Peyton’s four-step approach: differential effects of single instructional steps on procedural and memory performance – a clarification study

Markus Krautter¹
Ronja Dittrich²
Annette Safi²
Justine Krautter¹
Imad Maatouk²
Andreas Moeltner²
Wolfgang Herzog²
Christoph Nikendel¹
¹Department of Nephrology, ²Department of General Internal and Psychosomatic Medicine, University of Heidelberg Medical Hospital, Heidelberg, Germany

Background: Although Peyton’s four-step approach is a widely used method for skills-lab training in undergraduate medical education and has been shown to be more effective than standard instruction, it is unclear whether its superiority can be attributed to a specific single step.

Purpose: We conducted a randomized controlled trial to investigate the differential learning outcomes of the separate steps of Peyton’s four-step approach.

Methods: Volunteer medical students were randomly assigned to four different groups. Step-1 group received Peyton’s Step 1, Step-2 group received Peyton’s Steps 1 and 2, Step-3 group received Peyton’s Steps 1, 2, and 3, and Step-3mod group received Peyton’s Steps 1 and 2, followed by a repetition of Step 2. Following the training, the first independent performance of a central venous catheter (CVC) insertion using a manikin was video-recorded and scored by independent video assessors using binary checklists. The day after the training, memory performance during delayed recall was assessed with an incidental free recall test.

Results: A total of 97 participants agreed to participate in the trial. There were no statistically significant group differences with regard to age, sex, completed education in a medical profession, completed medical clerkships, preliminary memory tests, or self-efficacy ratings. Regarding checklist ratings, Step-2 group showed a superior first independent performance of CVC placement compared to Step-1 group ($P<0.001$), and Step-3 group showed a superior performance to Step-2 group ($P<0.009$), while Step-2 group and Step-3mod group did not differ ($P=0.055$). The findings were similar in the incidental free recall test.

Conclusion: Our study identified Peyton’s Step 3 as being the most crucial part within Peyton’s four-step approach, contributing significantly more to learning success than the previous steps and reaching beyond the benefit of a mere repetition of skills demonstration.

Keywords: Peyton’s four-step approach, skills-lab training, procedural skills

Introduction

Skills-labs have become a valuable methodological teaching approach in medical education. Skills-lab training allows for sustained deliberate practice in a “mistake-forgiving”, safe environment. The positive effect of skills-lab training sessions on learning outcomes and transfer of acquired skills to clinical practice have been shown in a variety of different settings. Specific ingredients have been shown to contribute to a successful learning experience in skills-labs, such as predefined learning goals and curriculum integration, validity of the simulated scenarios, sustained deliberate practice, and feedback. Although there is evidence that the instructional approach used to introduce novel clinical technical skills to learners also has a decisive impact on subsequent skills performance, little is yet known about differential effects of varying instructional methods.
In the literature, a variety of instructional approaches that include multiple steps to convey clinical technical skills to learners have been described.\textsuperscript{14,15} A model that is becoming increasingly prevalent in medical education is “Peyton’s four-step approach”,\textsuperscript{16–18} which consists of four clearly defined instructional steps:

- **Step 1 – “Demonstrate”:** The trainer demonstrates the skill at a normal pace and without additional comments.
- **Step 2 – “Talk the trainee through”:** The trainer demonstrates the respective skill while describing each procedural substep in detail.
- **Step 3 – “Trainee talks trainer through”:** The trainer performs the skill for a third time, based on the substeps described to him by the trainee.
- **Step 4 – “Trainee does”:** The trainee performs the skill on his/her own.

A previously published study has shown decisive advantages of Peyton’s four-step approach in undergraduate medical education over standard instruction.\textsuperscript{13} While the intervention group was instructed in gastric tube insertion by applying Peyton’s four-step approach, the control group received a standard instruction based on the instructional principle of “see one, do one”, the traditional approach in which the skill is first demonstrated and explained by the trainer and then performed by the trainee him/herself.\textsuperscript{13,19} The first independent performance of gastric tube insertion was videotaped and rated by independent video raters on the basis of binary checklists\textsuperscript{20} and the integrated procedural performance instrument for the assessment of procedural skills in a clinical context.\textsuperscript{21,22} Video analyses showed that the group instructed by Peyton’s four-step approach not only performed more single procedural steps more accurately and had a more professional and self-assured skills execution, but also performed the skill faster than the comparison group. The authors assumed that Step 3 of Peyton’s four-step approach might be of critical importance for learners’ superior skills acquisition, as it integrates mental representation and motor imagery,\textsuperscript{23–25} which might lead to a more elaborate motor learning. However, it remains unclear whether the results can be traced back to a single specific step of Peyton’s four-step approach.

There is ample evidence of a plethora of various cognitive aspects influencing motor skills acquisition, such as the focus of attention,\textsuperscript{26–30} visuospatial competencies,\textsuperscript{31,32} performance observation and motor imagery,\textsuperscript{25,33} working memory,\textsuperscript{34} as well as declarative and procedural memory capacities.\textsuperscript{35} Due to the fact that declarative memory constitutes an indispensable prerequisite for high-value skills performance,\textsuperscript{36–38} it might be speculated that Peyton’s four-step approach also affects declarative memory for skill performance. One might assume that motor imagination, verbalization, and instructing the teacher, as conducted in Step 3 of Peyton’s four-step approach, leads to a deeper and more elaborate declarative processing of the single steps to be learned in terms of the Level of Processing Theory.\textsuperscript{39} However, to our knowledge, no studies have been conducted to date on the interplay between declarative memory performance and skills performance in skills-lab training.

The presented study was designed to evaluate the hypotheses that 1) different groups of trainees show more elaborate skills performance depending on how many single steps of Peyton’s four-step approach were completed during the training; 2) the group of trainees who underwent Step 3 show superior procedural performance compared to the group of trainees who receive a repetition of Step 2 instead of Step 3; and 3) this superior performance is also reflected and can be detected in a free recall memory performance test for procedural steps 1 day after the training. Therefore, we performed a clarification study that prospectively investigated the differential learning outcome of the different steps of Peyton’s four-step approach.

**Methods**

**Participants**

Trainees were recruited via advertisement, and written consent was provided by all study participants. All 3rd-year medical students at our faculty were contacted and asked to participate in the study. Ninety-seven students were recruited (equivalent of a 40.4% participation rate) and received a book voucher as incentive for participation. Participants were assured that all data obtained would be treated anonymously. They were informed that the purpose of the study was to investigate different teaching models, but no further details were provided concerning the difference between study groups. Each study participant used a pre-prepared sheet to provide data on group characteristics (baseline data) such as age (in years), sex (female/male), previous education in a healthcare-related or medical profession other than as a paramedic or nurse (ie, physiotherapy, dental assistant, etc), and completed medical clerkships.

**Materials**

Participants underwent a preliminary test of working memory performance and a self-efficacy rating prior to their skills training session. The revised version of the Wechsler Memory Scales,\textsuperscript{40,41} (German version\textsuperscript{42}) was designed to
assess attention/concentration and working memory as well as logical, visual, and verbal short- and long-term memory. It comprises 13 subtests that are combined to form five primary indexes with retest reliabilities ranging from $\alpha=0.80$ to $\alpha=0.88$. We employed the subtests assessing short-term and working memory. In the digit span subtest, individual digit spans are slowly read aloud to the participant and are to be repeated by the participant from memory (digit span forward). The second part of the task consists of repeating the digit span in a reversed order (digit span backward). The two task components allow conclusions to be drawn regarding short-term memory capacity (digit span forward) and working memory capacity (digit span backward). The self-efficacy ratings consisted of five items using Likert scale ratings concerning knowledge of anatomy, materials, and the different steps of and the ability to perform a central venous catheter (CVC) insertion on a manikin and on a patient. Students who had already placed a CVC were excluded from the study. Refusal to participate had no impact on later evaluations or other assessments.

The assessment of the skills-lab training was conducted through video ratings of participants’ performances. Recordings of trainees’ performances on the manikin were independently rated by two clinically experienced and trained video assessors who were blind to both the aim of the study and its design. Assessors were provided with checklists comprising 39 binary items that reflected the procedural steps of inserting a CVC.

Memory performance for the single steps of CVC was assessed by an incidental free recall test that was conducted the day after the experiment. Free recall tests represent a measure of active memory retrieval performance. Participants were called by telephone and asked to name as many single steps of the CVC as possible in the correct order. For scoring of the test, the same checklists were used that were provided for video assessors in order to allow for maximum comparability of the results. The number of correctly recalled steps as well as the number of steps recalled in the correct order served as dependent variables. Procedural steps were considered to be correctly recalled regardless of the exact medical terms. The free recall test lasted for approximately 10 minutes.

**Study design**

To implement the study, a total of 97 voluntary participants were randomly assigned to four groups, with two groups being larger due to the main interest in differential effects of Peyton’s Step 3. Each group was taught how to insert a CVC using a manikin, and the groups were instructed with a varying number of steps or a modification of Peyton’s Step 3, respectively (Figure 1):

- **Step-1 group** received Peyton’s Step 1 ($n=20$).
- **Step-2 group** received Peyton’s Steps 1 and 2 ($n=20$).
- **Step-3 group** received Peyton’s Steps 1, 2, and 3 ($n=29$).
- **Step-3mod group** (modified) received Peyton’s Steps 1 and 2, followed by a repetition of Step 2 ($n=28$).

Each instructional procedure of individual groups was followed by a video assessment of the first independent performance of the task. CVC insertion was chosen as the

learning goal as it is a routinely performed clinical skill. At the same time, it is a complex procedure with a high cognitive load for novices. Moreover, it is associated with a number of serious complications, estimated at a rate of about 5.3 per 1,000 catheter days. Improper application can have hazardous adverse effects on patients, which are expensive to treat, including hemotherax and air emboli, and can even result in death.

The presented study was conducted over a period of 6 weeks alongside the regular curriculum at the University of Heidelberg. In light of the above-described study design, the University of Heidelberg Ethics Committee waived requirements for an ethical approval procedure. Ethical principles were adhered to.

Skills training sessions
Skills training sessions for all groups were conducted with video instruction for Peyton’s Steps 1 and 2 in order to maximize the standardization of teaching. In the video for Peyton’s Step 1, the teacher performed the CVC insertion without further explanation, while in the video for Peyton’s Step 2, detailed instructions for all necessary action steps were provided. Step 3 (Step-3 group) was performed with a student–teacher ratio of 1:1, in accordance with approved skills training standards. After watching the videos for Peyton’s Steps 1 and 2, trainees were asked to instruct the teacher to perform the CVC insertion by memory once. The teacher followed these instructions, checking with the trainee if false commands were given. Step-3mod group watched the video for Peyton’s Step 2 twice. The teaching of CVC placement was performed under consideration of checklists based upon international standards.

Participants did not receive feedback after video recording to prevent biasing the results of the free recall test.

Skills-lab teachers
Step-3 group was trained by two clinically experienced consultants in internal medicine (age, 33 and 36 years), both of whom have considerable experience as certified skills-lab trainers.

Statistical analysis
Data are presented as mean ± standard deviation. Data were compared using Student’s t-tests (continuous interval data) and Mann–Whitney U tests (categorical ordinal data), respectively. The distributions of group characteristics pertaining to sex, previous education in a medical profession, and completed medical clerkships were compared using chi-square tests. A P-value <0.05 was considered to be statistically significant. Inter-rater reliability for the two video assessors was calculated using a Pearson’s correlation. The software package STATISTICA (StatSoft, Inc., Tulsa, OK, USA) was used for statistical analysis.

Results
There were no statistically significant differences between the four groups with regard to age (mean age of all participants 22.8±2.3 years), sex (57 female, 40 male), completed education in a medical profession (10/97 participants), completed medical clerkships (10/97 participants), preliminary memory tests, or self-efficacy ratings (Table 1).

Effectiveness of the training
Ratings by the independent video assessors are presented in Figure 2. Regarding checklist ratings, Step-2 Group (76.9%±15.0) showed a significantly superior first independent performance of CVC placement compared to Step-1 Group (58.2%±15.1) (P<0.0001), and Step-3 Group

Table 1 Group characteristics of different participating groups; age (mean, SD), sex (f/m), education in medical profession, completed medical clerkships, digit span forward and backward, and self-efficacy rating prior to skills training (five items using Likert scale ratings; 6=completely agree; 1=completely disagree); results of Mann–Whitney U test and chi-square tests

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Step-1 group (n=20)</th>
<th>Step-2 group (n=20)</th>
<th>Step-3 group (n=29)</th>
<th>Step-3mod group (n=28)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>23.0±2.5</td>
<td>23.0±3.0</td>
<td>22.7±1.8</td>
<td>22.6±2.5</td>
<td>ns*</td>
</tr>
<tr>
<td>Sex (f/m)</td>
<td>12/8</td>
<td>12/8</td>
<td>17/12</td>
<td>16/12</td>
<td>ns*</td>
</tr>
<tr>
<td>Education in medical professiona</td>
<td>6/20</td>
<td>7/20</td>
<td>3/29</td>
<td>10/28</td>
<td>ns*</td>
</tr>
<tr>
<td>Completed medical clerkshipsb</td>
<td>1/20</td>
<td>3/20</td>
<td>2/29</td>
<td>4/28</td>
<td>ns*</td>
</tr>
<tr>
<td>Digit span forward (WMS-R)</td>
<td>8.4±1.6</td>
<td>9.9±1.3</td>
<td>8.9±1.7</td>
<td>9.6±1.7</td>
<td>ns*</td>
</tr>
<tr>
<td>Digit span backward (WMS-R)</td>
<td>8.5±2.0</td>
<td>8.6±1.8</td>
<td>8.2±1.8</td>
<td>9.2±2.0</td>
<td>ns*</td>
</tr>
<tr>
<td>Self-efficacy rating prior to skills session (6–1)</td>
<td>1.05±0.6</td>
<td>1.29±0.7</td>
<td>1.29±0.8</td>
<td>1.19±0.7</td>
<td>ns*</td>
</tr>
</tbody>
</table>

Notes: aCompleted education as paramedic, medical secretary, nurse, or occupational therapist; bcompleted clerkships in surgery, internal medicine, anesthesia, or general medicine; *ANOVA results; *Chi-square test.

Abbreviations: SD, standard deviation; WMS-R, revised version of Wechsler Memory Scales; ns, not significant.
Peyton’s four-step approach

(85.4%±14.8) was superior to Step-2 Group (P<0.009), while Step-2 Group and Step-3mod Group (82.9%±14.8) did not differ (P=0.055) significantly. Step-3 Group and Step-3mod Group did not differ significantly (P=0.32). Effect sizes for mean differences between Step-2 Group and Step-3 Group (d=0.72; power =0.79) and between Step-2 Group and Step-3mod Group (d=0.57; power =0.61) show moderate to strong values.

Results of incidental free recall test
Results of the incidental free recall test 1 day after training (delayed recall) showed a similar pattern of results and even revealed a significant difference between Step-3 Group and Step-3mod Group, with a superior memory performance shown by Step-3 Group (P<0.022). Effect sizes for mean differences between Step-2 Group and Step-3 Group show strong values (d=1.07; power =0.98), small to moderate values between Step-2 Group and Step-3mod Group (d=0.41; power =0.40) and moderate to strong values between Step-3 Group and Step-3mod Group (d=0.63; power =0.76).

Inter-rater reliability for independent video raters
Standardized inter-rater reliability for independent video raters was 0.91 for binary checklist ratings.

Discussion
This study prospectively investigated the learning outcome of the different steps of Peyton’s four-step approach with a focus on the differential implications of Step 3 (“Trainee talks trainer through”), which was compared to a modified Step 3 – a repetition of Step 2 (“Talk the trainee through”). Training effectiveness was examined with respect to participants’ subsequent first performance of independent CVC insertion. Objective video ratings with binary checklists, which focused on the correct stepwise performance, showed a significant superiority of Peyton’s Step 3 over the mere repetition of Step 2. These findings were also confirmed on the level of memory performance for the single steps of CVC as assessed by an incidental free recall test (delayed recall), which additionally revealed a superior memory consolidation of the trained skill.

We assume that Peyton’s Step 3 represents the crucial instructional substep in Peyton’s four-step approach and therefore leads to a superior performance of clinical skills as compared to standard instruction. This was also demonstrated in a previous study, in which the authors were able to show that Peyton’s four-step approach leads to a superior and faster procedural performance when students execute their very first gastric tube insertion compared to a standard instruction following a “see one, do one” approach. The current study was able to identify that this advantage of Peyton’s four-step approach can be predominantly accredited to Peyton’s Step 3:

While a commented repetition of a demonstration of CVC application (Peyton’s Step 2) has been shown to lead to a superior performance compared to a first uncommented demonstration only (Peyton’s Step 1), Peyton’s Step 3 (following Step 1 and Step 2) leads to a further advantage compared to...
a mere (third) repetition of a CVC application demonstration (here Step 3 modified), which is reflected in checklist and memory performance results. In other words, although both skills observation and motor imagery support motor skills acquisition, the combination of motor imagery and skills performance as inherent in Peyton’s Step 3 seems to be superior to skills observation only.

Our study is in line with previous findings from a randomized controlled trial comparing mental vs practical training. In that study, surgeons undergoing laparoscopic training were divided into three groups that received mental training, practical training, or no additional training, respectively. The mental training group outperformed both the practical training group and the control group and showed significantly higher task-specific checklist scores. These results support our findings of a specific importance of Peyton’s Step 3 in procedural skills acquisition. Motor imagery is a well-established training method in various sports and is used by athletes and coaches to increase performance. Furthermore, functional neuroimaging studies (fMRI) have demonstrated specific neural activation in the areas of the primary motor cortex and the cerebellum when subjects imagined body movements without actual muscular activity.

Regarding memory performance for single steps of the CVC procedure as assessed by an incidental free recall test, results showed that memory performance after an instruction following Peyton’s four-step approach (including the regular Step 3) is superior to a modified version of Peyton’s four-step approach including CVC demonstrations only. These findings are in line with the theory of retrieval practice, as Step 3 can be regarded as a correlate of retrieval practice. In a variety of studies, retrieval practice has been shown to greater enhance memory than studying alone. According to the levels of processing framework, self-referential encoding leads to a more elaborate processing of relevant information to be learnt compared, for example, to semantic or emotional processing operations. Indeed, in a variety of clinical patient studies, researchers were able to show that information encoded with a reference to the patients themselves resulted in a “deeper” processing and therefore superior encoding. In line with these assumptions, we expected Peyton’s Step 3, which includes motor imagery, to lead to a deeper encoding of single steps when compared to mere performance observation. This deeper encoding might have facilitated the retrieval of single procedural steps and therefore might have led to a reduced cognitive load, meaning that more cognitive resources were available for the procedural performance itself. In a study comparing laparoscopic training with either explicit or implicit motor learning, Zhu et al were able to show that implicit learning reduced EEG coactivation between verbal-analytic and motor planning regions, suggesting that verbal-analytic processes were less involved in laparoscopic performance. They conclude that implicit motor learning leads to the same level of technical proficiency as explicit motor learning but with greater neural efficiency. Deeper encoding of single steps, reflected in an eased recall in our study, might therefore prevent enhanced verbal-analytic processes that interfere with procedural performance. Superior declarative memory of single CVC steps as a prerequisite for CVC skills performance might also lead to a higher quality of skills execution in future procedural performances. Future research should assess whether self-referential motor imagery leads to a better skills performance than motor imagery that instructs the skills teacher. Previous research has shown that self-referential encoding leads to deeper encoding based on Levels of Processing Theory.

Several limitations of our study should be mentioned. First of all, group sizes were rather small. However, potential confounders were distributed equally across groups, and a previous study examining Peyton’s four-step approach used comparable group sizes. We did not examine performance of CVC skills on a real patient as this would have called for the trainer to intervene during the performance if the trainee made mistakes and would have consequently led to difficulties in objectively rating the performance. We did not evaluate objective competencies in CVC placement prior to the training, as this was a predefined exclusion criterion for participants in light of our aim to evaluate participants’ very first independent performance of the skill. Furthermore, a student-to-teacher ratio of 1:1 does not reflect normal skills-lab training conditions. However, our aim was to observe emerging effects as purely and in an as isolated a manner as possible in a research setting. Group 3 was the only group that had direct contact with a teacher and this could have influenced their performance as the engagement with a more senior person could have implications for a student’s motivation to do well. On the other hand, this could as well have led to an intimidation of trainees and thus negatively influence their performance. The Step-1 group and the Step-2 group did not have the chance to inquire upon possible ambiguities as they only received video training. Furthermore, results in the delayed recall test may not correlate with actual performance of the task; hence, all conclusions must be interpreted with this in mind. Finally, we used videos for Peyton’s Steps 1 and 2, despite the fact that the original approach calls for a direct trainer-trainee interaction. We purposely decided to do...
so in order to ensure that the training was completely equal for all students regarding Steps 1 and 2.

In summary, our study identified Peyton’s Step 3 as being the most crucial part of Peyton’s four-step approach, contributing significantly more to learning success than previous steps and reaching beyond the benefit of a mere repetition of skills demonstration. It remains unclear whether our findings are transferable to other settings such as the workplace. In order to adapt to workplace conditions, researchers would need to assess trainees’ performance in a longitudinal way, which means that the respective skill should be evaluated in different situations with several patients. Furthermore, in order to meet concerns of time management, it may be helpful only to train substeps of a skill so that, for instance, a patient does not need to hold still for a long time while a CVC is placed. Further studies should therefore evaluate the effectiveness of this instructional approach with respect to these considerations.

Disclosure
The authors report no conflicts of interest in this work.

References