Thinking Skills and Creativity 20 (2016) 1–16

Contents lists available at ScienceDirect

Thinking Skills and Creativity

journal homepage: http://www.elsevier.com/locate/tsc

# Supporting divergent and convergent production of test items for teachers in higher education

Silvester Draaijer[a,*], Judith Schoonenboom[b], Jos Beishuizen[a], Lambert Schuwirth[c]

[a] VU University Amsterdam, Faculty of Behavioural and Movement Sciences, Department of Research and Theory in Education, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands
[b] Department of Education, Faculty of Philosophy and Education, University of Vienna, Sensengasse 3a, 1090 Vienna, Austria
[c] Flinders University, School of Medicine, Sturt Road, Bedford Park, 5042, GPO Box 2100, Adelaide, SA 5001, Australia

ARTICLE INFO

*Article history:*
Received 23 January 2015
Received in revised form
20 November 2015
Accepted 23 December 2015
Available online 29 January 2016

*Keywords:*
Test item writing
Test development
Assessment
Diverging
Converging
Creativity
Concept mapping

ABSTRACT

Developing tests items for end-of-course tests in higher education is a recurring task for many teachers. Divergent production of test items is needed to create queries with original content, while convergent production is needed to design items that are clear and technically sound. Support for the test item design process, based on divergent and convergent production, could lead to better test items. An experiment to establish the extent to which the explicit application of diverging and converging support resulted in more and in more original test items and technically better test items was conducted in this study. The results showed an effect on the originality of the test items developed, but not on the technical quality of those items. The implications of these findings for test item design are discussed.



## 1. Background

Throughout education, test items such as such as true-false and multiple-choice questions form the building blocks of many achievement tests and play a major role in determining their quality. Especially teachers in higher education may have to design new test items for each of these exam that they create. These test items are referred to as in-house developed test items (Jozefowicz et al., 2002). Teachers need be able to produce test items for rote learning purposes as well as for higher-order learning outcomes; the latter require the teacher to produce novel conceptions of information that require the students to apply their knowledge of facts, concepts, and principles in familiar but new contexts. It requires knowledge and skill of both the subject matter and test item design to conceive such items (Downing & Haladyna, 2006; Mayenga, 2009). Authors in test item writing literature additionally suggest that it requires creativity to design those items (Osterlind, 1998; Popham, 1984; Rikers, 1988; Vale, 1995). Though in some studies and handbooks supporting creativity in designing such test items by means of examples, taxonomies or guidelines is described (Authors; Haladyna & Shindoll, 1989; Roid & Haladyna,

* Corresponding author at: Department of Research and Theory in Education, Faculty of Psychology and Education, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands.
*E-mail addresses:* s.draaijer@vu.nl, silvesterdraaijer@gmail.com (S. Draaijer), judith.schoonenboom@univie.ac.at (J. Schoonenboom), j.j.beishuizen@vu.nl (J. Beishuizen), lambert.schuwirth@flinders.edu.au (L. Schuwirth).

http://dx.doi.org/10.1016/j.tsc.2015.12.003
1871-1871/

1982; Scalise & Gifford, 2006), the authors of this manuscript are unaware of research in which the generation of test items with explicit respect to creativity or creativity stimulating interventions, was studied experimentally. The current study aims to provide an experimental empirical investigation into creativity stimulation for test item design. Specifically, the production of test items by university teachers was addressed in this study from the basic cognitive processes of divergent and convergent production (Acar & Runco, 2012).

The lack of research into the phenomenon of creativity in test item design possibly stems from the general held idea in the field of educational measurement that the main problem of test item design lies in ensuring the design of reliable tests and preventing any error in measurement (Brennan, 2006). In these fields, emphasis is put on rational methods and rational considerations in designing test items. The design process for test items is in educational measurement implicitly presented as solving a well-defined problem with a single best solution for a measurement problem. In terms of divergent and convergent production, convergent production is put central to ensure that test items discriminate as accurately as possible between the competent and less-competent student. For that reason, test items must be as clear as possible. In the design process it is emphasized that the teacher needs to prevent the design of flawed test items (Haladyna, Downing, & Rodriguez, 2002). Such flaws are for example the use of negations, spelling and grammar errors, and the use of vague wording. Flawed test items also include items that are too verbose, items with multiple or no correct answers, items with cues leading inadvertently to the correct answer or items with excessively implausible distractors. The absence of flaws is called the technical quality of test items in this study.

Test item design however entails more than an algorithmic or solely convergent problem solving processes. Test item design can also be regarded as a task to solve ill-defined creative design problems. In line with theories and findings from creative, ill-defined problem solving and design domains (Cross, 2004; Jonassen, 2004; Reitman, 1964; Rowland, 1993; Tripp & Bichelmeyer, 1990), divergent production must also be considered to be vital for designing test items, especially for test items extending beyond verbatim recall. Divergent production is needed to generate test items containing for example rephrased, reconceptualized, or associated and new facts, concepts, principles, procedures, problems, examples, counterexamples, misconceptions and plausible distractors on the basis of subject matter and instructional materials (Anderson, 1987; Haladyna, 1997, 2004). Such ideas cannot be deductively derived from the instructional materials. The cognitive process to generate such information includes exploring possible solutions for test items and relations between the possible solutions and the assessment problem to solve; what is it that the test item should elicit in answering a test item by a student?

In line with classic studies of creativity, indicators for the degree of divergent production can be expressed as originality (uniqueness), fluency (number of ideas), flexibility (number of categories), or elaboration (amount of detail) in the ideas that an individual typically produces (Guilford, 1967). The degree of divergent production by an individual in test item design can be linked to these indicators of creativity of test items and this is called the originality of test items in this study.

The question of both practical and theoretical value with regards to test item design is which specific practical applicable techniques lead to improved outcomes in the test item design process; this paper seeks to answer that question through an experimental approach in which a specific set of techniques is selected and used in a specific controlled test item design setting. In particular, it is studied if the selected techniques increase the yield of both diverging and converging production in designing test items.

In the experiment, participants were asked to design test items based on a piece of instructional material regarding the concept of nature versus nurture during a set time period in an actual test item design situation. The intervention was such that (1) the time period was split in two phases in which the participants were instructed to diverge in the first phase of that period and converge in the second phase of that period and (2) offering participants an informed set of techniques for diverging and converging. We begin by presenting the rationale for the selected combinations of techniques to support participants in the study's design task, after which the details of the experimental setup and results are presented and discussed.

## 2. Selected techniques for interventions in the test item design process

Three specific practical techniques were selected for support of interventions to improve the divergent production of test items. First, participants were provided with information about divergent production of test items and the importance of divergent thinking and production. Providing this information was intended to serve as priming for the task, which other research has shown results in an improved outcome with respect to creativity (Baumeister, Schmeichel, DeWall, & Vohs, 2007; Hacker, Dunlosky, & Graesser, 2009). Second, participants were instructed to develop a concept map (Novak, 1998) before undertaking the test item design task. Concept maps are easy to develop graphical representations of facts, concept, principles or procedures and their relations. Participants were also explicitly instructed to generate examples, counterexamples, misconceptions, unrelated but plausible facts, problems, and concepts based on recommendations by Roid and Haladyna (1982), all aimed at stimulating them to query understanding or application of knowledge and skill. Third, item shells (Haladyna, 1987) were provided to the participants. Item shells are hollow syntactic structures for test items to which test item designers can add the substance of the specific topic(s) for which test items are to be designed. Though not experimentally studied, Haladyna has described that using items shells improves the production of test item ideas aimed at designing test items for critical thinking and problem solving, especially for novice designers.

With respect to supporting the convergent phase, the participants were provided with information about the necessity of constructing unambiguous test items and supplied with Haladyna et al.'s (2002) guidelines. Examples of such guidelines