz, Preckel, Pekrun, & Hall, 2007; Johnson & Gronlund, 2009).

1.3.3. Effect of the assessment setting

Research indicated that test-takers experience more test anxiety in high-stakes than in low-stakes settings (cf. Bonaccio & Reeve, 2010; Nie, Lau, & Liau, 2011; Powers, 1986; Reeve, Bonaccio, & Charles, 2008; Selkirk, Bouchey, & Eccles, 2011). If test anxiety is causally related to test performance, the increased level of test anxiety experienced in high-stakes assessment situations may increase the size of the correlation coefficient between test anxiety and test performance (Powers, 1986).

1.3.4. Effect of the time-point at which trait test anxiety is measured

Trait test anxiety refers to the proneness to experience anxiety in different kinds of assessment situations and has been hypothesized to constitute a stable individual difference construct (cf. Spielberger & Vagg, 1995; Zeidner, 1998). In line with this hypothesis research indicated that measures of trait test anxiety were stable across time-points of measurement and comprised little variance attributable to the situation-specific factors (Hong, 1998; Keith et al., 2003). By contrast, state test anxiety has been defined as a fluctuating emotional state experienced in a particular assessment situation (cf. Spielberger & Vagg, 1995; Zeidner, 1998). Research indicated that trait test anxiety measured upon completing the entire admission test yielded higher test anxiety–test performance correlation coefficients than measuring trait test anxiety prior to completing the test items (cf. Seipp, 1991; Sommer & Arendasy, 2014; Zeidner, 1991). A possible explanation for this finding is that the level of state test anxiety experienced during test-taking may prime emotion-congruent memories which affect test-takers' answers to the post-test trait test anxiety questionnaire (c.f. Zeidner, 1998). Thus, trait test anxiety measured upon completing an admission test may reflect individual differences in trait- and state test anxiety.

2. Explaining the test anxiety–test performance relationship

Theoretical models, on the test anxiety–test performance relationship can be classified into two categories: models based on the *deficit hypothesis*, and models based on the *interference hypothesis* (cf. Hembree, 1988; Reeve & Bonaccio, 2008; Sommer & Arendasy, 2014; Wicherts & Scholten, 2010; Zeidner, 1998).

2.1. Deficit hypothesis

The *deficit hypothesis* posits that test performance and trait test anxiety are correlated, but test anxiety has no causal effect on test performance. Proponents of the deficit hypothesis postulate that the low performance of test anxious test-takers' results from deficits in their domain knowledge and their increasing awareness of these deficits during test-taking rather than from the interfering effect of test anxiety during the retrieval phase (cf. Birenbaum & Pinku, 1997; Cassady, 2004a, 2004b; Covington & Omelich, 1987; Klinger, 1984; Musch & Broder, 1999; Naveh-Benjamin, McKeachie, Lin, & Holinger, 1981; Naveh-Benjamin, McKeachie, & Lin, 1987; Paulman & Kennelly, 1984; Smith, Snyder, & Handelsman, 1982; Tobias, 1985; Zeidner, 1991, 1998). This variant of the deficit hypothesis has been supported by findings, which indicated that more test anxious test-takers use less effective study strategies and therefore develop a less integrated knowledge base (e.g. Birenbaum & Pinku, 1997; Cassady, 2004a; Naveh-Benjamin et al., 1981, 1987). Further support for this variant of the deficit hypothesis came from studies indicating that less test anxious test-takers also outperform their more test anxious counterparts in test-taking situations with low evaluative pressure (e.g. Birenbaum & Pinku, 1997; Cassady, 2004b; Covington & Omelich, 1987; Naveh-Benjamin et al., 1987). In its strongest wording the deficit hypothesis attributes the lower performance of test anxious test-takers entirely to deficits in constructing a well-integrated knowledge base of the subject matter domain during test preparation. Another variant of the deficit hypothesis postulates that more test anxious test-takers have deficits in attentional control processes, which are involved in the storage and planned retrieval of subject matter knowledge (e.g. Bishop, 2009). Both variants of the deficit hypothesis predict that individual differences in trait test anxiety do not induce measurement bias because test anxiety has no causal effect on test performance (cf. Halpin et al., 2014; Reeve & Bonaccio, 2008; Sommer & Arendasy, 2014; Wicherts & Scholten, 2010). The deficit hypothesis can also explain the findings on