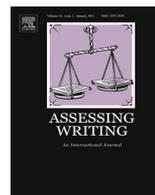




ELSEVIER

Contents lists available at [ScienceDirect](#)

Assessing Writing



Pair assessment of pupil writing: A dialogic approach for studying the development of rater competence



Lennart Jølle ^{a,b,*}

^a Department of Teacher and Interpreter Education, Sør-Trøndelag University College, Trondheim, Norway

^b Faculty of Humanities, Norwegian University of Science and Technology, Trondheim, Norway

ARTICLE INFO

Article history:

Received 25 June 2013

Received in revised form 12 January 2014

Accepted 18 January 2014

Available online 11 February 2014

Keywords:

Writing assessment

Pair assessment

Rater behaviour

Expertise

ABSTRACT

This paper reports on rating during the development of a Norwegian sample-based national assessment of L1 writing as a key competency. This assessment is to be officially introduced in August 2014. Novice members of a national rater panel to assess Year 8 pupils' texts were studied during three of their successive training sessions: in June 2011, November 2011 and April 2012. My purpose was to conduct an exploratory investigation into how the rating practice of novice raters might develop during such a preparatory stage. The raters in this study mainly assessed in pairs, and data sources were assessment dialogues. The analysis of transcripts showed that rater behaviour changed only to a minor extent towards an increased use of shared assessment resources. The quality of the assessment dialogues did not change much either, leaving the impression that raters often reached consensus without much discussion. Since reliable scoring is a collective task, however, it is argued that a more balanced use of resources, both those attained from teacher practice and those attained from being a member of a national panel, may together with more exploration-oriented dialogue be necessary to achieve sufficient reliability.

© 2014 Elsevier Ltd. All rights reserved.

* Correspondence to: Department of Teacher and Interpreter Education, HIST Trondheim, 7004 Trondheim, Norway.
Tel.: +47 73559077.

E-mail address: lennart.jolle@hist.no

1. Introduction

1.1. Background

In Norway, writing development and expectations of writing proficiency for compulsory school pupils are of current interest among policymakers, researchers and teachers. In 2006, Norway introduced a national curriculum where different key competencies were integrated in and adapted to each subject, and *writing* is one of these competencies (Knowledge promotion, 2007). Policymakers also wanted information about the pupils' writing performance, and in 2005 a first attempt to test pupils' proficiency in Norwegian (L1) writing as a key competency was made through a national writing test (Thygesen, Berge, Evensen, & Fasting, 2007). However, due to low interrater reliability the results were of such dubious value that they could hardly be reported back to the pupils or teachers, and further tests were postponed. Several explanations for the low reliability have been posited, such as the *level of expertise* among the raters and the *lack of a shared assessment culture* (Fasting, Thygesen, Berge, Evensen, & Vagle, 2009). The raters, who were all experienced teachers, had to deal with an unfamiliar theoretical construct for writing¹ when making their assessments, and this made the task doubly difficult for them. Even the most experienced teachers were turned into novice raters.² The lack of a common assessment culture might also be due to vague formulations in the various subject curricula as to what to expect from pupils at different year levels. Furthermore, there has been little research on the assessment of writing in a Norwegian context (Evensen, 2009).

Therefore, when it was decided to reintroduce a writing test from 2014, a number of critical factors needed to be addressed. By designing the test as a sample-based national assessment it became possible, both financially and practically, to establish a national assessment panel (The Norwegian Assessment Panel, NAP) to carry out the assessment. During a two-year period before the re-launching of the test approximately 90 teachers were recruited to the panel.³ NAP is now a semi-permanent panel that meets for three days on two occasions each year. The theoretical writing construct is presented to the raters on the panel, they participate in plenary and group discussions both on criteria for assessment and what to expect from pupils texts at different year levels, and they assess pupils' writings in pairs or groups according to a primary trait model. Using a five-level scale,⁴ the following six text domains are assessed: Communication, Content, Composition, Use of language, Spelling and Punctuation. Bearing the lack of shared standards in mind, the goal is not only to encourage the raters to use standards in common ways when assessing, but also to continue refining and developing these standards. A panel model, as used in this context, is expected to improve the reliability of assessments, that is, changes should take place (AERA, APA, & NCME, 1999; Breland, Bridgeman, & Fowles, 1999).

The present study should be understood with this background in mind. The purpose is to conduct a qualitative investigation into how the rating practice of novice raters within the panel develops over time.

¹ The Norwegian curriculum from 2006, where writing is introduced as a key competency, has encouraged Norwegian writing researchers to develop a theoretical writing construct that forms the basis of how to understand writing. A model, called Wheel of writing, displays the basic concepts in this construct, where acts and purposes of writing are central (Fasting et al., 2009; Wheel of writing, 2012).

² It has been important to move the conceptual understanding of writing from exclusively belonging to Norwegian language instruction (L1) to an understanding of writing as a key competency that makes it relevant in every subject for a number of purposes (DeSeCo., 2005). This turn has been difficult for teachers and raters to adapt to. What is meant by proficient writing in science has for instance not been a frequent topic of discussion among assessors of writing.

³ Since the NAP was in an early stage when the data material was compiled, the total number of raters increased from one session to the next as more raters were recruited. In June 2012, when the last tape recordings were made, the panel consisted of 90 raters.

⁴ The scale ranges from M1 (much lower than expected from most students at the actual level), through M3 (as expected), up to M5 (much better than expected from most students at the actual level).

Download English Version:

<https://daneshyari.com/en/article/344220>

Download Persian Version:

<https://daneshyari.com/article/344220>

[Daneshyari.com](https://daneshyari.com)