12th International Congress on Mathematical Education
Program Name XX-YY-zz (pp. abcde-fghij)
8 July – 15 July, 2012, COEX, Seoul, Korea (This part is for LOC use only. Please do not change this part.)

LEARNING TO INTEGRATE STATISTICAL AND WORK-RELATED REASONING

Arthur Bakker, Xaviera van Mierlo, & Sanne Akkerman

Utrecht University, Freudenthal Institute for Science and Mathematics Education

a.bakker4@uu.nl

People typically find it hard to use mathematics and statistics they have learned at school in work contexts. We used a boundary-crossing approach to help apprentices in secondary vocational laboratory education integrate the statistics learned at school with work-related knowledge. In five one-hour meetings three apprentices refreshed their statistical knowledge (e.g., correlation, regression, variation coefficient) in relation to the work task of comparing two measurement machines (including linearity, stability, reproducibility). Data collection involved video and audio recordings of the meetings and students' data from their research projects as well as interviews with two teachers and two workplace supervisors. An analysis of the apprentices' reasoning during the meetings revealed that the level of integrating mathematical-statistical and work-related reasons increased significantly and with medium effect sizes both at group and individual level.

Apprenticeship, boundary crossing, context-based statistics education, knowledge integration

PURPOSE AND PROBLEM

The purpose of this paper is to provide insight into how vocational students can be helped to integrate statistical knowledge they have mainly learned at school with workplace-related knowledge they have mainly developed during their apprenticeships. Discrepancies and transfer problems between school and out-of-school mathematics have been found for many years in several areas (cf. Lave, 1988; Nunes, Schliemann, & Carraher, 1993). It is therefore not surprising that students find it hard to use what they have learned at school in workplaces. It is also known that statistics as typically taught at school is quite different from what is typically used in workplaces (e.g., Bakker, Kent, Derry, Noss, & Hoyles, 2008). More generally, many researchers theorise such differences between school and work or daily life in terms of boundaries (for a review study see Akkerman & Bakker, 2011).

THEORETICAL BACKGROUND: TRANSFER AND BOUNDARY CROSSING

As many scholars have argued, adopting a transfer perspective on the problem has its limitations (e.g., Lave, 1988). Transfer is mostly considered to be the application of some general principle by a person in a new situation when confronted with a task. The concept is thus unidirectional and oriented on individuals performing tasks. In sociocultural traditions (e.g., Tuomi-Gröhn & Engeström, 2003), the alternative metaphor of boundary crossing has been proposed to capture the often more complex situation that people move not only forth but also back. Boundary crossing is therefore bidirectional, dynamic, and oriented on both individual and collective. We do not wish to imply that transfer does not exist or is not worth pursuing; we only suggest that the concept of boundary crossing draws attention to a wider range of relevant processes involved in learning to integrate theoretical knowledge such as statistics in work-based knowledge.

If not just transfer but boundary crossing is to be promoted in vocational education, what would it look like? In mathematics and statistics education, several researchers have explored the potential of boundary-crossing ideas in vocational and workplace settings. Wake and Williams (2007) invited mathematics college students to workplaces and discussed what they had seen. In collaboration with companies, Hoyles, Noss, Kent, and Bakker (2010) developed an approach to designing mathematical learning opportunities along with so-called technology-enhanced boundary objects – reconfigurations of workplace artefacts that involved some mathematics or statistics. Boundary objects are objects that are functional in different communities and fulfil a need, but typically not the same for each community (Star, 2010). For statistical examples see Bakker, Kent, Noss and Hoyles (2009).

In this paper we apply the insights gained in workplace training to vocational education, which is a form of education in between general education and workplace training. We focused on Dutch senior secondary vocational education (MBO), when students are mostly between 16 and 22 years old. The first year of MBO is typically school-based, but there is a gradual shift to work placement in the final year (apprenticeship). The day release programme of the final years, when apprentices come back to school one day per two weeks, is a particularly interesting boundary between general education and work. We were particularly interested in helping students cross possible boundaries between school and work situations. This paper studies how we can promote and study knowledge integration processes.

Boundary crossing is mostly conceptualised as the movements of people or objects between communities of practice (Wenger, 1998) or activity systems (Engeström, 2001). In our view, however, the sociocultural focus on practice and activity does not necessarily suffice to understand how apprentices can integrate the different types of knowledge that they learn at school (e.g., standard deviation, variation coefficient, correlation and regression) and develop during their apprenticeships (e.g., conditions that measurement machines should fulfil). To complement the one-sided focus in sociocultural theory on practice or activity and to avoid simplistic views on context, Bakker and colleagues introduced Brandom’s (1994) concept of *web of reasons* in educational theory (Bakker & Derry, 2011; Bakker et al., 2008; Kent, Bakker, Hoyles, & Noss, 2011).

This philosophical concept refers to the complex of interconnected reasons, premises and implications, causes and effects, motives for action, utility of tools for particular purposes that are at stake in particular situations. In workplace settings, some of these reasons are practical, some are statistical, and they are often weighed for their relative merits. Bakker et al. (2008) give an example from the car industry in which practical reasons outweighed statistical reasons for doing something else because the situation did not lead to customer complaints. Their analysis suggests that vocational students, apprentices and employees need to learn to integrate reasons of different nature (practical, statistical, mathematical etc.) in a web of reasons. Conceptualising this process as transferring statistical knowledge elements to a work context would be a one-sided view. Compared to statistical webs of reasons as learned at school, webs of reasons at work are different constellations of interconnected reasons.

BOUNDARY-CROSSING APPROACH AND QUESTION

We therefore did not focus on the transfer of statistical knowledge to work tasks, but designed an approach in which apprentices learned to include different types of reasons in their reasoning about work tasks. We drew on the boundary-crossing literature (Akkerman & Bakker, 2011) as a framework for action (diSessa & Cobb, 2004) to design what we came to call a *boundary-crossing approach*.

This approach in MBO laboratory education (clinical chemistry) entailed the following.

a. Apprentices were stimulated to formulate questions at work and ask them at school, and vice versa.

b. Workplace supervisors were invited at school to answer apprentices' questions and tell about how statistics was used in hospital laboratories (further abbreviated to 'labs').

c. In the meetings we used software that was also commonly used in the labs (Excel).

d. A boundary object was used as the connecting thread through the meetings. In this case it was a report of an apprentice's project of the previous school year. It was about comparing a new machine for measuring a concentration of a chemical substance with the old and reliable machine. We considered the report a boundary object because it served different functions in different communities. Initially it was the end product of an apprentice's project in a hospital lab (on method comparison) which contained results that were useful to the lab (whether the new machine was reliable and stable enough) and it was graded at school, so the apprentice could get his diploma. We used it to give the next generation of apprentices an idea of what kind of project they may be doing in their labs, to discuss the statistics needed for such projects, to stimulate apprentices to ask questions, and for teachers and supervisors to talk about their expectations of apprentices.

In this paper we evaluate one important aspect of our approach. We ask: *How well did the apprentices learn to integrate statistical-mathematical and work-related reasons in their reasoning about work tasks?*

METHOD

After ethnographic and survey research in laboratories, we designed an intervention using the aforementioned boundary-crossing approach. The participants were three apprentices (19 years old), one male, Ferdie and two female, Sylvana and Petra (all names are pseudonyms), in their fourth and final year of the highest level (4) of MBO laboratory education. The first author was their teacher for this intervention because the regular teacher did not feel comfortable enough about the statistics involved. Two supervisors from a hospital lab were invited in the third session to answer the apprentices' questions and discuss with them what they thought was important about the statistics required for method comparison.

Method comparison is a common project for many apprentices in clinical chemistry education, which involves a lot of statistics. The core of the comparison is pair-wise comparison of patient blood in both machines, leading to paired data to be represented in a scatterplot (Figure 1). Correlation and regression are applied to measure the degree of correlation (here $R^2 = 0.9634$) and, more importantly, the slope of the regression line (here 0.7464). However, lab technicians typically think in terms of linearity (is the correlation coefficient close to 1.000?), bias (does the slope deviate from 1.0?), stability and reproducibility (measured with the variation coefficient). What is not clear for most apprentices are the connections between statistical concepts and techniques such as variation coefficient, slope, correlation and regression on the one hand and lab technical concepts such as stability, bias and linearity on the other.

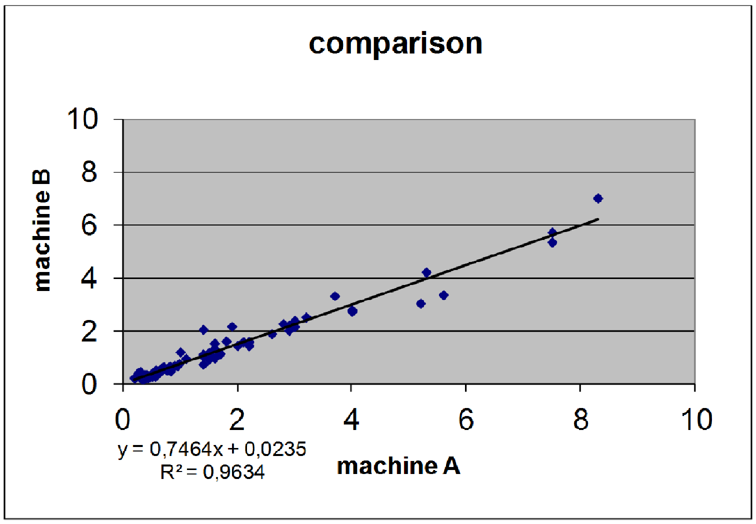

Figure 1: Comparison of machines A and B

The following data were collected: pre-interviews with two teachers were audio recorded, all classroom interaction was video and audio recorded by the second author. A brief questionnaire was handed out beforehand and discussed with the apprentices. All verbal interactions were transcribed verbatim.

Table 1: Levels of integrative reasoning

Level	Characterization
1	Statement about something statistical-mathematical *or* work-related but without explanation or reasoning
2	Reasoning with only statistical-mathematical *or* only work-related (non-statistical) knowledge.
3	Statement in which a statistical-mathematical fact *and* a work-related fact are combined.
4	Reasoning with both statistical-mathematical *and* work-related knowledge

To test if apprentices learned to better integrate both statistical and work-related knowledge in their reasoning about method comparison, we developed a coding system of what we call levels of integrative reasoning (Table 1). It is based on the following assumptions:

- Involving *both* statistical and workplace-related knowledge in a statement or reasoning is of a higher level of integrative reasoning than involving *either* statistical *or* workplace-related knowledge. Hence levels 3 and 4 are defined as higher than levels 1 and 2.
- *Reasoning* is in general of a higher integrative level than merely making a *statement*. A sign of reasoning is if students use if-then constructions or if cause-effect relationships are discussed. Hence level 2 is considered higher than level 1, and level 4 higher than level 3.

For the sake of being able to measure improvement in levels of integrative reasoning we quantified the levels at an interval scale from 1 to 4. Using Atlas.ti the transcripts were divided into fragments that covered one topic. This resulted in roughly 40 fragments per meeting, except in the third meeting when one supervisor talked quite long about particular topics (Table 2). In the first analysis round, codes were attributed to fragments of *group* interaction. Each of the fragments got one code – determined by the highest level of statement or reasoning in that fragment, irrespective of which apprentice made the statement or expressed the reasoning. Only correct reasoning was scored. Apprentices' statements or reasoning led by the teacher were not coded. In the second round of analysis we coded the fragments for each individual apprentice to test whether their *individual* integrative reasoning level increased.

RESULTS

Table 2: Numbers of codes for integrative reasoning level in each meeting

Meeting	M 1	M 2	M 3	M 4	M 5	Total
Level 1	29	25	11	21	17	103
Level 2	10	6	1	7	3	27
Level 3	3	9	7	7	20	46
Level 4	0	2	0	4	4	10
Total	42	42	19	39	44	186

As part of the first round of analysis (of the group interaction), Table 2 provides the number of codes per meeting and level, and Figure 2 shows the average level of integrative reasoning per one-hour meeting. A one-way ANOVA with linear contrast showed that the increase in average reasoning level was statistically significant, $F(1, 181) = 16.61$, $p = .000$, $\eta^2 = .08$, which is a moderate effect size according to Cohen (1988). This suggests that our goal of improving integrative reasoning level was accomplished by the boundary-crossing approach taken. In the first meetings, the apprentices mostly mentioned statistical or work-related facts, but in the later meetings, they more often included both statistical and work-related knowledge when reasoning about work-related problems.

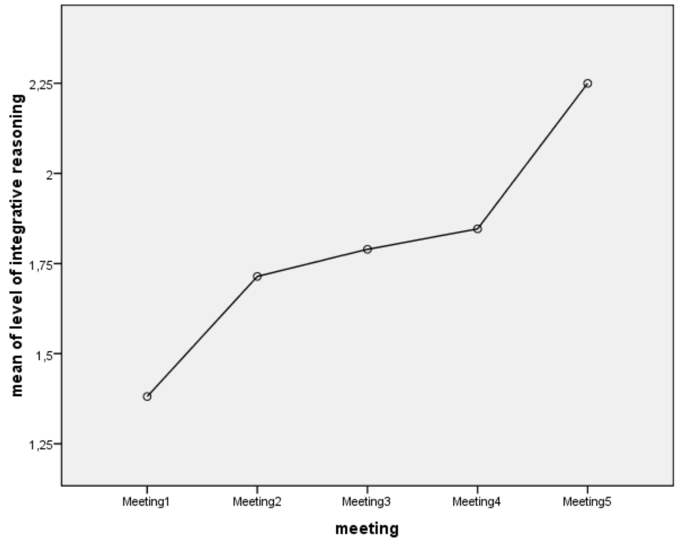

Figure 2. Mean level of integrative reasoning during group interaction per meeting

The second round of analysis of individual apprentices' reasoning indicates that their individual levels of integrative reasoning also improved. For all three, a one-way ANOVA with linear contrast led to statistically significant increase in integrative reasoning level (α = .05), all with medium effect sizes (between .05 and .15: Cohen, 1988). *Df* = 1 because we have used a linear contrast. Ferdie: $F(1, 90) = 5.34$, $p = .023$, $\eta^2 = .06$; Sylvana: $F(1, 109) = 13.91$, p = .000, $\eta^2 =.13$; Petra: $F(1, 68) = 5.19$ $p = .026$, $\eta^2 = .08$.

To give the reader a sense of the nature of improvement in integrative reasoning, we now turn to qualitative examples of each level in chronological order. When asked what is involved in method comparison, the apprentices mentioned several statistical techniques and concepts but had little idea which to use for a method comparison. Because they did not mention correlation, which is actually at the core of method comparison, the teacher then asked:

Teacher: Correlation, do you remember what it is?

Sylvana: Yes.

Teacher: Correlation coefficient?

Ferdie: Yes, that.

Sylvana: I think with that line.

Ferdie: Yeah, that's it.

Sylvana: I think it is something like this [drawing a straight line].

This fragment from meeting 1 was coded at level 1 because students only mentioned something statistical without statistical or work-related reasoning.

From meeting 2 we quote Sylvana:

Sylvana: If the results [of the new method] deviate too much [from the reliable old method] (…) then you cannot use the method, because then the patients' measurements are not all right. Only a specific deviation is allowed.

This was considered reasoning (indicated by "if … then" and "because") but using only work-related, non-statistical reasons (level 2). If she had shown understanding of the deviation earlier in the transcript in terms of a slope of the regression line, variation coefficient or a standard deviation, then we would have coded it at level 4.

In the third meeting one of the supervisors said they were satisfied with correlation coefficients of 0.9 for particular chemical substances. The supervisors and teacher stimulated the apprentices to find out what the norm at their own labs was. In the fourth meeting Petra reported:

Petra: But there are things that are different in my lab. (…) They [the supervisors] said that a correlation of 0.9 was enough, but at my lab, they say 0.099, uhm, 0.99.

Here she mentioned both work-related (norm in my lab) and statistical elements (correlation coefficient), but we coded this as a statement rather than an example of reasoning (level 3) because she just reported facts without an explanation.

To illustrate apprentices' reasoning at the fourth level we first need to discuss one of the dilemmas introduced by one supervisor in the third meeting. Assume a new machine, such as B in Figure 1, is going to be used because it is reliable (assuming $R^2 = 0.9634$ is fine), faster and cheaper than machine A, but systematically measuring lower than machine A. The slope of 0.7464 suggests that the bias is about -25.4%. What should the lab technicians decide? One option is to build in a correction factor (of 1/0.7464 = 1.34) into machine B's software so that measurement values are pretty much the same as before with machine A. Another option is to tell medical specialists that the values have gone down systematically. If a reference value of 0.5 mg/L of a particular substance (here a specific protein) in blood used to be the critical value for checking a particular disease (here thrombosis), the new reference value would become 0.7464* 0.5= 0.37 mg/L. Both options have advantages and disadvantages. In the first option, specialists do not have to get used to new reference values, but there is a risk in a reboot of the machine or installation of updated software that the correction factor is not carried over or that users are not aware of a correction factor being built in. In the second option, specialists will get confused because they have a sense of which concentrations of substances in blood are of clinical significance. If these values change because of one machine measuring differently, they will not be pleased. Moreover, comparison across labs or hospitals will become impossible.

This dilemma arose in the fifth meeting when the apprentices discussing Ferdie's data (represented with some adjustments in Figure 1), because machine B systematically measured lower than machine A. This gave rise to interaction at level 4 because mathematical elements (correction factor) were connected to work-related reasons (either for using a correction factor or changing the reference value).

Sylvana: If you assume that this one [machine A] is reliable and this one [machine B] lower, then you can/

Petra: build in a correction factor.

Ferdie: That would be possible but the point is that these are totally different measurements.

Sylvana: But then you take reference values as your starting point.

Ferdie: I think you cannot just build that [correction factor] in.

Sylvana: Otherwise you would have to adjust that one [pointing to the reference value].

Another interaction at level 4 concerned the clinical meaning of outliers, a statistical concept for which the apprentices had learned statistical tests. The reference value of thrombosis indication was 0.50 mg/L of a particular protein, below which the patient most likely has no thrombosis. When talking about whether 1.40 would be an outlier, Ferdie brought in some clinical reasoning:

Ferdie: Whether this protein is 1.40 or 2.06 has no meaning whatsoever.

Petra: No, it is just too high.

Ferdie: But if it were 0.56 or 0.38 then it would be a huge difference. Because at .38 specialists would say: I am not going to give a treatment [a value below .50 means no risk of thrombosis]. With this one [.56], they say: Here you go, anticoagulation [drug against thrombosis].

DISCUSSION

We asked how well apprentices in our boundary-crossing approach learned to integrate statistical- mathematical and work-related reasons when solving work tasks. The quantitative analyses at both group and individual level suggest that the apprentices' integrative reasoning levels improved significantly and with medium effect sizes. The qualitative illustrations give a sense of how the apprentices' reasoning became richer and how their webs of reasons were enhanced with statistical and work-related knowledge elements. In the last meetings they showed a rather sophisticated understanding of work-related dilemmas.

It is too early to conclude that boundary-crossing approaches are helpful in helping vocational students or apprentices integrate knowledge from different sources such as statistics and the laboratory in their reasoning. This teaching experiment only considered three apprentices in a setting in which all three worked on a similar work project (method comparison). More teaching experiments with larger groups in different vocational areas and with the regular teachers are desirable. Moreover, it is our experience (cf. diSessa & Cobb, 2004) that the quality of students' learning cannot be cleanly attributed to one characteristic of an approach (in this case a boundary-crossing approach). The quality of teaching, the suitability of the teaching materials, and participation by the apprentices are all influential as well.

Yet we think our approach is promising. The regular teacher was very positive about the approach. After the third meeting she exclaimed: "This is a feast. This is what it should be like." The apprentices also appreciated the approach and claimed they had learned a lot. One interesting effect of inviting the workplace supervisors was that they started negotiating with the teachers about what was possible and desirable in the curriculum. They asked about the curriculum, the books used, and mentioned some wishes for a regional meeting in which school and work could be attuned. In other words, unplanned boundary crossing between supervisors and teachers was the result of the invitation to the third meeting. This last point underlines the importance of widening the focus on transfer to boundary crossing processes at the level of broader practices.

We think that research in the vocational area could be relevant to general education as well, because context-based, project-based and other authentic approaches are explored in general education. Insight into how mathematics and statistics are used at work and how vocational students can be prepared for this practical usage should be a useful resource for general education as well.

As Lave (1988) wrote: "It seems impossible to analyze education – in schooling, craft apprenticeship, or any other form – without considering its relations with the world for which it ostensibly prepares people." This underlines the importance of studying relations between knowledge taught in education on the one hand and knowledge used in daily life or workplace settings on the other. Though she writes "it seems impossible" not to consider such relations, research at the boundary between school and work is still rare (e.g., Bakker et al., in press; Hahn, 2012; Roth, in press). The vast majority of studies in mathematics education deal with students and their teachers in general education as a relatively closed system, and a small minority of studies deal with workplace mathematics with little consideration of its relation to school mathematics.

One explanation for the lack of research on the transitions between school and work might be that vocational education is not a wide-spread educational system in many countries. Researchers are often unfamiliar with it and research in this area often requires hybrid expertise. Yet we encourage future research in this area because research in vocational education can help us understand how to bridge the gap between abstract and general mathematics and statistics typically taught at schools and situated workplace mathematics as typically found in workplaces.

Acknowledgment

We gratefully acknowledge funding by the Netherlands Organization for Scientific Research (NWO-PROO), under grant number 411-06-205 and Bert Nijdam and Rob Gould for their statistical advice.

References

Akkerman, S. F., & Bakker, A. (2011). Boundary crossing and boundary objects. *Review of Educational Research, 81,*132-169.

Bakker, A. & Derry, J. (2011). Lessons from inferentialism for statistics education. *Mathematical Thinking and Learning, 13*, 5-26.

Bakker, A., Groenveld, D. J. G., Wijers, M., Akkerman, S. F., & Gravemeijer, K. P. E. (in press). Proportional reasoning in the laboratory: An intervention study in vocational education. *Educational Studies in Mathematics.* Online first.

Bakker, A., Kent, P., Derry, J., Noss, R., & Hoyles, C. (2008). Statistical inference at work: The case of statistical process control. *Statistics Education Research Journal, 7*(2), 130-145.

Bakker, A, Kent, P, Noss, R., & Hoyles, C. (2009). Alternative representations of statistical measures in computer tools to promote communication between employees in automotive manufacturing. *Technology Innovations in Statistics Education, 3*(2). Retrieved from: http://www.escholarship.org/uc/item/53b9122r

diSessa, A., & Cobb, P. (2004). Ontological innovation and the role of theory in design experiments. *Journal of the Learning Sciences, 13*, 77-103

Cohen, J. (1988). Statistical power analysis for the behavioural sciences. New York: Academic Press.

Engeström, Y. (2001). Expansive learning at work: Toward an activity theoretical reconceptualization. *Journal of Education and Work, 14,* 133-156.

Hahn, C. (2012). Apprenticeship in higher education in France: An experimental device to help apprentices to link academic knowledge and work experience. *Journal of Vocational Education and Training, 64*(1), 75-86.

Hoyles, C., Noss, R., Kent, P., & Bakker, A. (2010). *Improving mathematics at work: The need for techno-mathematical literacies.* London: Routledge.

Kent, P., Bakker, A., Hoyles, C., & Noss, R. (2011). Measurement in the workplace: The case of process improvement in the manufacturing industry. *ZDM the International Journal on Mathematics Education, 43*(5), 747-756.

Lave, J. (1988). *Cognition in practice: Mind, mathematics and culture in everyday life.* Cambridge, MA: Cambridge University Press.

Nunes, T. A., Schliemann, D., & Carraher, D. W. (1993). *Street mathematics and school* mathematics. Cambridge, UK: Cambridge University Press.

Roth, W.-M. (in press). Rules of bending, bending the rules: The geometry of conduit bending in college and workplace. *Educational Studies in Mathematics*. Online first.

Star, S. L. (2010). This is not a boundary object: Reflections on the origin of a concept. *Science, Technology & Human Values, 35,* 601-617.

Tuomi-Gröhn, T., & Engeström, Y. (2003). Conceptualizing transfer: From standard notions to developmental perspectives. In T. Tuomi-Gröhn & Y. Engeström (Eds.), Between school and work. New perspectives on *transfer and boundary-crossing* (pp. 19-38). Amsterdam: Pergamon.

Wenger, E. (1998). *Communities on practice, learning, meaning and identity.* Cambridge: Cambridge University Press.

Williams, J. S., & Wake, G. D. (2007). Metaphors and models in translation between college and workplace mathematics. *Educational Studies in Mathematics, 64,* 345-371.