The Effects of Portfolio-based Advice on the Development of Self-directed

Learning Skills in Secondary Vocational Education

Wendy Kicken , Saskia Brand-Gruwel, Jeroen J. G. van Merriënboer, and Wim Slot

Educational Technology Expertise Centre and Netherlands Laboratory for Lifelong Learning Open University of the Netherlands

Abstract

This experimental study was designed to investigate whether supervision meetings, in which students receive specific advice on how to use a development portfolio to monitor their progress and plan their future learning, helps them to develop self-directed learning skills and improve their learning in the domain. In the first year of a hairdressing program in vocational education, supervision meetings were used to provide students with either specific advice or not. Students in the advice group (n = 21) formulated better learning needs, selected more suitable learning tasks, completed more practical assignments, and acquired more certificates than students in the feedback-only group (n = 22). Interviews also showed that students in the advice group appreciated the supervision meeting more and perceived them as more effective than students in the feedback-only group. Guidelines are provided for the use of development portfolios and the organization of supervision meetings in on-demand vocational education.

Introduction

In Dutch secondary vocational education, many institutes introduce on-demand education because it is nowadays acknowledged that students should be given more control over and responsibility for their own learning. In on-demand education, students are given the opportunity to direct their own learning by selecting learning tasks that fit their needs and interests. From a theoretical perspective, giving students the opportunity to direct their own learning may have positive effects on learning results because it gives them the opportunity to adapt the learning tasks to their learning needs, that is, students only need to learn what is directly relevant for them. This approach makes learning also more personally relevant and stimulates intrinsic motivation (Williams, 1996). In addition, letting students experience responsibility for their learning helps them to develop self-directed learning skills and to prepare for lifelong learning as independent learners (Kriewaldt, 2001).

However, from an empirical perspective, it has been found that students experience difficulties in directing their own learning, and that self-directed learning often has a negative influence on students' learning results (Van Velzen, 2002; Williams, 1996). In secondary vocational education, students often come from a tradition of teacher-directed learning in which it is mainly the teacher who assesses their performance, indicates their learning needs, and selects appropriate learning tasks for them to fulfill these needs (Shephard, 2000). A teacher-directed learning environment often makes little or no appeal to students' self-directed learning skills, which may hamper students to develop skills such as self-assessment of performance, formulation of learning needs, and selection of suitable learning tasks (Knowles, 1975). When students who are used to a teacher-directed learning environment suddenly enter an educational setting which demands them to direct their own learning, their lack of self-directed learning

skills may impede them in becoming successful independent learners. It is therefore of utmost importance that students who enter on-demand education are provided with sufficient support to develop their self-directed learning skills.

In a system of on-demand education, a portfolio with a focus on the student's learning process is often used for formative assessment. Compared to portfolios with a focus on learning products (e.g., showcase portfolios), such portfolios have been advocated by many theorists as promising tools to help students become reflective and self-directed learners. They are called development portfolios (Kicken, Brand-Gruwel, & van Merriënboer, 2008; Kicken, Brand-Gruwel, van Merriënboer, & Slot, in press), learning portfolios (Driessen, van Tartwijk, Overeem, Vermunt, & van der Vleuten, 2005; Klenowski, 2002) or process-folios (Järvinen & Kohonen, 1995; Seidel, Walters, Kirby, Olff, Powell, & Veenema, 1997). In this article, we will use the term 'development portfolio' to refer to a portfolio students use to describe and document multiple aspects of their own professional development over time. It may contain, for instance, reflections on prior performance, photographs or video clips of performance and products, formulations of weaknesses and strengths in performance, and decisions about courses of actions to improve performance.

Several studies reported that development portfolios are effective tools to help students reflect on their learning and to think about the development of their skills (Chen, Liu, Ou, & Lin, 2000; Driessen et al., 2005; Mansvelder-Longayroux, Beijaard, & Verloop, 2007; Zeichner & Wray, 2001). Especially in the field of teacher training, the construction of so-called 'teaching portfolios' by student teachers has been found to encourage them to think more deeply about teaching and to focus on key issues (Wade & Yarbrough, 1996), and to understand and improve their skill development (Mansvelder-Longayroux et al., 2007). In addition, research has shown

that students who developed a habit of reflection through the use of a portfolio, continued to reflect on their own performance well after the initial experience of portfolio construction (Grant & Huebner, 1998).

Studies on the effectiveness of portfolio use, however, also indicate limitations of the development portfolio as a tool to improve students' self-directed learning (Driessen et al., 2005; Wade & Yarbrough, 1996). Student characteristics and/or environmental factors can negatively influence the frequency or quality of portfolio use (Chen et al., 2000; Levett-Jones, 2005). For instance, students may lack experience in using a portfolio to direct their own learning, the portfolio may provide insufficient structure because of its open nature, or the learning environment may offer too little guidance on how to use the portfolio adequately. All these factors may lead to a suboptimal use of the portfolio or even no use of it at all, or they will mean that students add only superficial and short reflections to the portfolio, undermining its positive effects.

Recommendations to realize effective use of development portfolios pertain, on the one hand, to the proper introduction of the portfolio in the educational setting, and, on the other hand, to the guidance given to students once it has been introduced. When introducing the portfolio, students should be well informed on the concept of self-directed learning, and the aims and use of the portfolio to help them direct their own learning (Kicken, Brand-Gruwel, van Merriënboer, & Slot, in press; Levett-Jones, 2005; Van Tartwijk, Driessen, van der Vleuten, & Stokking, 2007). Next, the use of the development portfolio should be properly guided by including structured aspects or providing formats to balance its open-ended nature (Wade & Yarbrough, 1996). Ideally, the portfolio should be structured in such a way that it helps students to systematically assess their own performance (i.e. self-assessment), formulate their learning

needs, and select suitable future tasks to fulfill these needs. In particular, Driessen et al. (2005) revealed that coaching by a supervisor is one of the most crucial conditions for successful use of portfolios in medical education. Several other studies confirm this finding (Dagley & Berrington, 2005; Kjaer, Maagaard, & Wied, 2006; Smith & Tillema, 2003), showing that students are best guided, and also prefer to be guided, by a supervisor or mentor when constructing a development portfolio.

A related promising way to guide students in their use of a development portfolio is to provide them with specific advice on how to improve their self-directed learning skills. Two decades ago, researchers were already proposing learner-controlled instruction (for an overview, see Williams, 1996). In order to ensure that learners make effective decisions in such instruction, it is important to inform or advise them regularly. This counteracts the negative effects of little or incorrect prior knowledge (Hannafin 1984; Milheim & Martin, 1991; Steinberg, 1989; Tennyson & Buttery, 1980; Tennyson, Tennyson, & Wolfgang, 1980). Research in the field of computer-assisted learning has shown positive effects on learning when students are provided with advice on tool use (Clarebout & Elen, 2008; Gräsel, Fischer, & Mandl, 2000; Lee & Lehman, 1993) or advice on their development and on what their instructional needs are (Bell & Koslowski, 2002; Tennyson, 1980, 1981). For self-directed learning skills, similar positive effects may be expected of advice providing students with direction to effectively improve these skills through 'deliberate practice'. This construct is defined as time spent with the intention to improving one's performance. Effective application requires that students remain focused on the task and monitor their performance to retrieve cues for further improvement of performance (Ericsson, Krampe, & Tesch-Römer, 1993). The advice may, for instance, pertain to students' ability to assess their own performance (e.g., "the task you performed is assessed very differently by you and your

peer: Ask why s/he thinks your performance is weak"), to formulate learning needs (e.g., "you do not prioritize your weak points: Decide what is for this moment the most important aspect of your performance that needs to be improved"), and to select new learning tasks (e.g., "you frequently choose tasks that are not relevant for your learning needs: Familiarize yourself with relevant features of tasks available to you before making your final selection"). Well-designed advice consists of both feedback and feedforward. In this way, advice indicates strengths and weaknesses (i.e., verification) and, especially, provides constructive individualized information to improve learning (i.e., elaboration, see also Black & William, 1998).

Giving portfolio-based advice to students combines the advantages of on-demand education and teacher-directed education. The advice provides students with the necessary information to effectively direct their learning, and at the same time it gives them control over their own learning because they are free to choose whether and how to use the given information. Research on the use of advice has shown that students who direct their own learning on the basis of given advice perform equally well as students who learn under program or system control (Santiago & Okey, 1992; Corbalan, Kester, & van Merriënboer, 2006). Thus, giving specific advice seems to be a fruitful approach to making on-demand education more 'safe', limiting the risk that students undertake activities that do not contribute to learning at all.

The present study was designed to investigate the effects of giving students specific portfolio-based advice on the development of their self-directed learning skills. In an on-demand hairdressing program in vocational education, one group received feedback only and another group received specific portfolio-based advice (i.e., feedback and feedforward) in regular supervision meetings. Effects are studied on the development of self-directed learning skills (i.e., self-assessment of performance, formulating learning needs, and selecting new learning tasks),

learning results, and student perceptions of the effectiveness of the supervision meetings. First, it was hypothesized that students who received advice would better develop self-directed learning skills than students who received feedback only. Second, it was hypothesized that students who received advice would demonstrate better learning. Third, it was hypothesized that students who received advice would appreciate the supervision meetings more and perceive them as more effective than students who received feedback only.

Method

Participants

Participants were 43 students (42 female, 1 male; mean age = 18 years, *SD* = 1.2) in their first year of a three-year hairdressing program of a school for secondary vocational education in the South of the Netherlands. They were randomly assigned to the advice condition (*n* = 21) and the feedback-only condition (*n* = 22).

Four supervisors (all female; mean age = 32 years, *SD* = 6.78) participated in the study. Because teaching experience differed between supervisors, first the two experienced and next the two less experienced supervisors were randomly assigned to the conditions. Thus, one experienced and one less experienced supervisor participated in each condition.

Materials

Educational program. To become a certified hairdresser, students must in a period of three years acquire 8 main skills and 57 subskills. The eight main hairdressing skills are, in order of complexity: (1) caring and diagnosing hair (with 3 subskills, i.e., washing and shampooing the hair, massaging the scalp, and diagnosing the condition of hair and scalp); (2) permanent waving (with 6 subskills); (3) cutting hair (with 12 subskills); (4) blow-drying hair (with 4 subskills); (5) curling straight hair and straighten curled hair (with 11 subskills); (6) coloring hair (with 5

subskills); (7) cutting hair with a knife (with 6 subskills), and (8) styling long hair (with 10 subskills).

The pedagogy of the program may be characterized as on-demand education with a mix of theoretical and practical work inside and outside school. A typical school week consists of two days of practical lessons at school, one day of theoretical lessons at school, and two days of internship at a hairdressing salon. Students choose the main skill and subskills they want to develop and select the learning tasks to develop these skills. The learning tasks vary in complexity, authenticity, and amount of given support (van Merriënboer & Kirschner, 2007). With regard to *complexity*, learning tasks differ in the difficulty of practiced skills (e.g., cutting hair is more difficult than permanent waving) and combinations of skills (e.g., washing and cutting hair is more difficult than only washing hair). With regard to *authenticity*, learning tasks differ in the object on which the skills are performed (e.g., performance on a hairdressing dummy, a peer student, or a real client), the context in which they are performed (e.g., in school or in the hairdressing salon), and the constraints under which they are performed (e.g., with or without time pressure). With regard to *task support*, learning tasks differ in the level of students' independency (e.g., observing the teacher or peer student performing the task, performing only a part of the task, or independently performing the whole task).

After performing a number of equivalent tasks without support (called 'practical assignments'), a student may ask the teacher to summatively assess her performance on the main skill and subskills central to these tasks. When all subskills of a particular main skill (e.g., cutting hair, coloring hair) are up to standards and approved by the teacher, students are allowed to register for a formal examination of that main skill conducted by an external test institute. They receive a certificate after successful completion of the examination. During the school year,

students have several opportunities to participate in a formal examination. Per examination opportunity, students can register for a maximum of two formal examinations.

The current study was conducted in the first year of the hairdressing program. For this first year, which consisted of 40 study weeks, students were advised to fully develop at least three main hairdressing skills (i.e., take the formal examinations), and to start developing two other main skills.

Development portfolio. To help students take responsibility over their own learning process and make effective choices, a web-based development portfolio called Structured Task Evaluation and Planning Portfolio (STEPP; Kicken et al., in press) was designed and implemented. STEPP has three functionalities, helping students to (a) assess their own task performance, (b) formulate learning needs based on assessed shortcomings in task performance, and (c) select future learning tasks with characteristics that help to fulfill the formulated learning needs.

To assess their own performance on particular learning tasks, students select the performed skills from a hierarchical menu with all possible hairdressing skills and subskills. (Figure 1). Next, they assess the quality of their performance in relation to given performance standards (i.e., criteria, values, and attitudes), which are automatically shown by STEPP as soon as the student has selected a skill. Performance standards pertain to *criteria*, which indicate straightforward requirements in terms of time, accuracy, and order of activities (e.g., apply hair-dye in no more than 10 minutes); *values*, which indicate particular conventions and ways of working in the hairdressing profession (e.g., wear gloves during application of hair-dye), and *attitudes*, which indicate desirable aspects of behavior (e.g., be friendly to clients). For each selected skill, STEPP provides a matrix with standards on one dimension and scales for rating

the performance in relation to each standard on the other dimension. Students assess their own performance (i.e., self-assessment) by filling out the matrix. In addition to their self-assessments, students may also ask their teacher to assess their performance by filling out the same matrix. The self-assessments form the basis for the identification of learning needs. The students can formulate their needs for future learning in a separate textbox (Figure 1). For instance, a student may enter the text *"I need to learn to apply hair-dye in less time"* into the textbox.

< Insert Figure 1 about here>

For the selection of future learning tasks, STEPP offers students the opportunity to select skills and subskills that need further practice to meet the formulated learning needs (Figure 2). For all learning tasks that offer the opportunity to improve the selected skill(s) the student can indicate the desired level of complexity, authenticity, and available support. For instance, the student who wants to speed up hair-dying skills may do so by observing the teacher performing the skill of hair dying, by dying the hair of a dummy, by dying the hair of a real client, and so forth. This yields a working plan with learning tasks for every forthcoming week.

< Insert Figure 2 about here>

Finally, STEPP provides three structured overviews with all portfolio data necessary for discussing a student's progress during supervision meetings. The first screen presents all learning tasks performed by the student, together with the corresponding self-assessments and, if applicable, teacher assessments (Figure 3). The second screen gives an overview of all formulated learning needs. The third screen presents the working plan for the forthcoming week.

< Insert Figure 3 about here>

Supervision meetings. In both the advice and the feedback-only group, the supervisor provided feedback on the student's progress report and planning of learning over the last two

weeks, while discussing the three above mentioned STEPP overview screens in a fixed order. In the advice condition, the supervisor provided not only feedback but, in addition, advice on how to improve self-directed learning skills. Using the first overview (assessments), advice was given on how to improve the quality of the self-assessments. For instance, if discrepancies between self-assessments and teacher assessments occurred, students were advised to consult their teacher and ask for clarification. Using the second overview (learning needs), advice was given on how to trace and formulate learning needs. For instance, when needs were formulated in a general and abstract way (e.g., I should talk more), students were given tips on how to make them more specific (e.g., I need to think about interesting topics I could talk about with clients). Using the third overview (working plan), advice was given on how to select learning tasks appropriate to fulfill the formulated learning needs. For instance, when selected learning tasks were already mastered and mainly chosen because a student liked them, the student was advised to include also tasks that offered the opportunity to practice new skills or skills not yet mastered at an adequate level.

The supervisors in the feedback-only and advice condition structured their meetings according to different protocols (see the Appendix). All supervisors were trained to use the condition-specific protocol appropriately before data collection started. First they studied the protocol and discussed it with the experimenter. Then the experimenter demonstrated the use of the protocol and answered any remaining questions of the supervisors. For each supervisor, the experimenter also organized a simulated supervision meeting in which dummy portfolios were used and the experimenter played the role of the student. The supervisors were provided with feedback on their use of the protocol. They provided students only during the intervention period with the advice or the feedback.

Measurement Instruments

A large number of measures was used in this study. For reasons of clarity, Table 1 provides an overview of the variables portfolio-use, protocol use during supervision meetings, success and failure factors for participation in formal examinations, perceived effectiveness of supervision meetings, self-assessment, formulation of learning needs, task selection, and learning results. These variables and the instruments used to measure these variables will be further discussed in the following sections.

< Insert Table 1 about here >

Observation scheme supervision meetings. An observation scheme was constructed to check for the correct use of the protocols during the supervision meetings. For each observed supervision meeting, the supervisor could receive (a) 1 point for each of the three discussed screens (min-max: 0-3), (b) 1 point for providing feedback on students' progress (min-max: 0-1), and (c) 1 point for refraining from providing advice in the feedback-only condition or providing effective advice in the advice condition (min-max: 0-1). If supervisors scored less than the maximum of 5 points (i.e., 3 points for consulting the screens, 1 point for providing feedback, plus 1 point for refraining/providing advice), the supervision meeting was marked as inadequate.

Student interviews. A semi-structured interview consisting of four parts was conducted to measure: (1) actual portfolio use, (2) the contents of the supervision meetings, (3) perceived self-directedness in learning, and (4) perceived effectiveness of the supervision meetings. The questions concerning Part 1 and Part 2 served as control variables. The aim was to control for possible differences in portfolio use between the conditions, and to verify whether the contents of the supervision meetings for both conditions were in agreement with the condition-specific protocols. The other two parts (Part 3 and Part 4) measured dependent variables.

In Part 1, open-ended questions were asked concerning the usability of the portfolio and the degree to which students had integrated its use in their daily routine. Answers regarding the use of the portfolio were scored as (a) irregular use, (b) use only when reminded by the teacher, or (c) routine use. Regarding usability, students' answers were simply scored as (a) easy to use, or (b) difficult to use.

In Part 2 of the interview, students were asked to describe the topics that were discussed during the supervision meetings and the information given by the supervisor. Answers provided to this question were scored as (a) appropriate, or (b) inappropriate according to the condition-specific protocol.

To examine students' perceived self-directedness, Part 3 of the interview consisted of one open-ended question asking students to explain why they either had been able (success) or not been able (failure) to participate in one or more of the formal examination opportunities. Based on an analysis of the given answers, students' explanations of the *success* to participate in a formal examination were assigned to four categories: (a) frequent deliberate practice while focusing on inadequate aspects of performance, (b) drawing up a good year plan (i.e., what examinations do I want to register for and which practical assignments do I need to accomplish) and acting according to this plan, (c) taking care of practical preconditions to be able to carry out the work plan and year plan (e.g., make appointments with human models to practice hairdressing skills), and (d) intrinsic motivation or ask for help in case of declining motivation. Multiple classifications were possible. Based on an analysis of their given answers students' explanations of their *failure* to participate in a formal examination were assigned to four categories: (a) not being able to develop skills sufficiently before the deadline of the examination because the skill is experienced as difficult or complex, (b) lack of drawing up a year plan or the

discipline to act according to this plan, (c) not taking care of practical preconditions to be able to carry out the plan, and (d) lack of motivation. Again, multiple classifications were possible.

In Part 4 of the interview, one open-ended question was asked on the perceived effectiveness of the supervision meetings, that is, if the meetings were useful and helped to learn to direct own learning. The answers were categorized as (a) meetings were superfluous, (b) the supervisor helped me to analyze my performance, (c) the supervisor helped me to improve my process of task selection, and (d) the supervisor motivated me to further develop skills and to carry out my planning of tasks.

Quality of self-assessments. To rate the quality of students' self-assessments, log files were used to calculate the proportion of agreement between student and teacher assessments. The hairdressing skill 'permanent weaving' and its three subskills (i.e., three different weaving patterns) were used to calculate the proportion of agreement. This skill was selected and assessed at least five times, in both conditions, by both the students and their teacher. Each time the student performed the skill or one of its subskills, it was assessed by both the student and the teacher scoring the corresponding set of performance standards as either failed (F), satisfactory (S), or very good (V). The proportion of agreement between the student's and the teacher's assessment scores on the set of relevant performance standards was calculated per performed task, separately for each of the three subskills.

Diagnostic formulated learning needs. The coding system of Mansvelder-Longayroux, Beijaard and Verloop (2007) was adapted to code the learning needs retrieved from the log files. All learning needs were coded as either an 'assessment result' or a 'diagnosis'. A learning need was coded as an 'assessment result' when it literally repeated the performance standard which was rated as failed. For instance, if a student scored 'failed' for 'time on task' and then

formulated the learning need "*shorten the time*" or "*work faster*", this learning need was coded as an assessment result. The learning need was coded as 'diagnosis' if students identified potential causes of weak performance. For instance, taking the previous example, the learning need *"I should better master the technique of holding the comb and scissors in one hand to speed up my performance"* is diagnostic in nature because the student has diagnosed a possible reason for slow performance. Two raters were trained to code the learning needs after which they independently coded the learning needs. The inter rater reliability was .72, which is substantial (Landis & Koch, 1977). The proportion of learning needs formulated as a diagnosis was calculated for each student.

   Quality of task selection. Students received a score ranging from 0 to 4 for each task selection retrieved from their log files. A task selection consisted of four elements and for each element students could receive 1 point. Points were given for: (a) a relevant skill, (b) for a relevant learning need, (c) at an appropriate level of difficulty, and (d) with an appropriate level of authenticity. The relevance and appropriateness of a task selection were based on the registered assessments and formulated learning needs from the previous weeks. For example, a student's overview in the portfolio might indicate that she has difficulties placing the curlers during permanent weaving according to the pattern and firm enough when this task is independently performed on a human model. If this student selects a task in which permanent weaving (1 point for relevant skill) is practiced on a human model (1 point for appropriate authenticity level), but now with help from a teacher (1 point for appropriate, lower difficulty level) to improve the correctness and firmness of the pattern (1 point for relevant learning need), this student receives the maximum of 4 points for this particular task selection.

*Learning results.*The number of acquired certificates (max. = 4) during the formal examination and the number of practical assignments (i.e., tasks that are performed without any support) approved by the teacher (max. = 16) were counted for each student.

Procedure

For data collection, two periods were distinguished: a familiarization period of 10 weeks and the intervention period of 20 weeks. At the beginning of the familiarization period, all students followed two lectures in which the concepts of on-demand education and self-directed learning were explained. The use of the development portfolio and the attendance of supervision meetings in relation to the development of self-directed learning skills were also discussed. All students received two training sessions in the use of STEPP and were advised to update their portfolio at least once a week in order to practice their self-assessment skills, think about their learning needs, and plan future learning tasks. In the familiarization period, extra training sessions were scheduled for students who needed additional practice in using STEPP (e.g., students with low computer skills). The supervisors were also trained in the use of STEPP. During simulated supervision meetings, they practiced the use of the condition-specific protocol in combination with the three overviews from the portfolio (assessments, learning needs, and task selections). Also in the familiarization period, the collection of log files and approved practical assignments started and continued until the end of the intervention period. However, there was no opportunity for participation in a formal examination in the familiarization period.

In the intervention period, every two weeks a supervision meeting was scheduled for each student. These meetings were set up according to the condition-specific protocol for the advice group and the feedback-only group. The experimenter attended at random four supervision meetings of each supervisor (i.e., 16 meetings in total) to observe whether the protocols were

correctly used. During the intervention period, students had two opportunities to participate in a formal examination. For each examination, students could acquire a maximum number of two certificates. After this period, individual interviews were held with 15 randomly selected students from each condition. All student answers were directly typed out and the respondents were asked to verify the transcripts. Next, the answers were categorized and scored.

Results

To gain insight in how students used the portfolio, we provide information on the actual use of STEPP in this section. Observation results regarding the correct implementation of the protocols for the supervision meetings in the two conditions are presented next. Then, the results concerning the effects of giving advice in supervision meetings are described for (a) self-directed learning skills, (b) learning results, (c) perceived self-directedness, and (d) perceived effectiveness of the supervision meetings.

Actual Use of STEPP

Students participated in two practical sessions per week. This allowed them to perform at least three learning tasks per week, which could be assessed by themselves and, on request, by their teacher. Learning needs could be formulated for each task. In addition, student could select tasks to be performed in the next week(s). If students updated STEPP weekly, this would yield at least 30 (i.e., 3 learning tasks per week x 10 weeks) self-assessments, teacher assessments, and formulated learning needs after the familiarization period, and another 60 (i.e., 3 learning tasks per week x 20 weeks) after the intervention period. Actual STEPP-use was determined by analyzing the log files and responses to Part 1 of the student interview.

For the conditions within each period, Table 2 presents the means and standard deviations of the number of self-assessed tasks, the proportion of teacher assessments (i.e., the number of

tasks assessed by both the teacher and the student, divided by all self-assessed tasks), the number of formulated learning needs, and the number of task selections made. During the familiarization period, students in the feedback-only and advice condition updated STEPP with about 10 and 13 self-assessments, respectively. ANOVAs showed no significant differences in actual STEPP use between conditions, $F(1, 41) = 1.54$, $MSE = 70.70$, $p = .22$, $\eta_p^2 = .04$. With respect to teacher assessments, students asked for a teacher assessment in only 11 and 14 percent of the cases, respectively. Again, no differences were found between conditions $F(1, 41) = .61$, $MSE = 1.48$, $p = .44$, $\eta_p^2 = .02$. Students in both conditions formulated about 11 learning needs, $F(1, 41) = .05$, $MSE = 79.07$, $p = .83$, $\eta_p^2 = .00$, and STEPP was updated by students in both conditions with one task selection about every two weeks, $F(1, 41) = .51$, $MSE = 4.35$, $p = .48$, $\eta_p^2 = .01$.

<Insert Table 2 about here>

In the intervention period, students in both the feedback-only and advice condition updated STEPP with somewhat more than one self-assessment every two weeks, $F(1, 41) = .00$, $MSE = 140.69$, $p = .99$, $\eta_p^2 = .00$, and asked for a teacher assessment in 27 to 35 percent of the cases, respectively. No significant differences were found between conditions, $F(1, 41) = .40$, $MSE = 6.08$, $p = .53$, $\eta_p^2 = .01$. The number of formulated learning needs was close to the number of self-assessments, and no differences were found between both conditions, $F(1, 41) = .07$, $MSE = 269.28$, $p = .79$, $\eta_p^2 = .00$. Students in both conditions updated STEPP with about one task selection every three weeks, $F(1, 41) = .45$, $MSE = .51$, $p = .51$, $\eta_p^2 = .01$.

The interview results (Part 1) concerning integration of STEPP in daily routine reveal that 53 % of the students in the feedback-only condition and 60 % of the students in the advice condition integrated STEPP in their daily or weekly routine. In both conditions, 27 % of the interviewed students only used STEPP when the teacher reminded them to use it. Finally, 20 %

of the students in the feedback-only condition and 13% of the students in the advice condition reported irregular use. All students in both conditions reported that STEPP was easy to use. Chi-square analyses showed no significant differences between conditions.

Because actual STEPP use did not significantly differ between conditions in either the familiarization period or the intervention period, no covariates have been included in further analyses on the dependent variables.

Observation of the Supervision Meetings

To verify if the supervision meetings were held according to the condition-specific protocols, the experimenter observed and scored 16 supervision meetings (four of each supervisor). All 16 meetings received the maximum score of five points, indicating that all meetings were in agreement with the protocol. Moreover, results from Part 2 of the student interview revealed that the discussed topics during the meetings were in line with respect to the condition-specific protocol (mentioned by 95% of the students in both conditions). From these results it can be concluded that the condition-specific protocols were successfully implemented.

Self-Directed Learning Skills

To answer the question concerning the effects of giving advice during supervision meetings on students' self-directed learning skills, the proportion of agreement between the students' self-assessments and the teachers' assessments, the proportion of diagnostic learning needs, and the quality of task selections were determined for both conditions (see Table 3). As expected, no significant differences between conditions were found during the familiarization period. So, these variables were not added as covariates when analyzing the data for the intervention period. In case of significant differences between the two conditions, an additional

19

ANOVA was conducted with supervisor rather than condition as a fixed factor, to verify that the positive effect of condition was not caused by a specific supervisor.

Quality of Self-Assessments

The lower part of Table 3 indicates that the self-assessment skills are not fully mastered during the intervention period, because the proportion of agreement with the teacher is for both conditions below .65. Contrary to our expectations, ANOVA showed no significant difference between conditions, $F(1, 30) = 2.40$, $MSE = 0.63$, $p = .13$, $\eta_p^2 = .07$.

< Insert Table 3 about here>

Diagnostic Formulated Learning Needs

During the intervention period, students in the advice condition formulated significantly more learning needs with a diagnostic nature (61% of all formulated learning needs) than students in the feedback-only condition (32%), $F(1, 41) = 24.97$, $MSE = .04$, $p = 0.00$, $\eta_p^2 = .39$. In addition, no supervisor within condition effect was found, $F(1, 41) = 1.56$, $MSE = .04$, $p = .22$, $\eta_p^2 = .07$, meaning that the positive effect on the proportion of diagnostic learning needs was not caused by a specific supervisor.

Quality of Task Selections

Table 3 indicates that the average scores in the intervention period are above 3.5, meaning that students are quite capable to select suitable tasks. Students in the advice condition show a marginally, but not significantly, higher quality of task selections ($M = 3.91$; $SD = .17$) than students in the feedback only condition ($M = 3.70$; $SD = .33$), $F(1, 20) = 3.49$, $MSE = .07$, $p = .06$, $\eta_p^2 = .17$. No supervisor within condition effect was found.

Learning Results

The number of practical assignments approved by the teacher were counted to measure students' learning results. In the intervention period, students in the advice condition completed significantly more practical assignments (M = 10.05, SD = 4.68) than students in the feedback-only condition (M = 6.32, SD = 3.70), $F(1, 41)$ = 12.04, MSE = 16.53, p = 0.00, η_p^2 = .236. No supervisor within condition effect was found, $F(1, 41)$ = 2.45, MSE = 16.53, p = .10, η_p^2 = .11. The number of acquired certificates was also counted as an indicator of learning results (certificates could only be acquired in the intervention period). Students in the advice condition acquired slightly more certificates (M = 1.14, SD = 1.42) than students in the feedback-only condition (M = 0.59, SD = 1.10), $F(1, 41)$ = 3.04, MSE = 1.60, p = .0.09, η^2 = .07. Again, no supervisor within condition effect was found, $F(1, 41)$ = 1.07, MSE = 1.60, p = .35, η_p^2 = .05.

Perceived Self-Directedness

To determine students' perceived self-directedness, they were asked to explain the reason for their success or failure to participate in one of the formal examinations. Table 4 presents the percentages of student answers categorized according to the eight categories described in the Method section (student interview, Part 3). All 5 students from the feedback-only condition who participated in a formal examination reported that drawing a year plan helped them to participate. All 9 students from the advice-condition who participated in a formal examination mentioned deliberate practice as the most important success factor, and all but one (89%) also mentioned drawing up a year plan as an important success factor. Chi-square analyses showed no significant differences between the two conditions regarding their explanations for success.

Students' explanations for *not* participating in a formal examination reveal that 62 % of the 13 students in the feedback-only condition mentioned not drawing a well-defined year plan

as the most important failure factor, whereas only 17 % of the 12 students in the advice condition gave this as a reason, $\chi^2(1, N = 25) = 5.04, p = .02$. Thus, students in the feedback-only condition were more likely than those in the advice condition to blame their lack of planning skills. The other factors were not significantly different between conditions. As can be seen from Table 4, students in both conditions mentioned not taking care of practical preconditions and, to a somewhat lesser degree, difficulty with skill performance as important failure factors.

< Insert Table 4 about here>

Perceived Effects of the Supervision Meetings

Results regarding the perceived effectiveness of the supervision meetings (student interview, Part 4) showed that significantly fewer students in the advice condition (20%) than in the feedback-only condition (67%) perceived the supervision meetings as superfluous, $\chi^2(1, N = 30) = 6.65, p = 0.00$. More students in the advice condition (73%) than in the feedback-only condition (13%) mentioned that the supervision meetings had helped them to analyze their own performance, $\chi^2(1, N = 30) = 11.00, p = 0.00$. Finally, more students in the advice condition (67%) than in the feedback-only condition (13%) mentioned that the meeting helped them to improve their task selections, $\chi^2(1, N = 30) = 8.89, p = 0.00$. In both conditions, 73% of the students perceived the supervision meetings as motivating for further skill development and for carrying out their learning plans.

Discussion

The current study investigated the effects of providing students in secondary vocational education with portfolio-based advice on their ability to self-assess performance, formulate diagnostic learning needs, and make appropriate task selections. The hypothesis that students who receive advice develop their self-directed learning skills better than students who only

22

receive feedback is largely confirmed by the reported results. The effect of giving advice was evident in students' ability to formulate learning needs. Students who received advice were better able to diagnose possible cause(s) of their weaknesses and formulated relatively more diagnostic learning needs than students who only received feedback. Regarding task selection skills, a small effect was found in favor of the advice group. No difference was found between the two conditions, however, on self-assessment skills.

Our second hypothesis, that students in the advice condition would show higher learning results than students in the feedback-only condition, was supported. Students receiving advice checked off more assignments and passed slightly more formal examinations than students who only received feedback.

Finally, the third hypothesis that students who receive advice during supervision meetings will appreciate these meetings more and perceive the sessions as more effective than students who only receive feedback was also supported. Students apparently do not only want to be informed on the appropriate or inappropriate application of their self-directed learning skills (i.e., verification), but also want to be supported in learning how to improve these skills (i.e., elaboration).

It appeared from this study that supervision meetings in which students receive advice (i.e., feedback and feedforward) had different effects on the development of various self-directed learning skills. With respect to self-assessment skills, students did not reach a stage where they were able to assess their own performance at a sufficient level. The proportion of agreement with the teacher assessments was only about 65%. Training more extensively, observing the modelling of assessment skills by an expert, and discussing the differences between expert and own assessments are examples of instructional methods that might contribute to a better

development of self-assessment skills. Research by McDonald and Boud (2003) reveals that self-assessment skills can well be trained. In their study, students who received systematic self-assessment training outperformed students who had not been involved in this training. In addition, research by Falchikov and Boud (1989) reveals that learning to assess own performance is a complex cognitive process that develops over a relatively long period of time, and its accuracy depends on how well-defined the assessment standards are. Possibly, the training period in the current study was too short or students had too much difficulty with interpreting the given standards in order to obtain the desired effects.

With respect to task selection skills, it appears that providing feedback on task selection is to some extent effective, provided that students select tasks from a limited number of available tasks as in the current study. When the number of tasks to choose from is higher, it so happens that linking the information from the development portfolio (i.e., performance and learning needs) to the tasks' metadata (i.e., task features) in order to select the most appropriate task(s) is more difficult for learners (Kicken, Brand-Gruwel, & van Merriënboer, 2008). Thus, as the number of learning tasks from which the students choose increases, the need to provide good advice probably also increases.

Regarding the formulation of diagnostic learning needs, it appears that providing advice was clearly more effective than only providing feedback. Receiving advice prompted students to formulate diagnostic learning needs rather than learning needs that merely repeated the performance standards. In addition, results reveal that there was a positive effect of advice on students' learning results. This outcome is in line with findings on the effects of feedback (Black & William, 1998), indicating that more elaborate feedback has a positive effect on performance. Advising students to formulate learning needs in terms of cancelling causes of weak performance

can help them to gain more insight in the quality of their own performance. Identifying the causes of weak performance provides more direction for subsequent learning, resulting in a higher quality of deliberate practice which, in turn, has a positive effect on performance improvement (Ericsson et al., 1993). In on-demand education, the ability to formulate proper learning needs should therefore be one of the main topics to discuss with students in supervision meetings. This idea was also confirmed by the supervisors in our study, who conveyed that students often experienced problems with diagnosing their own performance.

Some remarks should be made regarding limitations of the design of the present study. First, the actual use of the development portfolio was rather low for all students, especially with respect to the number of selected learning tasks. In the interviews, students explained that they experienced the use of the portfolio as more and more superfluous, because after some time they "knew the assessment criteria and their learning needs by heart". To overcome this limitation, future research should be conducted in a learning environment in which heavier demands are made on the use of the development portfolio. Furthermore the portfolio should be adapted more to students' needs, supporting also the development of more complex self-directed learning skills in addition to helping them to become familiar with the assessment criteria and their personal learning needs. For instance, the portfolio could prompt students to unravel the processes that are causing weak performance and help them to draw up a plan for practice to improve performance.

A second limitation concerns the task selections made by the students. Due to practical reasons, such as problems finding human models on whom to practice hairdressing skills, the number and variety of learning tasks from which students could make their selection was rather small. This factor probably contributed to the small differences in quality of task selections between students in both conditions. To overcome this limitation, future research should be

conducted in a learning environment in which students can develop their skills by choosing appropriate learning tasks from a well designed, large task collection. The tasks available should be linked to metadata, such as difficulty level, support level, and focus of the task. Research on effects of different kinds of advice in a highly structured and well designed learning environment can provide more insight in the effects of advice on students' task selection behavior in relation to the quality of their performance.

A promising future approach to help students learn to direct their own learning is the use of reflective dialogue. Isaacs (1999) describes reflective dialogue as a process "where you become willing to think about the rules underlying what you do – the reasons for your thoughts and actions. You see more clearly what you have taken for granted" (p. 38). This technique was occasionally used during the practical training sessions. In the interviews, several students and supervisors indicated appreciating this communication technique because it systematically guides students in their thinking process. Future research could examine the different effects of reflective dialogue during instruction sessions, scheduled group sessions, and individual supervision meetings on students' self-directed learning skills and learning results.

Furthermore, future research could examine the possibility of programing advice or reflective dialogue into an automated system. For instance, the principles of reflective dialogue such as asking why and how questions could be integrated in the development portfolio, prompting students to engage in an analysis of their performance and to think about a plan to improve their performance. Designing such an automated system could assume some of the advising burden of supervisors. However, the combination of tool use (i.e., programmed advice) with advice provided by human agents should, in our opinion, remain part of the design of effective self-directed learning environments.

Another line of future research starts from the observation that, in addition to the three prompted self-directed learning skills, being able to plan for a longer period of time is also an important skill on which advice should be provided in the early stages of learning. Many students experienced problems with planning their learning trajectory for an extended period of time and/or with carrying out activities according to their plan, including taking care of preconditions. These barriers resulted in slow development of hairdressing skills and hence a small number of checked-off assignments and ineligibility to participate in formal examinations. Both students and supervisors indicated that the ability to plan a learning trajectory (e.g., for a semester or a whole school year) and to carry out activities according to plan are important skills that should be developed for efficient and effective learning in on-demand education. Future research should therefore also examine the effects of advice (e.g., learning contracts; Beitler, 2000; Knowles, 1986) on the ability of students to plan their learning over a longer period of time and taking care of preconditions.

It can be concluded from this study that the combination of using a development portfolio to stimulate students' self-directed learning skills and providing them with advice is a promising approach to improve their self directedness. This approach helps students to formulate directions for future learning which enhances the quality of their deliberate practice. However, to prevent students in secondary vocational education from practicing skills that are too easy or too difficult, with too much or too little support, or in an inappropriate context, it is advisable to gradually shift the responsibility for directing the learning process from the teacher to the students. Only when students demonstrate progress in using their self-directed learning skills, they should be given increasingly more responsibility to direct their own learning. In this way, students become confident self-directed learners.

Acknowledgments

References

Beitler, M. A. (2000). Contract learning in organizational learning and management development. In H. B. Long

(Ed.), *Practice and theory in self-directed learning* (pp. 63-70). Schaumberg, IL: Motorola University Press.

Bell, B. S., & Kozlowski, S. W. J. (2002). Adaptive guidance: Enhancing self-regulation, knowledge, and

performance in technology-based training. *Personnel Psychology,* 55, 267- 306.

Black, P., & William, D. (1998). Assessment and classroom learning. *Assessment in Education, 5*(1), 7-74.

Chen, G., Liu, C., Ou, K., & Lin, M. (2000). Web learning portfolios: A tool for supporting performance awareness.

Innovations in Education and Teaching International, 38, 19-30.

Clarebout, G., & Elen, J. (2008). Advice on Tool Use in Open Learning Environments. *Journal of Educational*

Multimedia and Hypermedia, 17 (1), 81-97.

Corbalan, G., Kester, L., & van Merriënboer, J. J. G. (2006). Towards a personalized task selection model with

shared instructional control. *Instructional Science, 34*, 399-422.

Dagley, V., & Berrington, B. (2005). Learning from an evaluation of an electronic portfolio to support general

practitioners' personal development planning, appraisal, and revalidation. *Education for Primary Care, 16,* 567-

574.

Driessen, E. W., van Tartwijk, J., Overeem, K., Vermunt, J. D., & van der Vleuten, C. P. M. (2005). Conditions for

successful reflective use of portfolios in undergraduate medical education. *Medical Education, 39*, 1230-1235.

Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of*

Educational Research, *59*, 395-430.

Ericsson, K. A., Krampe, R. T. H., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of

expert performance. *Psychological Review*, *100,* 363-406.

Grant, G., & Huebner, T. (1998). The portfolio question: A powerful synthesis of the personal and professional.

Teacher Education Quarterly, *25*(1), 33-43.

Gräsel, C., Fischer, F., & Mandl, H. (2000). The use of additional information in problem oriented learning

environments. *Learning Environments Research, 3*, 287-325.

Hannafin, M. J. (1984). Guidelines for using locus of instructional control in the design of computer-assisted

instruction. *Journal of Instructional Development, 7*(3), 6-10.

Isaacs, W. (1999). *Dialogue and the art of thinking together.* New York: Doubleday.

Järvinen, A., & Kohonen, V. (1995). Promoting professional development in higher education through portfolio assessment. *Assessment & Evaluation in Higher Education*, *20*(1), 25–36.

Kicken, W., Brand-Gruwel, S., & van Merriënboer, J. J. G. (2008). Scaffolding advice on task selection: A safe path toward self-directed learning in on-demand education. *Journal of Vocational Education and Training (60)*, 223 - 239.

Kicken, W., Brand-Gruwel, S., van Merriënboer, J. J. G., & Slot, W. (in press). Design and evaluation of a development portfolio: How to improve students' self-directed learning skills. *Instructional Science.*

Kjaer, N. K., Maagaard, R., & Wied, S. (2006). Using an online portfolio in postgraduate training. *Medical Teacher*, *28*, 708-712.

Klenowski, V. (2002). *Developing portfolios for learning and assessment.* London, UK: RoutledgeFalmer.

Knowles, M. (1975). *Self-directed learning: A guide for learners and teachers.* New York: Association Press.

Knowles, M. (1986). *Using learning contracts.* San Francisco, CA: Jossey-Bass Inc., Publishers.

Kriewaldt, J. (2001). A thinking geography curriculum. *Interaction*, 29. Retrieved November 01, 2007. Available from http://www.gtav.asn.au/interaction/issues/v29n4_dec01/metacognition.htm

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*, 159-174.

Lee, Y. B., & Lehman, J. D. (1993). Instructional cueing in hypermedia: A study with active and passive learners. *Journal of Educational Hypermedia and Multimedia, 2*(1), 25-37.

Levett-Jones, T. L. (2005). Self-directed learning: Implications and limitations for undergraduate nursing education. *Nurse Education Today, 25*, 3*63-368*.

McDonald, B., & Boud, D. (2003). The effects of self assessment training on performance in external examinations. *Assessment in Education, 10* , 210-220.

Mansvelder-Longayroux, D., Beijaard, D., & Verloop, N. (2007). The portfolio as a tool for stimulating reflection by student teachers. *Teaching and Teacher Education, 23*, 47-62.

Milheim, M. D., & Martin, B. L. (1991). Theoretical bases for the use of learner control: Three different perspectives. *Journal of Computer-Based Instruction, 18*(2), 99-105.

Santiago, R. S., & Okey, J. R. (1992). The effects of advisement and locus of control on achievement in learner-controlled instruction. *Journal of Computer-Based Instruction, 19*(2), 47-53.

Seidel, S., Walters, J., Kirby, E., Olff, N., Powell, K., & Veenema, S. (1997). *Portfolio practices: Thinking through the assessment of children's work.* Washington, DC: NEA Professional Library.

Shephard, L. (2000). The role of assessment in a learning culture, *Educational Researcher, 29*(7), 4-15.

Smith, K., & Tillema, H. (2003). Clarifying different types of portfolio use. *Assessment and Evaluation in Higher Education*, *28*, 625-648.

Steinberg, E. R. (1989). Cognition and learner control: A literature review. Journal of Computer-Based Instruction, 16(4), 117-124.

Tennyson, R. D. (1980). Instructional control strategies and content structures as design variables in concept acquisition using computer-based instruction. *Journal of Educational Psychology*, *72*, 525–532.

Tennyson, R. D. (1981). Use of adaptive information for advisement in learning concepts and rules using computer-assisted instruction. *American Educational Research Journal, 18*, 425-38.

Tennyson, R. D., & Buttery, T. (1980). Advisement and management strategies as design variables in computer-assisted instruction. *Educational Communication and Technology Journal, 28*(3), 169-176.

Van Merriënboer, J. J. G., & Kirschner, P. A. (2007). *Ten steps to complex learning*. Mahwah, NJ: Erlbaum / Routledge.

Van Tartwijk, J., Driessen, E., van der Vleuten, C., & Stokking, K. (2007). Factors influencing the successful introduction of portfolios. *Quality in Higher Education*, *13*(1), 69-79.

Van Velzen, J. H. (2002). *Instruction and self-regulated learning: Promoting students' (self-) reflective thinking* Unpublished doctoral dissertation, University of Leiden, The Netherlands.

Wade, R. C., & Yarbrough, D. B. (1996). Portfolios: A tool for reflective thinking in teacher education? *Teaching & Teacher Education, 12*(1), 63-79.

Williams, M. D. (1996). Learner-control and instructional technologies. In D. H. Jonassen (Ed.), *Handbook of research for educational communications and technology* (pp. 957- 982). New York: Simon & Schuster Macmillan.

Zeichner, K., & Wray, S. (2001). The teaching portfolio in U.S. teacher education programs: What we know and what we need to know. *Teaching and Teacher Education, 17*, 613-621.

Appendix

Protocols for Supervision Meetings in Both Conditions

Overview screen		Feedback-only Condition	Advice Condition
1. Assessments	Number of assessments	- If insufficient: tell to assess at least one skill or sub skill per week. - If sufficient: very good.	- If insufficient: advise to assess at least one skill or sub skill per week and explain that this helps to diagnose the weaknesses - If sufficient: very good and explain why.
	Number of teacher assessments	- If insufficient: tell to ask at least one teacher assessment for each skill or sub skill per week - If sufficient: very good	- If insufficient: advise to ask teacher for assessment because this helps to become aware of (causes for) weaknesses. - If sufficient: very good and explain why
	Agreement between teacher assessments and self-assessment	- If 2 or more discrepancies: point out the differences and tell to try to minimize the differences in assessment If 0-1 discrepancies: point out the difference + very good.	- If 2 or more discrepancies: point out the differences and advise to ask the teacher for clarification of the assessment to understand the differences in assessment - If 0-1 discrepancies: point out the difference + very good
2. Learning needs	Number of learning needs	If learning needs are not always formulated, tell to always formulate at least one learning need.	If learning needs are not always formulated, advise to always formulate a learning need, because this helps to become conscious of weaknesses
	Formulation of learning needs.	If formulated as assessment result, tell to try to use more own words, not merely copy the criteria. Tell to formulate more in terms of behavior and in terms of what to do in order to improve their performance.	If formulated as assessment result, advise to formulate learning needs more as an indication of what caused the weakness and in terms of behavior, and what the student should be working on. Diagnose the weakness with the student to give an example how to formulate the learning need.
	Comparison with learning needs formulated by others	Point out whether they are the same or not.	If learning needs of student and other assessor are not the same, advise to ask the teacher more often for an explanation of the learning needs she has pointed out. Ask teacher also during practical lessons for help while diagnosing own weaknesses. Advise to learn from teacher regarding diagnosis of weaknesses.
3. Working plan	Working plan	If nothing is selected, tell to always make a task selection and have the student make a selection directly after the meeting	If nothing is selected, advise the student to always make a selection and explain that this helps to deliberately practice weaknesses and helps to remind of practical preconditions.
	Selected skill	If not appropriate, only mention that it is not appropriate and that the student needs to look it over again, or leave it as it is, when she wants to.	If not appropriate, advise to select an other skill. Explain why the selected skill is not appropriate and provide advice to help the student in the process of task selection.
	Selected difficulty level	If not appropriate, only mention that it is not appropriate and that the student needs to look it over again, or leave it as it is, when she wants to.	If not appropriate, advise to select an other difficulty level. Explain why the selected skill is not appropriate and provide advice to help the student to select the appropriate level..
	Selected level of authenticity	If not appropriate, only mention that it is not appropriate and that the student needs look it over again, or leave it as it is, when she wants to.	If not appropriate, advise to select an other level of authenticity. Explain why the selected level is not appropriate and provide advice to help the student to select an appropriate level.

Table 1

Overview of Measurement Instruments Used in the Study

Variable	Measurement Instrument	Description	Score
Portfolio-use	Semi structured student interview Part 1	Usability and integration of portfolio use in daily routine	2 and 3 categories respectively
Protocol use during supervision meetings	Semi structured student interview Part 2	Topics of supervision meetings	2 categories
Success and failure factors for participation in formal examinations	Semi structured student interview Part 3	Factors influencing failure and success to participate in formal examinations	4 categories each
Perceived effectiveness of supervision meetings	Semi structured student interview Part 4	Perceived effectiveness supervision meetings	4 categories
Self-assessment skill	Quality of self-assessment	Proportion of agreement between student's and teacher's assessment score on a set of performance criteria.	0 – 1
Formulation of learning needs	Diagnostically formulated learning needs	Proportion of diagnostically formulated learning needs.	0 - 1
Task selection	Quality of task selection	Four aspects of task selections are scored. One point for each correctly selected aspect. Total score is calculated.	1 - 4
Learning results	Acquired certificates	Total number of acquired certificates is counted	0 - 4
	Approved practical assignments	Total number of approved practical assignments is counted.	0 - 16

Table 2

Means and Standard Deviations for the Actual Use of STEPP in Both Conditions for Both

Periods

	Feedback-only condition		Advice condition	
	M	*SD*	*M*	*SD*
Familiarization Period				
# Self-assessments[a]	9.77	7.5	12.95	9.27
Proportion of teacher assessments[c]	.11	.11	.14	.15
# Learning needs[a]	11.36	9.93	11.95	7.66
# Task selections[a]	4.59	4.10	5.24	3.24
Intervention Period				
# Self-assessments[b]	13.41	14.73	13.43	7.80
Proportion of teacher assessments[c]	.35	.28	.27	.14
# Learning needs[b]	15.82	21.10	14.48	9.20
# Task selections[b]	6.32	7.8	8.62	10.29

[a]Max. = 30. [b]Max. = 60. [c]Max. = 1.

Table 3

Means and Standard Deviations for Self-Directed Learning Skills in Both Conditions for Both

Periods

	Feedback-only condition		Advice condition	
	M	*SD*	*M*	*SD*
Familiarization Period				
Quality of self-assessment[a]	.46	.25	.59	.11
Diagnostic formulated learning needs[b]	.46	.31	.48	.27
Quality of task selections[c]	3.72	.25	3.80	.24
Intervention Period				
Quality of self-assessment[a]	.55	.16	.63	.16
Diagnostic formulated learning needs[b]	.32	.22	.61	.20
Quality of task selections[c]	3.70	.33	3.91	.17

[a] Proportion of agreement between student's and teacher's score on set of assessment criteria, min-max = 0-1.
[b] Proportion of diagnostic formulated learning needs, min-max = 0-1. [c] Score for task selections, min-max = 0-4.

Table 4

Perceived Self-Directedness Expressed as Reasons for Success and Failure to Participate in

Formal Examinations

Reasons for Success	Feedback-only condition ($n = 5$) Percent	Advice condition ($n = 9$) Percent
Success because of deliberate practice	60	100
Success because of year plan	100	89
Success because of practical preconditions	40	78
Success because of motivation	40	11
Reasons for Failure	Feedback-only condition ($n = 13$)	Advice condition ($n = 12$)
Failure due to skill difficulty	53	33
Failure due to lack of year plan	62	17
Failure due to practical preconditions	62	58
Failure due to lack of motivation	31	8

Figure Captions

Figure 1. Self-assessment and formulation of learning needs functionality in STEPP: Overview of skills (left column), standards for performance assessment (table), and possibility to formulate learning needs (textbox).

Figure 2. Task selection functionality in STEPP: Overview of skills (left column), table to indicate the selected task(s) in terms of skill, level of support, authenticity and learning needs, and options to view all formulated learning needs and the previous task selection (right).

Figure 3. Overview page in STEPP: All assessments by all assessors (right) of the performance on one particular task (left). F = fail, S = satisfactory, V = very good.