Preliminary evaluation of the efficacy of an intervention incorporating precision teaching to train procedural skills among final cycle medical students

Sinéad Lydon, Nadine Burns, Olive Healy, Paul O’Connor, Bronwyn Reid McDermott, Dara Byrne

ABSTRACT

Introduction A lack of preparedness for practice has been observed among new medical graduates. Simulation technology may offer one means of producing competency. This paper describes the application of a simulation-based intervention incorporating precision teaching (PT), a method of defining target skills, assessing individual progress and guiding instructional decisions, which is used to monitor learning and the development of behavioural fluency in other domains, to procedural skills training. Behavioural fluency refers to accurate and rapid responding that does not deteriorate with time, is resistant to distraction and can be adapted into new, more complex responses.

Method This study used a between-groups design to evaluate the efficacy of a simulation-based intervention incorporating PT for teaching venepuncture among 11 medical students. The intervention consisted of timed learning trials during which participants carried out the skill in pairs and received corrective feedback. Two control groups of 11 untrained medical students and 11 junior doctors were also included in the study.

Results Intervention group participants required an average of five trials and 21.9 min to reach the criterion for fluency. The intervention group demonstrated significantly higher accuracy in venepuncture performance than either control group. Improvements persisted over time, did not deteriorate during distraction, generalised to performance with patients and performance of an untargeted skill also improved.

Conclusions The outcomes of this preliminary study support the application of PT within medical education. The implications of these data for clinical and procedural skills training are explored and suggestions are made for further research.

INTRODUCTION

Research has consistently shown that new medical graduates feel underprepared for clinical practice. Similarly, educational supervisors perceive new medical graduates to lack the procedural skills competency necessary for successful practice. This ill-preparedness is not trivial; patient outcomes worsen, and efficiency of care decreases, during the first weeks of clinical practice by a new intake of junior doctors.

The lack of skills competency among new medical graduates has led to the recognition that traditional approaches to skills training are insufficient for developing competency and that there is a need for the evaluation of novel skills training approaches. Simulation-based Medical Education (SBME) has emerged as a popular solution to these problems, allowing students to perform procedures, make mistakes and receive feedback in an environment that replicates real life. However, although SBME has been identified as an ethical and effective way to teach clinical and procedural skills, there is a lack of consensus regarding best practice for employing simulation technology to teach skills and the frequent lack of consideration educational framework, or instructional design, may be criticised.

Outside of medical education, the search for expertise in particular skill repertoires has led to a growing interest in behavioural fluency. Behavioural fluency has been described as “that combination of accuracy plus speed of responding that enables competent individuals to function efficiently and effectively in their natural environment”. It differs from mastery, which refers to high levels of accuracy alone, without any consideration of the pacing or flow of performance. Behaviour that is learnt to fluency is Retained for longer periods of time, individuals can Endure performing the skill for longer durations, individuals can Adapt the behaviour learnt by performing it as part of a new, more complex, compound behaviour, individuals can Perform the skill at a rate that makes it functional and their performance is Stable regardless of distraction (REAPS). Precision teaching (PT) has been defined as ‘a system for defining instructional targets, monitoring daily performance, and organising and presenting performance data in a uniform manner to facilitate timely and effective instructional decisions’. In this way, PT does not dictate instruction but rather offers a means of evaluating any educational programme implemented. The careful definition and monitoring of behaviour using PT allows for the ascertainment of individual differences in performance and the development of behavioural fluency. Components of an intervention incorporating PT typically include: (1) establishing a criterion rate for performance of the targeted skill (ie, a time in which an expert can comfortably complete the behaviour with total accuracy); (2) providing opportunities to practice the behaviour in the context of timed trials and delivering feedback regarding errors and speed of performance (referred to as frequency building); (3) charting of the data from trials using the standard celeration chart, and; (4) changing teaching tactics when the standard celeration chart indicates that performance is not improving.
incorporating PT typically include many of the elements of mastery learning\textsuperscript{6} and deliberate practice\textsuperscript{16} including baseline testing, the targeting of a clearly defined behaviour, the provision of corrective feedback and the continuation of practice until a passing standard is achieved.\textsuperscript{17, 18} However, the crucial difference between such interventions and interventions including PT is the focus on the speed of performance, as the ‘flow’ of behaviour is the hallmark of fluency and high accuracy alone is insufficient for producing the outcomes of behaviour fluency (ie, REAPS).\textsuperscript{13, 14} PT has been successfully employed within interventions for a range of target skills across a variety of educational contexts.\textsuperscript{14, 19, 20}

The current study aimed to explore the efficacy of an intervention comprising of frequency building, peer tutoring and PT to improve performance of venepuncture by medical students. The goals were to: identify whether behavioural fluency could be established in a procedural skill, whether there was a difference between the effectiveness of the intervention delivered and ‘training as usual’, and whether there was a difference in the level of performance between students who received the intervention and qualified junior doctors.

**METHOD**

**Experimental design**

A between-groups design was used with three groups: (1) an intervention group of fourth year medical students that received the intervention; (2) a control group of fourth year medical students that did not receive intervention (control group one) and (3) a control group of junior doctors who did not receive intervention (control group two). As part of this design, similar to that employed by a recent study of the effects of deliberate practice on blood product prescribing,\textsuperscript{21} control group participants were tested just once, at the conclusion of the intervention phase. The examination of performance across the groups at this time was deemed to be most appropriate as it allowed for the best comparison of the intervention’s effects to education as usual, or usual practice, and allows us to infer whether the intervention offered improvements in performance beyond what could be expected from usual teaching or clinical conditions. The performance of venepuncture was compared among these groups and the performance of peripheral intravenous cannulation (PIC) during the post-testing phase was compared among the two medical student groups. The rationale for including PIC was to identify whether the learning had transferred to another needle-based task. Figure 1 provides an overview of the activities of each group.

The accuracy of intervention and control participants’ performance was compared. The rate of performance was assessed in the intervention group throughout the study, as the aim was to develop behavioural fluency (ie, accurate and well-paced performance). However, it was considered more appropriate to compare the groups on accuracy than speed for two reasons. First, the absence of behavioural fluency in the control groups is evidenced by the lack of accuracy, as without total accuracy one cannot be fluent regardless of performance speed. Second, duration data do not accurately capture performance differences between the groups as some individuals could complete the task more quickly by omitting steps.

**Ethical approval**

Ethical approval was obtained from the research ethics committee of the participating university.

**Targeted behaviour**

Venepuncture was targeted for improvement in the current study. Venepuncture has been identified as an essential skill for medical practice.\textsuperscript{22, 23}

---

**Figure 1** Flow diagram depicting the various study phases for the intervention and control groups.

Participants

Recruitment

Convenience sampling was used to recruit 33 participants from a large public tertiary teaching hospital with an attached medical school in the Republic of Ireland.

Intervention group

This group consisted of 11 fourth year medical students. Six were men and five were women. Mean age was 23 years (SD=1.1; range 21–25). All participants reported some previous experience in carrying out venepuncture, either through simulation or with patients.

Control group one (fourth year Controls)

This group consisted of 11 fourth year medical students. Four were men and seven were women. Mean age was 22.9 years (SD=1.3; range 21–25). All participants had previously carried out venepuncture on patients.

Control group two (junior doctors)

Eleven junior doctors were recruited. Four were men and seven were women. Mean age was 24.7 years (SD=0.66; range 24–26). All participants had completed 11 months of clinical practice and were experienced in taking blood from patients.

Setting

The study was carried out at a simulation laboratory adjoining the participating medical school and hospital. Intervention group participants also attended the hospital’s phlebotomy outpatient clinics.

Training materials

In order to promote the likelihood of generalisation of learning to the natural environment, this study used the materials supplied for performing venepuncture in the hospital. Reusable simulation training pads that replicated the antecubital fossa were filled with artificial blood and used. Other necessary equipment included hand sanitiser gel, a tray, a sharps disposal bucket, gloves, one needle, one disposable tourniquet, one 2% chlorhexidine and alcohol wipe, a cotton ball, a band aid and four vacutainer bottles. The vacutainer bottles used were those most frequently used in the hospital. In addition to these materials, each observer used a data collection sheet, a timer and a copy of the task analysis when assessing performance.

Procedure

Preparatory phase

A task analysis, a list of 20 necessary steps deemed essential for the satisfactory performance of venepuncture, was developed (see online supplementary material 1). The use of PT requires the ‘pinpointing’ of the targeted behaviour. Effective pinpointing involves the describing of a targeted behaviour in terms of its observable physical movements. Effective pinpointing benefits the teacher by facilitating the accurate assessment of performance (and performance changes), and also benefits the learner by making it completely clear what behaviour is required. In this study, behaviour was pinpointed by operationally defining each step of venepuncture to provide a clear, simple explanation of the physical movements and materials involved. The breakdown of venepuncture into discrete steps for the task analysis was conducted with the help of a phlebotomist, who had over 25 years of experience teaching venepuncture and using the university and hospital protocols concerning best practice for performing venepuncture. Face validity was ensured by having two senior doctors review the document.

The phlebotomist involved in developing the task analysis was asked to carry out venepuncture using the simulation materials three times in order to establish an ‘expert’ or fluent rate of performance of the skill. The median time of these three performances, each completed with 100% accuracy, was used to establish a fluency criterion. The criterion rate was set at 4:48 min and below which required participants to complete the task without exceeding the expert’s median duration by more than 10%.

Intervention phase

This phase was completed by the 11 intervention group participants only (see figure 1).

Baseline

Participants attended the simulation laboratory for baseline testing. Participants were asked to carry out venepuncture to the best of their ability using the simulation materials and were informed that the accuracy and speed of their performance would be assessed. Following completion of the baseline probes, participants were presented with the task analysis but were not provided with corrective feedback.

Frequency building, precision teaching and peer tutoring intervention

Participants were invited to attend a series of brief teaching sessions, in the simulation laboratory, in which the intervention was delivered. Participants were introduced to the components of the intervention, the format of teaching, and the criterion for achieving behavioural fluency (ie, 100% accuracy within 4:48 across two consecutive trials). Participants were grouped into random pairs and were given copies of the task analysis, timers and scoring sheets to score and time their fellow students’ performances. Owing to the uneven number of participants, there was one group of three in each session. Participants within this group rotated and took turns to ensure that each participant had the opportunity to act in each role (ie, performer and assessor).

For each frequency building trial (ie, each separate performance of venepuncture) during the intervention phase, one member of each pair acted as the peer ‘tutor’ for the participant carrying out the trial. Once the trial was complete, the peer tutor gave specific feedback, using the task analysis, on the steps completed correctly and incorrectly and the trial duration. A trained observer was stationed with each pair but only provided input if there was a lack of consensus regarding the accuracy of completion of any step or if the observer had useful feedback beyond that offered by the peer tutor. The scoring sheet provided two assessment metrics for each step of the task analysis: a binary indication (ie, performed accurately, not performed/not performed accurately) of the correctness of performance and a free text space. Observers used the free text data to inform the feedback they provided to the learner regarding the accuracy of their performance of each step (eg, angle of needle insertion, completeness of handwashing). The opportunity to practice incorrect steps was also provided at the end of each trial following the provision of feedback. Observers monitored participants’ data using the standard celeration chart in order to identify any potential need for supplemental instructional tactics if participants were not progressing. However, participants progressed quickly and no such need was identified.
**Interobserver agreement**

While participants collected data on their training partner(s), trained observers simultaneously but independently scored the performance during 92.6% (n=50) of training trials. Interobserver agreement (IOA) was calculated by dividing the total number of opportunities for agreement between two observers by the number of agreements observed. The level of agreement between the two sets of observers was found to be 95.6%.

**Post-testing phase**

A number of post-tests were conducted 8 weeks after the cessation of intervention in order to assess for the characteristics of behavioural fluency.

**Retention**

This post-test was conducted in the same manner as baseline testing.

**Stability**

The stability of the behaviour, or the persistence of the accuracy and speed in the presence of distraction, was assessed. Participants were asked to perform the task while a video of a hospital emergency was playing loudly on a nearby screen, providing an audible and visual distraction. Further, a confederate entered and moved around the room in order to provide an additional source of distraction.

**Generalisation (novel environment)**

Participants were asked to perform venepuncture on patients presenting at the hospital’s phlebotomy outpatient clinic. This post-test was intended to determine whether transfer of learning to the clinical environment was observed.

**Generalisation (novel skill)**

Participants were asked to carry out PIC in the simulation laboratory. This post-test was conducted to assess if skills training in venepuncture had extended to another similar but untargeted needle-based skill (see online supplementary material 1 for the PIC task analysis used).

**Control group testing**

Participants from the control groups were invited to visit the simulation laboratory individually. They were asked to carry out venepuncture in test conditions equivalent to those used for baseline testing in the intervention group. Data were collected on accuracy and duration but no feedback was provided.

**Statistical analysis**

A one-way repeated measures analysis of variance (ANOVA) was used to compare the intervention group participants’ performance across the various intervention phases (ie, baseline, final intervention trial and post-tests). Post hoc testing was conducted using dependent t-tests.

The comparison of the intervention group’s performance at the conclusion of training with that of the two control groups’ performances was conducted using a one-way between participants ANOVA. Post hoc Tukey tests were conducted. The comparison of the intervention group and control group one’s performance of PIC was conducted via independent samples t-test.

**RESULTS**

**Intervention group performance**

The intervention group achieved behavioural fluency in a mean of 4.9 trials (SD=1.1; range: 3–7), with a mean duration of 4.24 min per trial. The mean total training time required to achieve fluency was 21.93 min (SD=6.59; range: 12.23–36.38).

The percentage of participants completing each step of the task analysis correctly across all of the study phases is presented in Table 1. A one-way repeated measures ANOVA was conducted to compare intervention group participants’ performance of the skill during the baseline, on achieving behavioural fluency (ie, the last training trial), the retention post-test, the stability post-test and the generalisation to novel environment post-test. This test indicated a significant difference between performance across conditions, F(1,9)=4323.6, p<0.001. In order to determine where the differences lay, post hoc dependent t-tests were performed. A Bonferroni correction was applied to control for these multiple comparisons and α was set at 0.005. Post hoc testing revealed that accuracy at baseline (M=43.6, SD=16.6) was significantly lower than accuracy at the final intervention trial (M=100; SD=0), the retention trial (M=99.5; SD=1.5), the stability trial (M=99.5; SD=1.5) and the generalisation to novel environment trial (M=93; SD=15.3), all p’s<0.001. No significant differences in performance accuracy emerged from the additional comparisons conducted, all p’s>0.005.

**Comparison with control groups**

Table 1 displays the percentage of participants in each control group who completed the steps of the task analysis accurately. A one-way between participants ANOVA was conducted to compare the post-training accuracy of the intervention group to that of control group participants tested at the same timepoint. There was a significant difference in accuracy between the three groups, F(2,30)=74, p<0.001. Post-hoc comparisons using the Tukey test indicated the mean accuracy for the intervention group (M=100, SD=.00) was significantly higher than control group one (M=64.1, SD=7.4) and control group two (M=73.2, SD=10.1).

Mean performance accuracy of the intervention group on the skill of PIC was also compared to that of control group one. Results showed that the intervention group performed PIC with significantly more accuracy (M=86.1, SD=6) than did control participants (M=48.1, SD=15.3), t(13.04)=7.65, p<0.001.

**DISCUSSION**

This study examined the effects of a simulation-based intervention incorporating frequency building, peer tutoring and PT on medical students’ ability to perform venepuncture. PT allows for the assessment of individual performances and the ascertainment that instruction has resulted in behavioural fluency or behaviour that is accurate and fluid. When compared to traditional educational processes in other domains, the use of PT alongside learning programmes has been shown to result in higher student achievement.18 19 20

These data support a growing body of research suggesting that SBME leads to better skill proficiency than traditional training approaches,6 24 with the intervention group outperforming both their peers and junior doctors. These results compliment prior research that has shown that greater clinical experience does not engender the same level of competency as achievable through SBME.21 25 The use of interventions incorporating PT to establish behavioural fluency in procedural skills would minimise the potential risk to patients and promote best standards.
for education and practice by ensuring learners receive training in an individualised yet systematic way to ensure that all become fluent. In our study, medical students were able to accurately assess their peers’ performance. Therefore, the delivery of training in this manner should not require additional staff beyond those involved in typical training activities. In other domains, the addition of PT to educational programmes has also been found to offer a highly effective, but low resource-intensive, approach to education. Therefore, the delivery of training in a manner that allows learners to assess their peers’ performance is a valuable tool in education and practice by ensuring learners receive training in an individualised yet systematic way to ensure that all become fluent.

Results from the post-tests indicated that participants who received the intervention achieved behavioural fluency, completing the target behaviour ‘accurately’ and ‘without hesitation’ post-intervention. Participants retained the skill over 8 weeks, displayed stability when performing the skill in the presence of distraction, remained highly accurate when they transferred their learning to real patients and their learning appeared to generalise to an untrained needle-based skill. Being able to fluently perform basic skills ensures clinical competence and decreased errors leading to greater patient safety. The generalisation of learning to a novel, untargeted skill in the current study is a particularly positive finding. This generalisation suggests that becoming fluent in one procedural skill may have a positive impact on other skills. This finding could allow for a reduction in the time taken to achieve fluency in other skills resulting in training that is both time-efficient and cost-efficient.

The current study has a number of important strengths including the use of two control groups in order to best explore outcomes. The use of these control groups allowed for the demonstration of the superiority of SBME incorporating PT over both standard skills training and clinical experience. The demonstration of the generalisation of learning to the clinical environment is another strength. There is a recognised lack of research supporting the transfer of SBME learning to clinical situations. The application of PT which is supported by a strong theoretical background and much empirical data in other educational settings is also a positive element of this study.

However, there were also limitations. The small sample size and use of convenience sampling must be noted. Participants were not randomised to the experimental groups due to practical considerations; students in the intervention group had to be available to attend training sessions at specified times and the schedules of some students precluded this. Further, the non-testing of control group participants at baseline, as well as post-intervention, may be considered an additional methodological weakness. The stability assessment task used may also be criticised. We opted for a task that included a number of perceived naturalistic distractors. However, the distraction potential of the stimuli added was distracting but future research may wish to use established distraction tasks or to validate the distraction task used. Further, healthcare practitioners may believe that particular elements of the task analysis are not required or should be carried out differently. However, the task analysis was compiled in accordance with hospital and medical school protocols and reviewed for face validity by subject matter experts. Therefore, it is ‘one correct way’ to perform venepuncture and acceptable within the clinical environment. Further research evaluating the effects of PT could be conducted using larger sample sizes, a more rigorous experimental design and could extend training to other, more complex clinical and procedural skills.

---

Table 1 Percentage of participants performing each step of the venepuncture task analysis correctly across groups and study phases (n=11 for control group 1, control group 2 and the intervention group, total n=33)

<table>
<thead>
<tr>
<th>Step</th>
<th>Intervention group</th>
<th>Control group 1 (%)</th>
<th>Control group 2 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Baseline (%)</td>
<td>Final training trial (%)</td>
<td>Retention Post-test (%)</td>
</tr>
<tr>
<td>1. Hand hygiene with gel</td>
<td>45.5</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>2. Palpates suitable vein</td>
<td>9.1</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>3. Hand hygiene with gel</td>
<td>0</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>4. Applies gloves</td>
<td>63.6</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>5. Applies tourniquet</td>
<td>36.4</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>6. Sani cloth</td>
<td>27.3</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>7. Allows to dry</td>
<td>9.1</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>8. Opens phlebotomy needle</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>9. Inserts needle, bevel up</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>10. If no blood enters bottle…</td>
<td>27.3</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>11. Fills bottle in correct sequence and agitates</td>
<td>0</td>
<td>100</td>
<td>90.9</td>
</tr>
<tr>
<td>12. Releases tourniquet</td>
<td>63.6</td>
<td>100</td>
<td>90.9</td>
</tr>
<tr>
<td>13. Removes needle and applies cotton wool</td>
<td>63.6</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>14. Closes safety on needle</td>
<td>45.5</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>15. Disposes of sharps</td>
<td>72.7</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>16. Labels at bedside</td>
<td>9.1</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>17. Checks patient’s arm</td>
<td>45.5</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>18. Applies bandaid</td>
<td>81.8</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>19. Clean up</td>
<td>9.1</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>20. Hand hygiene with gel</td>
<td>9.1</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

All fourth year participants (intervention and control group) were on placement in general practice settings and attending teaching sessions during the time between baseline and conclusion of the intervention.
CONCLUSION
The use of PT to monitor the impact of an intervention on performance allows for the determination that an intervention has resulted in behavioural fluency, or behaviour that is automatic or effortless and persists over time and in challenging conditions. This preliminary study demonstrates that PT can be readily incorporated within a SBME intervention to teach a procedural skill and that only a short duration of training, and minimal staff input, is required to produce competence exceeding that of practicing doctors. Our data, along with those from other recent applications of behavioural measurement and training strategies within medical education, suggest a great potential for applied behaviour analytic techniques to contribute positively to medical education that must be explored by future research.

Acknowledgements The authors wish to thank Dr Aoife McTiernan, Ms Mary McLoughlin and Mr Tommie Mellett for their assistance with this study.

Competing interests None declared.

Ethics approval NUI Galway’s Research Ethics Committee.

Data sharing statement All data collected as part of this study are described in the submitted manuscripts. We are happy to consider sharing data with anyone who should request this.

REFERENCES