Empowerment of Trainee Teachers; Students and Teachers Discuss and Explore Workplace Dilemmas by Means of an International Virtual Community of Practice

Roelien Bos-Wierda, Ron Barendsen, Peter Joore, Hans G. K. Hummel, Toine van den Bogaart

NHL Hogeschool, Leeuwarden, The Netherlands

Abstract

In this research paper the implementation of a new method of peer consultation will be described. Face to face (f2f) peer consultation or “intervision” is traditionally a structural part of Dutch teacher training programmes under the assumption that “intervision” – or the discussion of workplace dilemmas during traineeship – will bridge the gap between workplace learning and formal learning and that “intervision” will lead to short-term improved professional performance. However, extensive research has demonstrated that this is not the case. What can be demonstrated, however, is that participants develop a shared professional discourse (Joyce & Showers, 2002).

Moreover, face-to-face intervision has a number of limitations (Wierda & Barendsen, 2011) which will be discussed in this paper. To meet these drawbacks an online intervision tool was developed and tested with a group of student teachers (n=46).

Hence, the overall hypothesis of our research study is that online intervision – in combination with f2f intervision - leads to a stronger development of the shared professional discourse than f2f intervision alone. The online intervision tool was developed as an integral part of the multifunctional educational platform MySchoolsNetwork (Wiera & Barendsen, 2011). We have not been able to demonstrate a development of shared professional discourse as a consequence of f2f intervision alone. When complemented with online intervision the data did show a substantial degree of growth of a shared professional discourse.

Index Terms— Workplace dilemmas, Peer Consultation, E-pedagogy, Community of Learners

I. INTRODUCTION

In the Netherlands, the development of teacher competencies by means of workplace learning lies at the foundation of teacher education. Peer consultation or intervision has always been seen as the bridge between workplace learning and academic training. The underlying thought is that intervision will bring about the transfer between academic learning and experiences in the workplace. However, research has shown (De Haan, 2001) that intervision will not lead to an immediate improvement of workplace performance. However, this does not mean that students regard the intervision sessions as useless. An exploratory research study (Wierda & Barendsen, 2011) has shown that 86% of the students questioned see f2f intervision as a valuable addition to their internship. They indicate that they find comfort in the fact that they can share experiences – and sometimes frustrations - with their peers.

A large-scale research study by Joyce and Showers (2002) has shown that peer consultation leads to the development of a shared professional discourse, which in its turn is one of the conditions for a continuing competency development, once the student has graduated. Reason enough to explore whether the practice of intervision cannot be brought to a higher, more productive level by facilitating deep learning, exchange of workplace experiences with peers and experts and the practical exploration of the conceptual framework.

In order to do so we first examined the f2f sessions to determine possible shortcomings and drawbacks. The f2f sessions typically last 1-3 hours and are attended by 8-10 students and a staff member leading the discussion. A strict protocol (incident method) is used: each student describes a recent workplace dilemma to the group, two or three of which are chosen. In order to clarify the dilemma the members of the group ask questions to obtain missing information. The next step is that group members formulate an analysis of the situation. Finally advice is formulated and the owner of the dilemma gives a reaction. By following this procedure participants learn not to jump to conclusions but to base their analysis and advice on the basis of a thorough understanding and analysis of the dilemma. The most prominent limitations of the f2f method are:

- Not enough time to discuss more than 2-3 dilemmas; rest of the dilemmas remain undiscussed and unresolved
- Not enough time for deep reflection
- No time and opportunity to consult relevant literature
- Dilemmas are not recorded and soon forgotten; no knowledge construction
- Limited expert input

It was our assumption that the drawbacks mentioned above could be countered by the added use of an online intervision tool. In the online programme students first upload the dilemmas they encounter during their
In this research study 2 sub-questions have been formulated:
1. What is a Shared Professional Discourse (SPD)?
2. How can SPD (and its development) be measured?

**Sub-question 1:** What is a Shared Professional Discourse?
Donald Freeman describes professional teacher discourse as follows: “Like any community [a teacher training course] shares … a common discourse. The discourse creates a set of facts [and] perceptions which are accepted as givens by members of the discourse community. …These … become the socially constructed facts on which the shared professional discourse … is based. The new professional discourse serves two purposes: it allows teachers to make themselves part of the discourse community (social function), and enables them to identify aspects of their experience and thus to organize and develop their conceptions of teaching (cognitive function).” (Freeman 1992)

For this study Freeman’s definition has been interpreted as follows:
- “Shared perceptions” become manifest in the professional discussions of workplace dilemmas, more specifically in analysis and advice.
- “Shared facts” become manifest in the professional’s conceptual framework

**III. EMPIRICAL RESEARCH METHOD**

This research pilot depended on the department’s organization of groups and timetables; therefore we have opted for a quasi-experimental research design. Data have been collected from an expert group, a test group and a control group.

The expert group consists of teachers with a minimum of three years of professional experience. The students in the test group and the control group are all second year language students of the teacher training faculty of the NHL University of Applied Science in the Netherlands. Four groups have been formed on the basis of the department’s organizations of groups: two test groups consisting of second year students of English and two control groups consisting of second year students of Dutch, German, French and English. Each of the four groups had their own tutor.

In the experimental test group the weekly f2f intervention sessions were complemented by the online tool whereas the control group only had the f2f sessions. Both groups started with a session explaining the procedure of the incident method, after which both groups had the same programme for three weeks. After that the test group started the online programme to complement the f2f sessions.

See figure 2 for the program.
A. Instruments

For this research study students had to complete a pre-test and a post-test. Central in both tests was an authentic classroom dilemma (uploaded by a student teacher via the online tool some time before the research started). A questionnaire was formulated around this dilemma in which the research population was invited to look at the dilemma from different angles, analyse it and formulate an advice. In order to analyse the elicited data, the Content Analysis Method was used to elicit keywords/dimensions that will be described later.

B. Participants

The research population was formed by all second year language students of the NHL teacher training faculty (n=46) (English, French, German and Dutch) in the academic year 2010-2011. The pre-test was completed by 100% of the research population whereas 65% completed the post test. There are two reasons for the lower response during the post-test:

1. some students had dropped out in the course of the year,
2. the post-test was submitted to the students in the final week of their internship; some students needed extra time to fulfill their obligations at the professional development school and therefore did not show up in class.

C. Pre-test and post-test analysed by means of the Content Analysis Method

This research study focuses on the presence c.q. the development of an SPD, or – in other words – the ability to communicate with peers about the professional context in a critical, well-founded and constructive way. The question arose how this could be measured on the basis of the test results, which consisted of textual material. It is important to realize that this study does not focus on “student performance”, “student participation” or the production of the correct answer. “Critical thinking is not just limited to the one-off assessment of a statement for its correctness, but a dynamic activity, in which critical perspectives on a problem develop through both individual analysis and social interaction” (Garrison, 1992).

In order to analyse the texts by means of the content analysis method a set of keywords/dimensions had to be developed as a basis to analyse and score the test results. For this purpose the Interaction Analysis Model (Gunawardena, 1997) was used, as this model seemed to fit the process that we were going through perfectly. Moreover, the inherent interrater reliability of this model would give added validity to this research study. The model was adapted to the specific needs and context of this study in order to develop a useful set of dimensions.

The Gunawardena model is represented by the following procedure: “The Gunawardena et al (1997) Interaction Analysis model used to analyze the phases of learning outlines five phases of knowledge co-construction that occur during (online) debate. These include: Sharing/ Comparing, Exploration of Dissonance, Negotiation/Co-construction, Testing Tentative Constructions, and Statement/Applications of Newly-Constructed Knowledge”

The model was operationalized for this research study as follows:

- The data from the experts’ pre-test were analyzed and coded by the researchers (independent from each other, for the sake of interrater reliability) Each of the researchers made a list of the keywords with the highest frequency.
- Subsequently the results were shared and compared.
- Next, the differences were exposed and identified.
- Then positions and views were clarified and negotiations about the different interpretations of the essence of professional discourse in this particular case took place.
- Finally, a list of keywords with the highest frequency was drawn up, after which these keywords were clustered into five groups. Thus, 5 dimensions were defined to cover all aspects of the professional discourse in the advisory reports of the experts. See figure 3 for the 5 dimensions.

D. How was the SPD and its development measured?

In order to measure the development or growth of the SPD a pre-test and post-test were developed. For the scoring of these tests the Interaction Analysis Model was used again. This means that researchers scored dimensions and keywords mentioned in figure 3 independent from each other, followed by discussions and negotiations about the different interpretations, resulting in interrater valid scores.

The pre-test was submitted to a group of experts. The average number of relevant dimensions, as mentioned by the experts in their advice, was calculated. This average would later serve as a benchmark to measure the students’ advisory texts. The assumption here is that expert teachers’ discourse – by definition – represents the Shared Professional Discourse.

The test group and control group of students did the same test, which was analysed by scoring the relevant dimensions in the advisory texts. The difference between the average of the experts and the test group and control group was then established.

After the intervention, the post-test was submitted to all groups. The post-test procedure was similar to the pre-test, the only difference being that a different workplace dilemma was submitted for analysis and advice. The difference between the benchmark score of the experts and the test group and control group scores was
established to find out whether the hypothesis of this research study could be confirmed.

IV. RESULTS

The pre-test shows that the experts’ discourse was – on average - more fully-developed than the student teachers. Figure 4 shows that the experts’ scores are higher on all dimensions except visible behaviour (i.e. advice formulated in terms of visible behaviour). The control group scored 1.46 and the test group had 1.80. So, the difference between experts and test group was 1.09 and between experts and control group 1.43, the difference between test group and control group being 0.37.

In the post test (figure 5) the experts’ average was considerably higher than the students’ scores; the experts again scored higher on almost all dimensions, except visible behaviour. The experts had an average score of 3.3, the test groups scored 2.23 and the control group had 1.29. Thus, the difference between experts and test group is 1.1, the difference between experts and control groups is 2.04. The difference between test group and control group is 0.94.

The Wilcoxon-Mann-Whitney-test

Pre-test and post-test were taken on an anonymous basis. This means that there are no individual data available. In order to prove or disprove the hypothesis that the intervention of f2f intervision, complemented with online intervision, leads to an increased development of the SPD. The Mann-Whitney-Wilcoxon-test was used (Wilcoxon, 1945; Mann & Whitney, 1947). This is a statistical test of the difference between the distributions of data collected in two experimental conditions applied to unmatched groups of subjects but comparing the distributions of the ranks of the scores. The data are arranged from low to high, all values having an ordinal number. Then the averages of the ordinal numbers of the control group and the test group are calculated. The next step is that these averages are compared. If it turns out that the larger part of all ordinal numbers of the one group are higher than those of the other group, then the conclusion can be drawn that there is a significant difference (Hollander et al., 1973).

With this test a 69.7% validity has been established for the hypothesis that the addition of the online tool to f2f intervision leads to an increased development of the SPD.

V. DISCUSSION AND RECOMMENDATIONS

On the basis of the results of this research study there is reason to assume that online intervision has added value in the context of workplace learning. However, further research is necessary. The following aspects and improvements should be taken into consideration:

- The average score for the control group decreased from 1.46 to 1.29. The distance to the experts’ average score has grown significantly.

These data provide information about relative and absolute scores on the pre- and post-test separately, however the hypothesis of this research study can neither be supported nor rejected.

The Wilcoxon-Mann-Whitney-test

Pre-test and post-test were taken on an anonymous basis. This means that there are no individual data available. In order to prove or disprove the hypothesis that the intervention of f2f intervision, complemented with online intervision, leads to an increased development of the SPD. The Mann-Whitney-Wilcoxon-test was used (Wilcoxon, 1945; Mann & Whitney, 1947). This is a statistical test of the difference between the distributions of data collected in two experimental conditions applied to unmatched groups of subjects but comparing the distributions of the ranks of the scores. The data are arranged from low to high, all values having an ordinal number. Then the averages of the ordinal numbers of the control group and the test group are calculated. The next step is that these averages are compared. If it turns out that the larger part of all ordinal numbers of the one group are higher than those of the other group, then the conclusion can be drawn that there is a significant difference (Hollander et al., 1973).

With this test a 69.7% validity has been established for the hypothesis that the addition of the online tool to f2f intervision leads to an increased development of the SPD.

On the basis of these data the following conclusions can be drawn:

- The experts scored considerably higher in the post-test than in the pre-test.

- The average score for the test group increased from 1.8 to 2.23, however the comparison to the experts’ scores shows no significant increase in relative terms.
The original assumption of this research study was that SPD can be measured from three different angles; in this paper the focus was on data collected from written advisory texts. In a follow-up study attention will be paid to the development of the conceptual framework and on the development of professional competency-awareness.

ACKNOWLEDGMENT
With special thanks to Dr. Barbara Schwarz-Bechet for giving advice and feedback.

REFERENCES

AUTHORS
First A. Roelien Bos-Wierda is with the NHL University of Applied Sciences, Leeuwarden, 8900 CA Netherlands (e-mail: r.wierda@icw.nhl.nl).
Second B. Ron Barendsen is with the NHL University of Applied Sciences, Leeuwarden, 8900 CA Netherlands (e-mail: r.m.barendsen@iec.nhl.nl).

Manuscript received 30 March 2012.
Published as submitted by the author(s).